Working Paper Series
Department of Economics
University of Verona

A Gini measure of Economic Segregation

Umut Turk, John Östh

WP Number: 14 August 2017

ISSN: 2036-2919 (paper), 2036-4679 (online)

A Gini measure of Economic Segregation

Umut TÜRK*& John ÖSTH†

umut.turk@univr.it

Abstract

This paper proposes a new measure of spatial segregation by income that uses the Gini index as the basis of measurement. Gini Index of spatial segregation (GSS) is a ratio of two Gini indices comparing inequality between neighbourhoods to inequality between individuals at the macro area where neighbourhoods are nested. Unlike other measures of income segregation found in literature, the index uses individualized neighbourhoods. Depending on the population density and proximity between individuals, an individualized neighbourhood is defined both as an area constituted within a radius or as a population-count method around an individual geo-location. The GSS is suitable for the measurement of residential segregation by any continuous variable. It is sensitive to spatial configuration of areas, easy to compute and interpret and suitable for the comparative studies of segregation over time and across different contexts. An empirical application of the index is illustrated using data from Sweden covering the entire population in the years 1994, 2004 and 2014. We show how the definition and scale of the neighbourhood dramatically affect the measures of economic segregation.

*Economics Department, University of Verona, Verona Italy

†Department of Social and Economic Geography, Uppsala University, Uppsala Sweden

1 Introduction

The measurement of residential segregation by income (hereinafter Economic Segregation) has attracted relatively less attention than measurement of residential segregation by race or occupation. Economic and racial segregation share many factors in common: both are distinctively spatial phenomena, may occur from similar dynamics and are often empirically entangled (Reardon, 2011). Whereas the literature studying income segregation faces the challenge of measuring segregation along a continuous dimension, hence it cannot easily borrow indices from the racial segregation literature. Racial segregation refers to the uneven distribution of people belonging to different groups across physical space, while economic segregation amounts to quantify the income homogeneity or diversity in the areas of residents.

The efforts to identify segregation in urban spaces is not new to literature. Previous studies by sociologist, economists and geographers theorized the distinctive distribution of social classes in cities. The patterns of segregation among people by ethnicity, race, social class and among institutions, roads and variety of economic activities would be observed in cities growing radially from a core as rings outwards (The Concentric Zone model by Burgess (1928)) or as star-shaped, sector base (The Second Theory Model by Hoyt (1939)) or in a multiple core fashion, where the number and size of cores (nuclei) vary highly for different cities prevailing from historical development, geography and culture (Multi-core Model by Harris & Ullman (1945)). Most residential segregation studies use these theories as a basis to understand urban structure. The similar patterns of residential segregation has been confirmed repeatedly. The effects and causes of the segregation have been shown to prevail from households sorting across neighbourhoods with differential public goods/services that are excludable for location (Tiebout, 1956; Epple et al., 1984). Similarly, by the preferences of neighbourhood racial composition (Schelling, 1969), by education and income (Jargowsky, 1996), by exogenous factors such as changes in spatial distribution of opportunities and due shifts from manufacturing activities to service-oriented economies (Morenoff & Tienda, 1997) and by demographic changes: female participation in the market, aging population thus changes in demographic composition of cities (Wyly et al., 1998). Moreover, the effects of segregation have been shown to be related to inequality in growth (Burgers & Musterd, 2002; Reardon & Bischoff, 2011) and the distribution of top 1% income in the space (Essletzbichler, 2015). However, the scale in which these factors become evident has started to attract scholarly attention only recently.

