eman ta zabal zazu

Universidad del País Vasco

Euskal Herriko Unibertsitatea

Master's Thesis

Measuring economic segregation in the US-Metropolitan areas

University:	University of the Basque Country UPV/EHU
Study Program:	Master's in economics: Empirical Applications and Policies
Author:	Robert Kurtz (20th of October 1996)
Supervisor:	Casilda Lasso de la Vega
Academic Year:	2020/2021
Date of Submission:	23th of September 2021

Empirical Applications and Policies

Table of Contents

List of Abbreviations

ACS	American Community Survey
EI	Entropy Index
GC	Gini Coefficient
HEI	Household Equivalent Income
HTI	Household Total Income
IEI	Income Ethnical Inequality
IESI	Income Ethnical Segregation Index
IPUMS	Integrated Public Use Microdata Series
IS	Income Segregation
J&K	Jargowsky & Kim
MA	Metropolitan Areas
NSI	Neighborhood Sorting Index
SSI	School Separation Index
SR	Square root
TIEI	Total Income Ethnical Inequality
TBI	Theil Between Index
US	United States of America

List of Tables

List of Formulas

Glossary

Decomposability

Decomposability is a strong property that is also stating the mathematical expression of an index. Furthermore, it implies that segregation as a value can be written as the total sum of in our case segregation or inequality of all the subregions, plus a between term. Therefore, it is the sum of the inequality or segregation among the subgroups. Consequently, if one of the sums increase the total sum does so as well. Consequently, decomposability always implies subgroup consistency.

$$Decomposability = Sum\ of\ Inequality\ of\ Subgroups + The\ "Between"$$

Income Inequality

Expresses how unevenly income is distributed amongst a population. Income is defined as the yearly household's disposable income. Consequently, the less monotone the dispersion the higher the degree of inequality. Inequality only concerns the differences within a certain population and does not put it into relation with subgroups for example ethnicity. (OECD, 2015); (Lynch, John W, 1998)

Income Segregation

IS concerns the allocation of primary units based on household income as well as a secondary variable that can be for example regional factors such as MA, education or ethnical socioeconomic properties. (Jargowsky, 1996)

Metropolitan Area

A region represented by an extensive populated urban core area and its suburban enclosing territories, which have a common and equal administrative jurisdiction, comparting industries, infrastructures, and housing. (John S. Adams, Barbara J. Van Drasek & Eric G. Phillips, 1999)

Sample

In our case we selected three samples as the basic data to operate with. One sample displays data of the respective year, whereby all the variables remain the same throughout the years. The IPUMS data samples that were selected are 1990 5%, 2000 5% and the 2010 ACS. The observations for the 3 years accumulate to more than 8 million.

Sample Weights

Important for the regressions are furthermore the "sample weights", where a higher importance is given to households that rather display the average characteristics of an US-American household in that respective year. Thus, a household with twelve children would be rather extraordinary and therefore obtain a lower weighting for the variance estimations.

Subgroup Consistency

Generally, not only for inequality or segregation but any kind of indexing we can talk about this important property. To explain what it means in our case of income and ethnical segregation let us imagine that for one of the subgroups the segregation increases. Subgroup consistency implies that as long as the other subgroups remain unchained, total segregation should also increase.

Variable

The specific variables selected for the analysis will be described in detail in the "Data section" in part 3. The variables are mostly social economic information for the US population that can be individually selected on the IPUMS website.

Abstract

This paper aims to provide empirical evidence on “Income Ethnical Segregation” by constructing an index that analyses US Census Data from 1990 until 2010. Following the groundwork of the “SSI” the results can be considered as decomposable and thus also subgroup consistent. The evidence for the biggest 10 MA shows a clear increase of segregation and inequality even though average income climbs as well throughout the years. Same implies that significant differences amongst races have been observed that were especially revealed for the poorer MA such as Los Angeles-Long Beach, CA. As the “Income Ethnical Segregation” can be seen as one component of the overall score of the Entropy Index, which consists of the “Within – Inequality part” and the “Between – Segregation part”, it goes hand in hand with previous assumptions that a higher segregation also estimates higher inequality. The reader will not only be guided through the various indicators and empirical evidence for the case, but moreover also learn about practical application of previously established indices based on the Theil-Between Index. Thus, the main takeaway will be an understanding of the scientific requirements and its empirical application additional to the real analysis on US-Microdata timeseries, when working with segregation measures that include the problematic and continuous variable that is “Income”.

Keywords: Income segregation, income ethnical segregation index, inequality, development analysis, entropy index, household equivalent income, US - metropolitan area, decomposability and subgroup consistency.

1. Introduction

Empirical analysis from Census Data over the past 40 years has shown that income inequality in the United States has increased drastically, with top 1% earners making more than 21% of the total income generated in the year 2012. Whereas back in 1970 that percentile of top earners only accounted for 8,5% of the whole population (Stanford Center on Poverty and Inequality, 2021). This study aims to empirically review whether this trend was also observable for the field of income segregation, where not only the inequality within a certain subgroup is being compared but moreover the degree of disparity amongst subgroups.

Segregation refers to the extent that individuals are classified into different characteristics such as race, level of education or household income in regard to for example the area or neighborhood in which they are living. In other words, if we are setting up demographic grouping on "race" we are speaking of ethnic segregation which is based on a non-ranking and categorical variable. Moreover, segregation can be done in the economic field of Income as well, which is the primary variable of interest in this study. The index we are creating can consequently be referred to as "Income Ethnical Segregation Index" with main variable of interest being "Household Equivalent Income" and a secondary categorical variable that is "race". Thus, "Income Ethnical Segregation" concerns the extent to which households with different income levels belong to different ethnical groups.

However, previous research struggled when segregating, for continuous ranked variables such as Household Income and corresponding methodology has been rather short. Common economical tools known from the field of social welfare and inequality such as the Gini-coefficient have been proven to fail making profound and consistent regressions due to subgroup inconsistencies as well as in-decomposability for continuous and ranked variables (Jargowsky, 1996). Lasso de la Vega and Volij in 2020 show that the SSI index which they introduced (is fact the "Second Measure of the Theil-Between") is the only one that regarding

IS fulfills subgroup consistency and decomposability as well as other crucial properties considered segregation. Consequently, this study ambitions on continuing groundwork and empirically putting same into practice by reviewing US census data from over three decades and to contribute to the general problem of constructing a profound methodology which succeeds to measure segregation for continuous variables. It demonstrates how a continuous variable "Household Equivalent Income" is interlinked with the district of living/ regional factors as well as ethnicity. The 10 biggest US-Metropolitan Areas are compared to obtain results whether richer areas tend to show more inequality and segregation or vice versa. From the microdata also socioeconomic information such as race is being made available wherefore, the IS index can be tested on this categorical variable as well. Where presumably people of color have a higher score of inequality due to ongoing opportunity biased as well as a higher number of poor people in relation with "whites". The outline of this thesis was limited to the ten biggest Metropolitan Areas, as the total of 384 MA in the US would spread the scope of this work and would not allow profound data analysis in regard to determining factors for regions etc. Three years 1990, 2000 and 2010 have been selected to serve as vivid demonstration of the development taking place in the US. More recent data was lacking information about the crucial data on the MA which was only made available by the US governmental Data Base "IPUMS" until the year 2011. Finally, the developed index should be considered as relevant as it can be applied to other countries as well once same data is applicable. However, also the scope of this work must be fenced due to the limited dimensions of this Master's Thesis. Thus, a further investigation of the empirical application of the introduced "Income Ethnical Segregation Index" for MA can be executed as well as more in-depth subcategory differentiations.

In the following section "2. Literature Review" the scientific situation of the matter is being demonstrated as well as where the practical relevance of the new index lays. The rest of this thesis is structured as follows: The section "3. Data" shall provide detailed information on the

data source and important background information concerning the samples and variables selected for the regressions. In the passage “4. Methodology” the methodology utilized will be displayed with the aim that the reader has a clear vision of what theoretic models have been applied and customized to provide profound and efficient estimations in the field of “Income Ethnical Segregation” whilst fulfilling the crucial property of decomposability. Thereafter, the theoretic framework is put into empirical practice and a close analysis on the surrounding factors for the 10 biggest US-Metropolitan Areas will be evaluated in the section “5. Income Segregation in the US-Metropolitan areas”. Also the analysis of the segregation and inequality developments will be closely observed and demonstrated before definitely concluding most important takeaways and findings in the final “6. Conclusion”.