To quantify the relative clusters of people by socioeconomic characteristics, any measurement has to aggregate them into some spatial unit so called "neighbourhood". The measurements are likely to vary depending on the definition of neighbourhood chosen. Especially, when relied on some predefined administrative unit such as census tracts or municipalities findings can be erroneous. This is what has become known as the modifiable areal unit problem (MAUP) (Openshaw, 1984; Wong, 2004). The MAUP occurs both with the scale problem, where the same data portrays different spatial patterns for its varying levels of aggregates,and with the zoning problem, where altering the grouping schemes produce different results even if the units are of the same scale. In particular to racial segregation analysis, the scale problem has been recognized and addressed in several ways (Wong, 1993, 1999, 2005; Reardon et al., 2008, 2009). Since the residential segregation by definition relates to clusters of people, the way in which the geography is handled becomes not only a statistical issue but also a crucial strategy to study the effects and causes of segregation. A way to address this issue is to construct scalable egocentric local environments. Depending on the definition of neighbourhood either a set of varying radius (see Lee et al., 2008, for racial segregation) or varying population sizes around an individual location so called k nearest neighbours (knn) used (see Östh et al., 2015, for interaction among racial groups).

The aim of the present work is to construct an index to perform a comprehensive investigation of residential segregation by income with a particular attention to the scale problem. The proposed index compares the inequality between individualized-neighbourhoods to the inequality between individuals at the macro area where neighbourhoods are nested. It is flexible in deciding the definition and size of neighbourhoods.

Section 2 recalls some relevant tools used in income segregation measurement. Section 3 presents the new GSS index. Section 4 provides an empirical application to Swedish context. Neighbourhoods are considered both as the area constituted within a radius that is drawn around each individual location and as the nearest population-count approach using the information of residential coordinates. Observing the diverse patterns in segregation measurements in response to the definition of neighbourhood chosen, we propose a variant of k nearest neighbour algorithm that makes use of spatial weights matrix (see Getis, 2009). In return, this new approach communicates both the spatial distribution of individuals and the population density constituted in each neighbourhood. The paper shows how the definition and scale of the neighbourhood influence the measures of economic segregation and the use of individual neighbourhoods permits managing the related weakness of the traditional tools and to obtain robust results.

2 The Gini indices of segregation

Despite the problems associated with the dissimilarity index, it remains the most diffused index of segregation. It measures the degree to which the minority proportions of areal units differ from the minority proportion of the whole area in absolute terms (James & Taeuber, 1982). Hence, the dissimilarity index is designed to measure the unevenness in the distribution of two population groups. The fundamental issue associated with the index is that it is sensitive to the share of minority population in different spatial units and to the size of the overall areal unit (Cortese et al., 1976). On the other hand, it is insensitive to the reallocation of minority groups among areal units where minority proportion is less or higher than the overall area's minority proportion (James & Taeuber, 1982). Even so, the dissimilarity index is widely used for the measurement of economic segregation in a similar manner to racial segregation. Typically, the index computes the uneven distribution of two population groups defined under and above to a given income threshold. However, this approach discards considerable amount of information hidden underneath income distributions (Abramson et al., 1995; Massey et al., 2003, see).

There are many other alternative indices found in literature: the index of exposure, relative concentration, absolute centralisation and spatial proximity (Massey & Denton, 1988), nearly all inequality indices can measure dichotomous /categorical segregation (Kim & Jargowsky, 2005). These include entropy and Atkinson indices for analyses of evenness in distributions.

In the present paper we focus on the Gini segregation indices. In its original form, the Lorenz curve is a representation of sorted cumulative percentage of total income as a function of cumulative percentage of total households (Lorenz, 1905). Whereas, the Gini coefficient is a measure of the area between Lorenz curve and the line of perfect equality, normalized by the total areal under the 45 degree line. As for the measurement of racial segregation, the index performs similar to the dissimilarity index, where the Gini coefficient represents the area between Lorenz curve normalized by the total area under 45 degree line for the minority populations weighted across all pairs of areal units (Massey & Denton, 1988). It takes a maximum value 1 when the minority and majority members of the society are perfectly segregated and 0 for no segregation. This form of the Gini is limited to measure the segreation along two population groups only. Reardon & Firebaugh (2002) proposed extensions of six segregation indices measuring multi-group segregation, including the Gini. As a function of the disproportionality in group proportions across organizational units, the index is interpreted as the weighted sum of the weighted average absolute difference in group proportions between all possible pairs of units over multigroups. Kim & Jargowsky (2005) developed a version of the Gini segregation index that accounts for continuous variables. By this extension the Gini becomes suitable for the measurement of income segregation, where the Gini for income disproportionality among neighbourhoods is normalized by the individual level Gini.