2. Literature Review

When speaking of segregation especially in the economic field of income segregation one name seems to always fall, which is Paul A. Jargowksy who for many US relevant scientific contributions laid the groundwork in his paper "Take the Money and Run: Segregation in U.S. Metropolitan Areas" back in 1996. He documents that economic segregation especially for Afro-American and Hispanics increased in US-MA especially during the 1980s. It will thus be interesting to work with more recent data in this paper and to check whether this trend is also observable in more recent times. In order to measure economic segregation Jargowsky introduces the so-called "Neighborhood Sorting Index" or NSI. In other words it is the variance correlation of household income in relation with different "neighborhoods" within a MA and can be expressed as:

Formula 1: Neighborhood Sorting Index

$$\boldsymbol{NSI} = \frac{\sigma_{Neighborhood}}{\sigma_{Household}} = \frac{\sqrt{\frac{\sum_{n=1}^{N} h_n(\bar{y}_n - \bar{y})^2}{H}}}{\sqrt{\frac{\sum_{i=1}^{H}(y_i - \bar{y})^2}{H}}}$$

Where y is household income, i indexes households, h_n represents the number of households in the respective neighborhood and H and N embodies the total number of households and neighborhoods. It is the ratio of the variance of the neighborhood households and the total variance of the district. Based on the NSI Jargowsky is making comparisons for the 10 largest MA which is exactly what this thesis is aiming at with the difference of being indexed entirely different as well as with more recent census data. Furthermore, Jargowsky in his paper 1996 bases the ethnical part of the segregation only on three "races" which are White, Black and Hispanic. Later he claims that the NSI is an adequate measure of segregation also for the continuous variable that is household income. However, later on we will see that this statement is by now considered outdated and there is new research proving that this claim does not hold

because of subgroup inconsistency. However, what is indeed equal to the new "Income Ethnical Segregation Index" that will be introduced in the Methodology section of this paper is his approach of spatial or in other words geographical segregation in terms of MA and the combination of household income and ethnical segregation. (Jargowsky, 1996)

In his work from 2005 "The GINI Coefficient and Segregation on a Continuous Variable" Jargowsky again mentions the difficulties of measuring IS due to the problems for crucial continuous variable which is income. The original approach from Massey and Eggers back in 1991, which suggests to establish different categories for different amounts of income (e.g. to divide the population into rich, middle and poor) has been heavily criticized for discarding information and for potentially confounding fundamental distributions in regard to segregation. Furthermore, laying the borders for being considered "rich" etc. is to an extend highly subjective and simplifying. Thus, Jargowsky intends to establish his NSI further by comparing it with the well-known inequality measuring tool which is the Gini-Coefficient that has been introduced firstly in the year of 1912 by the Italian Statistician C. Gini and is derived from the Lorenz Curve. The inequality measure for a population hereby lays between zero and one whereas zero is representing perfect equality and one total inequality (Gini, 1955). It is globally known to report about state of income inequality (Deltas, G., 2003). Indeed, Jargowsky once he came up with the empirical conclusions of his comparison between the NSI and the G^S noted a very high correlation between the two measurements indicating for the same observed phenomenon in the data. However, as Jargowsky later claims the GC does not serve as consistent measurement for income segregation as it fails to be subgroup consistent and therefore also not decomposable. (Kim, J., and P. A. Jargowsky, 2005)

Therefore, Lasso de la Vega and Volij show in their paper from 2020 published in the International Economic Review called "The Measurement of Income Segregation" theoretically as well as empirically, that the construction of a new index is necessary. This introduced index

to face the problem of measuring IS can be seen as the second measure of Theil's index for measuring income inequality and comes from the family "Generalized Entropy Indices". It displays income segregation for schools, where "schools" like in our example "race" can be seen as the subcategory and secondary segregation variable and where a higher amount of segregation is assumed for schools in districts with higher economic inequality. Below the formula for the SSI is described in detail in the appendix and will be assessed closely later on in the Methodology section as it basically follows the same exact principle as for the later introduced "Income Ethnical Segregation Index".

Formula 2: School Separation Index

$$\boldsymbol{SSI}(\boldsymbol{x}) = \text{"}Income\ Ethnical\ Segregation\ Index(x)\text{"}$$

Thus, instead of working with ethnicities in different MA, Lasso de la Vega and Volij work with the segregation between schools, while also treating with the primary continuous variable that is income. They proof that the SSI is characterized to fulfill all the properties such as subgroup consistency as well as decomposability et cetera and put same into practice by reviewing the assumptions of the index with real data provided by the "Education Quality Agency of Chile". (Lasso de la Vega and Volij, 2020)

Finally, on that basis the "Income Ethnical Segregation Index" will be constructed to face the problem of solving for IS in the US-Metropolitan areas. As shown earlier in this section previous US concerning Income Segregation attempts such as the one from Kim and Jargowsky in 2005 did not fulfill all the properties necessary when treating with characterizing the continuous variable "Income". The SSI however, that was proven to achieve all the surrounding criterions can thus be applied for different subgroups e.g., "race" when analyzing real data. In the following section the empirical evidence used for the analysis and its roots are depicted to get a better understanding of the data before diving into the methodology and indexing.

3. Data

Every empirical and scientific work needs some actual data as the foundation for subsequential analysis and conclusions. As we are looking at economic segregation for various years, in this case over three decades, data needs to be available and similarly subdivided to have the same variables used for the regressions and comparisons. Whereas European databases seem to lag consistency over longer time periods, the data from the "Integrated Public Use Microdata Series" or short "IPUMS USA" accomplishes to assemble, preserve, and interlink census microdata on Social, Economic, and health Research. IPUMS provides freely accessible data, which includes decennial censuses from 1990 until 2010 as well as American Community Surveys (ACS) from 2000 until the present. The data extracts can be individually selected in terms of variables and samples (see Glossary for further description). In this case three samples with each 27 variables have been selected leading to an overall of more than eight million observations, that represent a total population of 2,7 million weighted individuals for each year. Due to this extensive number, it makes sense to isolate with the help of Stata specifications such as to one year e.g., 2010 and to focus on ten Metropolitan Areas e.g. "New York, NY-Northeastern NJ". After this process there are still more than 100.000 observations, however the data analysis can be executed more precisely and efficiently in the beginning. The Family income was selected to be the basis economic variable to apply the HEI as described in the Formula in Appendix 1 as well as in the following Methodology part. Also, geographic information specially the Metropolitan Areas, as well as Ethnicity of the individuals that was postliminary used for the comparison of the segregation. Unfortunately, the data for the MA was only available until the year of 2011 not allowing for more recent IS analysis. In the following section of this paper the fundamental development of the methodology will be outlined such as the crucial variable for IS "Household Equivalent Income" as well as the underlying theoretic models and more.

4. Methodology

4.1 Household Equivalent Income

When constructing an Index to measure and compare Income Segregation one of the essential questions is which unit to use as comparison basis. Whereas income of individuals could in theory be compared but show by experience a far greater variance due to age, family constellation and other factors, the common measurement tool in the economic field of IS has been the "Household income". In other words, individuals with the same household identification number who belong to the same family or communal residence are bundled together to have a more evenly spread medium of income. Whereby the main breadwinner usually is represented by the parents and the children if any do not contribute to that statistic as they are not subsidized financially. However, since family sizes can vary greatly and to make the comparison basis even more equal, we introduce a variable that we call "Household Equivalent Income" (HEI). It is calculated by dividing the Total Household Income by the square root of the amount of family members (see Formula for HEI in the Appendix). The HEI can now be utilized to serve as the comparison basis for the indexing as it aims as demonstrating what an adult person household would earn hypothetically. Important to mention is that for the calculations the sample weights have been used that can e.g., be found in the household weight.

4.2 Cleaning the Data

Stata is the data analysis program of choice, which is used for the applications of the regressions and the indices. The IPUMS data includes information about sample wages, which hold important information about relevancy of individuals for the variance calculations. As aforementioned units with a higher statistical probability to appear in the set receive a higher weight than units that rather do not display the sample average e.g., a family with more than 10 family members. To cut out misleading data in example if the HEI is displayed as infinite or negative number, which is not sensible it can be disregarded by eliminating those values with

the “drop” operate. Furthermore, with the “tab” function the Metropolitan Areas are displayed and show how many people are living in each district. Consequently, the isolating process can be started as for the indexing the total amount of hundreds of MA would clearly exceed the scope of this thesis. Therefore, the 10 biggest MA have been selected to serve as a basis for the IS like it has been done similarly in the paper of Jargowsky in 2005. Further isolating can also be used for the comparisons between the years as well as the different MA by only keeping their respective values.