The forms of the Gini index listed above do not account for the spatial configuration of areas, therefore, are subject to the problems associated with MAUP. Also to "checkerborad phenomenon" occuring when an index is insensitive to spatial proximity between areas(White, 1983). A spatial version of the Gini is proposed by Dawkins (2004), it measures racial segregation given the spatial proximity of neighbourhoods. The decomposition of the index produces within and between components and also a residual term that captures the correlation of neighbourhood's own position and the position of its neighbours when ranked with some proximity among neighbourhoods. Extending the standardized spatial Gini index , Dawkins (2007) proposed spatial ordering index calculated from either a nearest neighbour or a monocentric spatial ordering of neighbourhood per capita income and the Gini index of between-neighbourhood income segregation. The index represents a ratio of two covariances where numerator is the covariance between neighbourhood aggregate income and spatial reranking of neighbourhood whereas denominator is the covariance between neighbourhood aggregate income and the average ranking of neighbourhood. However, the index does not address the scale issue. Table 1 shows these Gini segregation indices with their corresponding papers.

3 The GSS Index

In this section we introduce the Gini index of spatial segregation (GSS). Given a population of N individuals, let y_i be the income of individual i and μ_{is} be the average income in individual i's neighbourhood. The latter can be either radii-based (considering people comprised within a circle of a given radius drawn around individual i) or a count of the k nearest neighbours around i's location. Therefore, the shape (s) of the neighbourhood varies for the definition chosen and the size (n_{is}) can be set to meet various scales of geography. Note that for any population size $n_{is} \neq N$, μ_s will differ from μ.

The GSS is a weighted measure of income homogeneity/diversity in the areas of residents. It is defined as the ratio of two Gini indices as follows:

$$GSS(y,n) = \frac{\frac{1}{N^2\mu_s}\sum_i\sum_j|\mu_{is}-\mu_{js}|}{\frac{1}{N^2\mu}\sum_i\sum_j|y_i-y_j|} \tag{1}$$

where

$$\mu_{is} = \frac{\sum_{j\in s} y_j}{n_{is}} \tag{2}$$

The GSS index is the ratio of the between neighbourhoods inequality I_B to the individual level inequality I_G. It takes a minimum of 0 (no segregation) in two case scenarios: if the numerator is zero thus the between spatial inequality is zero or when the size of the neighbourhood is equal to the size of the whole study area: $n_{is} = N$. While the index takes maximum value 1 (perfect segregation) if the distribution of individualized-neighbourhood average incomes is identical to that of individual incomes thus when $I_B = I_G$ or when the size of the neighbourhood $n_{is} = 1$ every neighbourhood consists of one individual only.

The GSS measures the extent to which neighbourhoods differ from each other in terms of income, without concerning the fairness in income distributions. In a given area, neighbourhoods might be populated solely by lower income groups. In such a case the GSS shows lower values. This implies a lower residential segregation by income for this particular area.

The index has several advantages. First, in contrast to Kim & Jargowsky (2005) the index is sensitive to the spatial configuration of neighbourhoods so that it overcomes the checkerboard problem. Second, it does not require the use of predefined administrative units, hence it is not subject to the MAUP. Finally as a ratio of two Gini indices GSS preserves several properties of the index. It respects to the Pigou-Danton principles of transfers, it is less sensitive to outliers, deviations from normality and it is suitable for the segregation measure of continuous variables. Finally normalizing the index by individual level income provides an ideal research environment for comparative studies among different contexts and over time.

4 Empirical Application

In this section we test how the GSS performs with an empirical application to Swedish register data for the years 1994, 2004 and the latest year available in Place database 2014. The database contains the disposable income and residential coordinates of entire population. A common problem shared by previous studies of segregation is the choice of predefined administrative units of analysis. This study uses residential the coordinates that are available for each individual as 100x100m grid units in the database.