4.3 Isolating for the Segregation Process

As mentioned in the introduction about the different types of segregation, fortunately the provided data also includes information about ethnicity, which allows for a further application of the Theil-Between Index. This time however, for the non-ranking and categorical variable which is called “race”. The variable is subdivided into nine different categories which are the following: White, Black-African/American, Indigen-Native, Chinese, Japanese, Other Asian or Pacific Islander, other race, two major races and three or more major races. Consequently, in the following part this allows comparisons amongst different ethnicities, which can also be counter checked with the hypothesis that different subcategories show different average incomes and inequalities. Generally, we analyze if and to what extent in a metropolitan area e.g., NY, the equivalent household income is unequally distributed within ethnic groups of individuals and how this inequality has evolved over the years. Finally, the results are compared with other metropolitan areas and conclusions are drawn. Subsequently, the IS process can be started by applying the Theil-Between Index for the different years and regions (MA). Same is done with a theta level at zero as it is required for the calculations since $\theta\ (Theta) = 0$ corresponds to the second Theil index, which is the one used.

4.4 Theil Between Index – Income Ethnical Segregation

The "Theil Between Index" is known as the Theil's second measure of income inequality and is primarily used as a statistic to determine the level of economic inequality, however also as determinant for racial segregation. It was firstly introduced by the Dutch econometrician Henri Theil back in 1967 and can be seen as a specific application of the general entropy index. The "Between" component thereby provides information about the inequality amongst subgroup such as "race". Therefore, and this is the vital difference we speak for the "Between" of the segregation score, as we know that when dealing with inequality of a society and the unit is no longer the individual but the subgroup (here "race") it is segregation and not inequality. The total amount is the sum of the subcategories indexed divided by their percentage of overall population of the MA.

Formula 3: The "Between" (Income Ethnical Segregation Index)

$$\boldsymbol{IESI(x)} = \sum \frac{n_c}{n_x} * ln\left(\frac{\mu_x}{\mu_c}\right)$$

Where x is the whole population of the respective metropolitan area; c is each subcategory, for instance the white individuals, black individuals, etc. Furthermore, n_x is the total sum of population in the corresponding metropolitan area and μ_x its total mean income. On the other hand, μ_c is the mean income of the respective subcategory e.g., mean "White HEI". For instance, in our case, if we consider nine different ethnicities we will have, n_1: number of white individuals, n_2: number of black individuals, until we have the last n_9: number of three or more major races individuals using respective sample weights. Furthermore, μ_1: is the mean income of the white individuals, μ_2: the mean income of the black individuals et cetera. At the same time those variables are all referring to a specific metropolitan area or for the median calculation to all ten of them that will be calculated.

The outcome will be an interpretable coefficient that predominantly lays between zero and one but is not bound by its upper limit and can easily exceed one. Especially, for populations whose percentage rather makes a small proportion of the total population e.g., the "Japanese ethnicity" which oftentimes were much higher scored than larger population ethnicities such as e.g., "Whites". Generally, the higher the score the higher the respective segregation.

To better understand the Index, we will look separately at each side. The left-hand-side simply shows the weighting of the subgroup. In other words, clearly the bigger the subgroup the larger the effect for the total score. The right-hand-side shows the real segregation based on the "Household Equivalent Income", where the score will be multiplied with the natural logarithm. Let us imagine that we have perfect inequality, so each subgroup income is exactly the total population mean income. Then, we would have a quotient of one that we multiply with the logarithm leading to a zero. Thus, the total segregation would also be zero as if one of the factors of a multiplication is zero its product will always be zero as well. Surely, this case will unlikely occur in reality and thus we can analyse the segregation score in order to make inequality statements amongst subgroups. The "Household Equivalent Income" will be shown for all the MA and subdivided into the race categories used. Thus, not only can we make assumptions on the inequality amongst MA but moreover within the inequality amongst races within one MA

4.5 The "Within" and the Entropy Index

Earlier it has been displayed what makes the difference between Income Ethnicity Inequality and Income Ethnical Segregation. Thus, for the analysis of the financial inequality within a subgroup we refer to as "The Within" which serves as direct comparison bases for "the Between". Assumingly, MA with a higher Segregation score will also show a higher tendency for the overall Inequality score expressed by the Entropy Index that combines "The Within and Between" in one score.