We begin by creating individualized-neighbourhoods for each individual's geo-location in the country. As far as the approach to the neighbourhood is concerned, the best method that renders the spatial boundary of the neighbourhood must be chosen. The optimal method potentially varies among different context and especially with the population density at the interested area. Both the radii and knn approaches have their advantages and disadvantages. The radii-based neighbourhood depicts the geography constituted within a predetermined radius, therefore the space as the point of interest. This is a desirable way of studying geography when the analyses focus on locations, services: parks, recreational areas and when there are no changes expected in terms of population density in the space defined. But when the concern is the spatial relationship between individuals, knn approach might be more appropriate. As a population-count method knn successfully illustrates the interaction possibilities between individuals, when the areas are not populated too sparsely. In this paper we make use of both approaches and we propose an intermediate method that benefits the advantages of two.

First, we construct neighbourhoods based on a *knn* algorithm. Each neighbourhood contains the average disposable income earned by nearest 100 200 400... 51200 working-age (20-64) people for each residential location as follows: $\mu_{ik} = \sum_i \frac{y_i}{k}$. Since we work on the entire Sweden, the physical separation between neighbours is an issue that we have to handle especially for the northern parts of the country. For this reason, we introduce a distance decay model $f(\beta, d_{in})$: a function of distance between individual i and their k nearest neighbours $n = 1, 2, 3, ..., k$ with a distance decay factor β. This operation spatially weights the observations, so that as the distance between i and k nearest neighbours increases ks' relative contribution to the average income decreases. Therefore, for densely populated metropolitan areas, the neighbourhood average incomes remain similar to those produced by *knn* algorithm without a decay factor. The computations are carried out by the EquiPop software (Östh, 2014). The EquiPop permits introducing a decay factor prior to computations start and produces outputs both with and without spatial weigths in an efficient way.

Since the daily interaction behavior of residents is not feasible to extract from the data and given the spatially disaggregate nature of income distributions, β is derived mathematically by half-life models[1] as 0.0001153

[1]Common way to determine distance decay factor is to use spatial interaction models(SIMs) where observed flows of people between origins (O) and destinations(D) are regressed over distances between all possible O-D pairs. This operation reduces the overall deviation from the mean commuting distance in a given population. Mathematically derived half-life models (HLMs) are valid alternatives to SIMs when the flows of people are not observed and in the presence of spatially highly disaggregate data. HLMs use median value as departure. This is because the median commuted distance always occur at a distance where half of the population commute longer and half of the population commute shorter distance. Therefore, knowing the maximum distance from i to its *kth* neighbor, we can say that the probability of interacting with neighbours equals 0.5 at the observed median distance. Then for the decay function to describe this probability at various distances, the probability-value

with an exponential decay function. Then the spatial weights function becomes $\exp(-\beta d_{in})$ for each pair of neighbours.

Table 2 shows the GSS values for three years where each value corresponds to a different size of neighbourhood and every second column of a given year reports the index weighted with decaying distance. While the last row shows the overall Gini of disposable income for each year. The slight increase in the Gini from 1994 to 2004 is reflected by the GSS measured at neighbourhood size $k = 100$, for larger k values instead a similar segregation pattern yields. Therefore, from 1994 to 2004 residential segregation by income has increased only at a very small geography i.e. among 100 nearest neighbours. Moreover, the GSS values for the year 2014 show that the increase in inequality at individual level is reflected at any scale of geography.

Furthermore, using the spatial analyst tools available in ArcGIS, we construct radii-based neighbourhoods where the average disposable income in i's neighbourhood is measured as $\mu_{ir} = \frac{\sum_{j\in r} y_i}{n_{ir}}$. Computations are repeated for the radius sizes: $r = 100m, 1km, 5km, 25km$. The estimates reported on Table 3 show increasing index values parallel to increase in individual level Gini over years. There is no direct equivalence between r and k in how much geography is depicted as we move from one definition to other. Fig.1 offers a useful picture how the GSS varies between years and for different neighbourhood sizes and definitions. On the left-hand side a similar pattern is observed for the years 1994-2004, whereas the GSS in year 2014 (grey line) lies above for all k values. Therefore, what we observe from knn approach is that the residential segregation by income remained at a similar rate from 1994 to 2004 despite a slight rise in the overall inequality and it increased in 2014. The radii-based approach (below) instead shows a clear ranking among years 1994, 2004 and 2014 with a similar response to different r values. Compering pictures in the Fig.1, it is evident that the radii-approach exhibits a higher level of segregation

will decay from one at no distance towards almost zero at far, far away (see Östh et al., 2016, for details).

than knn. But the GSS values with decay factor display a similar pattern to radius. Both show a decreasing at a decreasing rate for increasing size of neighbourhoods. The difference between k-nearest and radii-based neighbour approaches becomes clear as we move to the analysis at the municipality level.

To show how economic segregation varies by geographic locations within the country, we compute the GSS separately for 290 Swedish municipalities. Each value represents the ratio of the inequality between average incomes earned in the bespoke neighbourhoods of people who live in the same municipality and the total inequality in the country. We use both radii and k-nearest neighbour aggregates and for the k-nearest neighbour approach we reporte values both with and without a decay factor (=0.0001153). The results for the year 2014 with different radius and k values are shown in Fig.2. The colours correspond to the fixed intervals of GSS values for all maps. This makes easier to compare the values obtained by the two approaches to neighbourhoods.

By looking at the maps for smaller to higher r and k values, a lens scans economic segregation from block level to larger units of localities such as census tracks. Smaller r and k values may communicate a residential segregation in a couple of buildings and as the scale gets larger the GSS may communicate the economic segregation in an area including schools, stores, play grounds etc. The radius and knn approaches display different patterns especially for lower values of r and k. As stated above, the reason for this is that the radii-based approach focuses on the geography only, meaning that the number of people living within a given radius varies between locations (and time) and this is not catered for by this approach. This is evident especially on the first row of Fig.2 with r=100 meters, at this scale the GSS values are very high. They vary between 0.4-0.6 for all municipalities. Even for a small k value as 200 (may be equivalent to a block in a densely populated area), a much lower segregation is observed, close to zero in some municipalities but still retaining the high GSS values for metropolitan areas.

As opposed to radii, knn approach focuses on people and neglects how far they live from each other. This becomes a relevant issue especially for the sparsely populated areas in the northern parts of the country where kth neighbour might reside kilometers away from i. The second and the third rows of the figure offer a useful comparison for this respect. For a smaller value $k = 200$ both maps display a similar pattern, while for intermediate k levels, decayed GSS values capture some of the segregation pattern similar to radii approach. Therefore, the maps on the third row lie somewhere between radii and knn maps, rendering both the number and the geographic distribution of people. Fig.3 shows the change in GSS values from 1994 to 2014. The maps are organized in the same way as in Fig.2. In the northern parts of the country economic segregation mostly remained the same and decreased in couple of northern municipalities. While in the metropolitan areas such as Stockholm, Malmö and Göteborg, it has increased, even for higher aggregates of people i.e. for larger r and k values.

In the next step we report spearman[2] rank-order correlations between the computed GSS values and several characteristics of the municipalities to explore the properties of areas signaling the current state of the residential segregation by income. The municipalities are characterized based on the variables provided by Statistics Sweden. We use the information on the changes in employment by an application of the Martin resilience-employment index (Martin, 2012) measuring the growth rate of employment overtime and expressed as the change over time $\Delta = t - t + 1$ and employment levels at local (Er) and national level (EN).

$$REI_i = \frac{(\Delta Er/Er) - (\Delta EN/EN)}{|\Delta EN/EN|}$$

We find a significant positive correlation between REIs calculated for the time intervals 1993-1994 and 2013-2014 and the GSS indices for 1994

[2]Spearmans correlation coefficient is a statistical measure of the strength of a monotonic relationship between paired data.

and 2014, respectively. This can be interpreted as a degree of job/housing balance, that the population tend to cluster both as employed/unemployed, therefore employment growth is associated with economic segregation and by the earnings generated from different types of jobs. We find negative correlation with cost-equalization grants per municipality, this suggests that the equalization grants ensure a degree of residential mixture by income. Moreover, there is a significant positive correlation with election participation rates and strong positive correlation with the number of low/high educated people in areas. These two correspond to the sorting behavior of individuals by political orientation/politicization and education.