5. Income Segregation in the US-Metropolitan areas

5.1 Introduction to the findings

As in the previous part the Methodology has been explained in detail, now consequently we are putting same into empirical practice. But to look at IS on race, we need to get familiar with the data that is present firstly. Consequently, the primary step is to analyze which are the biggest metropolitan areas throughout the years. Thus, the below overview *Table 1* for the population amongst the 10 biggest US-MA over the years from 1990 until 2010 can be observed. As shown throughout the years, the global trend of total population rising can also be concluded for the sample data.

TABLE 1: POPULATION OVERVIEW

			Year		
#	**Metropolitan area**	**Total Frequency**	**1990**	**2000**	**2010**
1	New York, NY-NJ	446.718	137.919	153.724	155.075
2	Los Angeles-Long Beach, CA	353.487	109.146	121.286	123.055
3	Chicago, IL	216.587	67.571	73.051	75.965
4	Dallas-Fort Worth, TX	139.091	35.462	45.908	57.721
5	Washington, DC/MD/VA	134.204	36.669	44.819	52.716
6	San Francisco-Oakland-Vallejo, CA	128.863	39.219	43.956	45.688
7	Philadelphia, PA/NJ	122.988	38.752	41.466	42.770
8	Houston-Brazoria, TX	116.400	30.280	38.325	47.795
9	Boston, MA/NH	106.199	31.232	36.503	38.464
10	Detroit, MI	100.712	32.788	34.360	33.564
	Total	**1.865.249**	**559038**	**633398**	**672.813**

With the biggest three MA, New York, LA, and Chicago leading the charts throughout the years, only few changes can be determined in the ranking of the rest e.g., the increase of the Dallas population. Furthermore, the largest MA areas can be considered as highly representable for whole US population as still more than 1,8 Million observations for the analysis from basically all relevant areas spread among the United States is embodied. The biggest one New

York, NY-NJ with a combined observation amount for the three years of 446.718 is remarkably larger than the 10th place represented by Detroit, MI but which has been famous through his industrial car manufacturing scene since long time. For the total segregation and inequality score naturally, bigger MA affect the score more than smaller ones. However, certainly when treating with segregation looking at the population size is not sufficient. Therefore, we introduce in *Table 2* the ethnical distribution amongst the years.

TABLE 2: ETHNICAL DISTRIBUTION

			Year					
Ethnicity	**Total Frequency**	**(%)**	**1990**	**(%)**	**2000**	**(%)**	**2010**	**(%)**
White	1.246.846	66,85	401.491	71,82	402.411	63,53	442.944	65,83
Black/African American	257.304	13,79	77.491	13,86	91.836	14,5	87.977	13,08
American Ind. /Alaska Nat.	7.107	0,38	1.948	0,35	2.563	0,4	2.596	0,39
Chinese	44.757	2,4	9.955	1,78	14.273	2,25	20.529	3,05
Japanese	8.667	0,46	2.942	0,53	2.749	0,43	2.976	0,44
Other Asian or Pacific Islander	99.143	5,32	20.784	3,72	31.846	5,03	46.513	6,91
Other race	159.650	8,56	44.427	7,95	65.719	10,38	49.504	7,36
Two major races	39.114	2,1	NA	NA	20.854	3,29	18.260	2,71
Three or more major races	2.661	0,14	NA	NA	1.147	0,19	1.514	0,23
Total	**1.865.249**	**100**	**559.038**	**100%**	**633.398**	**100%**	**672.813**	**100%**

By far "White" is the ethnicity that is predominantly represented for the 10 biggest MA with a total amount of 1,246 million out of 1,856 million observations overall making a percentage share of 66,85% or about 2/3 of total population. Later, we will see that therefore changes in IS for aforementioned ethnical group is affecting overall segregation the most. The 2nd largest share for the US population is making "Black and African American" citizens that overall display 13,79% of total population. "American Indian and Alaska Natives" make only a relatively small portion of less than a half percent, since Latinos are not considered for this