Fig.4 displays how the spearman correlations vary for the GSS indices computed for increasing k values in 1994 and 2014. In both years a constant association is observed with REI and election participation rates for any scale of geography, whereas in both years the correlation with cost-equalization spending and number of low/high educated people decreases after $k = 6400$. This is probably because at this scale we exceed the municipality size. Interestingly, despite the increase observed in economic segregation from 1994 to 2014 (see Table 2), the exact pattern of sorting by education is observed in the two years in response to varying scales of the neighbourhood.

Sweden has three levels of government; the central government(staten), county council (landstinget) and municipality (kommunen). All three are allowed to tax personal income and the municipalities provide public services and are subject to income, cost and structural equalization grants. The first two are purely re-distributive, the income equalization grants are transferred from the high income municipalities to the ones with lower incomes and the cost equalization grants are transferred to municipalities with less favourable cost structure from the ones with better conditions. The structural equalization instead is a grant from the central government to municipalities with a small population or a high share of unemployment.

The strong local government sector in Sweden offers a useful research environment for the study of economic segregation, particularly useful to possible policy implications for other contexts. Using OLS regressions, we test the effects of all three equalization grants took place in 2013 on the GSS values computed with *knn* (decayed) approach to neighbourhoods for the year 2014 (see Appendix 1). We find that economic segregation decreases with both income and structural equalization grants at any scale of neighbourhood. Moreover, looking at the cost equalization grants for different structural costs, again the transfers are associated with a degree of residential mixture by income. For instance, the grants to equalize costs of heating, public transportation, upper secondary and compulsory schools decrease the economic segregation in the following year. Whereas the municipalities that receive higher grants for the costs of streets&roads and the grants devoted to children with foreign background on average experience higher degree of economic segregation. A possible interpretation as far as the grants for streets&roads are concerned is that the municipalities in need of grants due their relatively less favorable infrastructure may show higher physical separation between classes. Similarly, the municipalities that are eligible to receive grants for children with foreign background constitute higher minority population that potentially cluster in neighbourhoods due to relatively lower earnings.

5 Conclusions

So far, the segregation measures found in literature have been mainly developed to measure the extent to which individuals are clustered by groups: typically race, ethnicity, gender in occupations. The residential segregation by income instead has not received much attention in the literature. The most of the existing studies of the latter use the indices originally developed for racial segregation by dividing population into two categories; being under and above to a given level of income i.e. poor and not poor. By restricting the analysis to two groups, the indices do not make full use of the available information. Moreover, nearly all existing indices are aspatial in nature, that they do not take into account the distribution of individuals in space. Although there exits few spatial ones, they are rather difficult to compute and nearly all use some administratively defined area for the unit of analysis.

In this paper, we offer a new measure of residential segregation by income based on a individualized-neighbourhood approach and therefore makes use of the full information on the income distribution of residents and their distribution in space. The proposed index allows to handle the geography flexibly as neighbourhoods can be constructed by both radii and *knn* approaches and the scale can be set to varying levels to meet distinctive characteristics of different contexts. This last point allows the index to avoid robustness issues associated with MAUP and checkboard phenomenon, the problems that may severely distort the sensitivity of the results of spatial analyses. Additionally as a ratio of two Gini indices, the index has the advantage of preserving desirable properties of the Gini. It respects to the Pigou-Danton principles of transfers, it is less sensitive to outliers, deviations from normality and finally it is suitable for the segregation measure of continuous variables.

Moreover, using the Swedish register data we have tested the efficiency of the index. We have used both approaches to individualized neighbourhoods and by employing spatial weights matrix based on the distance between neighbours we have proposed an intermediate approach that benefits the advantages of both. In particular to Sweden, the estimates suggest that the economic segregation has remained at a similar degree from 1994 to 2004. Although it has increased from 1994 in 2014 in parallel to rise in inequality, correlation analysis has shown that the individuals sort into neighbourhoods almost identically in both year. Additionally, the analysis on the influence of transfers among municipalities has illustrated that as the smallest local governments in Sweden, the municipalities can reduce the economic segregation through equalization grants of income, and several structural costs.

References

Abramson, A. J., Tobin, M. S., & VanderGoot, M. R. (1995). The changing geography of metropolitan opportunity: The segregation of the poor in us metropolitan areas, 1970 to 1990. *Housing Policy Debate*, *6*(1), 45–72.

Burgers, J., & Musterd, S. (2002). Understanding urban inequality: a model based on existing theories and an empirical illustration. *International Journal of Urban and Regional Research*, *26*(2), 403–413.

Burgess, E. W. (1928). Residential segregation in american cities. *The Annals of the American Academy of Political and Social Science*, *140*(1), 105–115.

Cortese, C. F., Falk, R. F., & Cohen, J. K. (1976). Further considerations on the methodological analysis of segregation indices. *American sociological review*, 630–637.

Dawkins, C. J. (2004). Measuring the spatial pattern of residential segregation. *Urban Studies*, *41*(4), 833–851.

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

Epple, D., Filimon, R., & Romer, T. (1984). Equilibrium among local jurisdictions: Toward an integrated treatment of voting and residential choice. *Journal of public economics*, *24*(3), 281–308.

Essletzbichler, J. (2015). The top 1% in us metropolitan areas. *Applied geography*, *61*, 35–46.

Getis, A. (2009). Spatial weights matrices. *Geographical Analysis*, *41*(4), 404–410.

Harris, C. D., & Ullman, E. L. (1945). The nature of cities. *The Annals of the American Academy of Political and Social Science*, *242*(1), 7–17.

Hoyt, H. (1939). The structure and growth of residential neighborhoods in american cities. *Washington, D.C.: Federal Housing Administration.*

James, D. R., & Taeuber, K. E. (1982). *Measures of segregation.* JSTOR.

Jargowsky, P. A. (1996). Take the money and run: Economic segregation in us metropolitan areas. *American sociological review*, 984–998.

Kim, J., & Jargowsky, P. A. (2005). *The gini coefficient and segregation on a continuous variable.*

Lee, B. A., Reardon, S. F., Firebaugh, G., Farrell, C. R., Matthews, S. A., & O'Sullivan, D. (2008). Beyond the census tract: Patterns and determinants of racial segregation at multiple geographic scales. *American Sociological Review*, *73*(5), 766–791.

Martin, R. (2012). Regional economic resilience, hysteresis and recessionary shocks. *Journal of economic geography*, *12*(1), 1–32.

Massey, D. S., & Denton, N. A. (1988). The dimensions of residential segregation. *Social forces*, *67*(2), 281–315.

Massey, D. S., Fischer, M. J., Dickens, W. T., & Levy, F. (2003). The geography of inequality in the united states, 1950-2000 [with comments]. *Brookings-Wharton papers on urban affairs*, 1–40.

Morenoff, J. D., & Tienda, M. (1997). Underclass neighborhoods in temporal and ecological perspective. *The Annals of the American Academy of Political and Social Science*, 59–72.

Openshaw, S. (1984). The modifiable areal unit problem. *Geo Abstracts University of East Anglia.*

Östh, J. (2014). *Introducing the equipop software–an application for the calculation of k-nearest neighbour contexts/neighbourhoods.*

Östh, J., Clark, W. A., & Malmberg, B. (2015). Measuring the scale of segregation using k-nearest neighbor aggregates. *Geographical Analysis*, *47*(1), 34–49.

Östh, J., Lyhagen, J., & Reggiani, A. (2016). A new way of determining distance decay parameters in spatial interaction models with application to job accessibility analysis in sweden. *EJTIR*, *16*(2), 344–363.

Reardon, S. F. (2011). Measures of income segregation. *Unpublished Working Paper. Stanford Center for Education Policy Analysis*.

Reardon, S. F., & Bischoff, K. (2011). Income inequality and income segregation 1. *American Journal of Sociology*, *116*(4), 1092–1153.

Reardon, S. F., Farrell, C. R., Matthews, S. A., OSullivan, D., Bischoff, K., & Firebaugh, G. (2009). Race and space in the 1990s: Changes in the geographic scale of racial residential segregation, 1990–2000. *Social Science Research*, *38*(1), 55–70.

Reardon, S. F., & Firebaugh, G. (2002). Measures of multigroup segregation. *Sociological methodology*, *32*(1), 33–67.

Reardon, S. F., Matthews, S. A., OSullivan, D., Lee, B. A., Firebaugh, G., Farrell, C. R., & Bischoff, K. (2008). The geographic scale of metropolitan racial segregation. *Demography*, *45*(3), 489–514.

Schelling, T. C. (1969). Models of segregation. *The American Economic Review*, *59*(2), 488–493.

Tiebout, C. M. (1956). A pure theory of local expenditures. *Journal of political economy*, *64*(5), 416–424.

White, M. J. (1983). The measurement of spatial segregation. *American journal of sociology*, 1008–1018.

Wong, D. W. (1993). Spatial indices of segregation. *Urban studies*, *30*(3), 559–572.

Wong, D. W. (1999). Geostatistics as measures of spatial segregation. *Urban Geography*, *20*(7), 635–647.

Wong, D. W. (2004). The modifiable areal unit problem (maup). In *World-minds: Geographical perspectives on 100 problems* (pp. 571–575). Springer.

Wong, D. W. (2005). Formulating a general spatial segregation measure. *The Professional Geographer*, *57*(2), 285–294.

Wyly, E. K., Glickman, N. J., & Lahr, M. L. (1998). A top 10 list of things to know about american cities. *Cityscape*, 7–32.

Tables and Figures

Table 1: Gini indices of Segregation [a]

	Gini Segreagation	Binary	Multigroup	Continuous	Spatial
Massey & Denton (1988)	$G = \sum_i \sum_j \frac{t_i t_j \lvert p_i - p_j \rvert}{2T^2 P(1-P)}$	✓			
Reardon & Firebaugh (2002)	$G = \sum_m \sum_i \sum_j \frac{t_i t_j}{2T^2 I} \lvert \pi_{im} - \pi_{jm} \rvert$		✓		
Kim & Jargowsky (2005)	$G = \frac{\frac{1}{2N^2\mu} \sum_i \sum_j \lvert y_{ni} - y_{nj} \rvert}{\frac{1}{2N^2\mu} \sum_i \sum_j \lvert y_i - y_j \rvert}$	✓		✓	
Dawkins (2007)	$G = \frac{Cov(Y_j, \hat{R}_{j(n)})}{Cov(Y_j, \hat{R}_j)}$			✓	✓

[a] T total population, ti and t_j total populations at i and j, P total minority population, p_i minority population at i, m number of groups, π_{im} proportion in m at i, y_i income of i household, y_{ni} average income at n, μ overall average income, N total number of households, Y_j aggregate income at j, $\hat{R}_{j(n)}$ spatial re-ranking of neighbourhood, $\hat{R}_j$ average ranking of neighbourhood

Table 2: GSS for different k values

k	GSS_94	GSSdecay_94	GSS_04	GSSdecay_04	GSS_14	GSSdecay_14
100	0,315	0,319	0,323	0,327	0,328	0,333
200	0,287	0,294	0,287	0,293	0,298	0,306
400	0,264	0,272	0,260	0,269	0,274	0,286
800	0,241	0,253	0,235	0,247	0,2548	0,270
1600	0,220	0,235	0,215	0,231	0,237	0,258
3200	0,200	0,220	0,199	0,218	0,221	0,248
6400	0,179	0,204	0,180	0,207	0,204	0,237
12800	0,156	0,188	0,158	0,192	0,182	0,222
25600	0,139	0,177	0,139	0,180	0,161	0,209
51200	0,121	0,172	0,120	0,177	0,140	0,205
GINI(Individual)	0,257		0,262		0,332	

Table 3: GSS for different r values

r	GSS_96(radius)	GSS_04(radius)	GSS_14(radius)
100m	0,436	0,475	0,497
1000m	0,290	0,327	0,373
5000m	0,212	0,264	0,302
10000m	0,186	0,236	0,267
25000m	0,158	0,186	0,209
GINI(Individual)	0,257	0,262	0,332