category but moreover could be embodied in “Other race” or “Two or Three major “races”. Example given a person with a mother that is considered as “African American” and a father that is a “White American” or any other ethnicity can impossibly only be considered for one category and is hence a “Mixed race”. Therefore, even those classifications of more than one race are introduced but lack information on detailed ethnical segregation information. Finally, the Asian fraction is divided into “Chinese”, “Japanese” and other “Asian or Pacific islander”. Moreover, it is remarkable that the percentage of “Whites” decreases from 1990 until 2000 significantly from 71,8% to 63,53% but increases back to 65,8 until the year of 2010. It can thus not be concluded that an ongoing integration process is observed at least for the present data until 2010. Following Census data from the “US-Census Bureau” in the year of 2019 for the whole US-Population of 328,2 Million inhabitants 76,3% is considered as “White alone”. However, this number cannot directly be put into relation with the numbers from 1990, 2000 and 2010 since we are comparing an entire population with a reduced sample representing the 10 largest MA. Therefore, many rural and less economic relevant areas are not considered where presumably more “White” population is living due to a less developed internationalization process etc. Asiatic “Ethnicities” on the other hand clearly show continuous growth amongst the years and present in the sample data.

Thus far for introduction of the analysis section we have been seeing a small overview on the population and ethnicities in general. However, investigating in the field of IS it makes sense to analyze the economic situation and make comparisons amongst MA throughout the years. Therefore, *Table 3* shows as in the Methodology section introduced the crucial Income variable that is yearly “Household Equivalent Income” for the different areas. Clearly amongst all MA Washington, DC wins the race for the highest average income with ca. 50.000 per annum. On the other hand, worst performer overall is Los Angeles-Long Beach, CA with only ca. 34.000 annual earnings. Directly, when seeing LA performing so poorly with the international prestige

that is has with renown commercial markets such as “Hollywood” etc. the question comes to the mind how this situation can arise. Obviously, the answer to that must lay in the high inequality that can be presumed caused by historically sky rocking high unemployment rates that by far exceed the US median (US- Office of Financial Management, 2020). Over the years for all 10 MA, a total number of 79,466 persons did not even acquire one dollar of equivalent income being literally with zero earnings displayed in the data. Finally, the gap between richest and poorest MA is also clearly increasing from only 9.318$ in 1990 to ca. 14.000$ in 2000 and up to almost 25.000$ difference in 2010 indicating increasing Inequality.

TABLE 3: EQUIVALENT INCOME OVERVIEW

Equivalent Income Overview				Year		
Metropolitan area	**Total Population**	**Average Income**	**Rank**	**1990**	**2000**	**2010**
New York, NY-Northeastern NJ	446.718	$ 40.87	4th	$ 30.147	$ 36.815	$ 54.444
Los Angeles-Long Beach, CA	353.487	$ 33.989	10th	$ 22.773	$ 34.495	$ 43.439
Chicago, IL	216.587	$ 38.258	6th	$ 26.837	$ 40.390	$ 46.366
Dallas-Fort Worth, TX	139.091	$ 37.523	7th	$ 25.240	$ 37.740	$ 44.896
Washington, DC/MD/VA	134.204	$ 50.002	1st	$ 32.092	$ 48.460	$ 63.771
San Francisco-Oakland-Vallejo, CA	128.863	$ 43.605	3rd	$ 24.180	$ 47.059	$ 56.955
Philadelphia, PA/NJ	122.988	$ 38.813	5th	$ 27.179	$ 38.180	$ 49.967
Houston-Brazoria, TX	116.400	$ 36.209	8th	$ 23.895	$ 35.406	$ 44.654
Boston, MA/NH	106.199	$ 45.709	2nd	$ 29.923	$ 46.385	$ 57.885
Detroit, MI	100.712	$ 34.542	9th	$ 25.563	$ 38.849	$ 38.905
Total	**1.865.249**	**$ 39.953**	**#**	**$ 26.783**	**$ 40.378**	**$ 50.128**

However, also data is showing that average income throughout time is constantly increasing. It is moreover remarkable that we observe for both extrema (richest and poorest) very little changes and with little exceptions MA perform quite constant.

5.2 Income Segregation over the years

In the following table and graph additionally to the regional distinction now the first hint on the segregation can be determined. The nine subgroups displaying their respective "Ethnicity" can now be put into comparison. Again, keeping in mind that the primary variable that is income has like all the other comparisons been used by the annual Household Equivalent Income.

TABLE 4: INCOME SEGREGATION OVERVIEW

Income Segregation Overview		**Race**								
Metropolitan area	**Average Income**	White	Black	American Ind	Chinese	Japanese	Other Asian / Pacific Islander	Other race , nec	Two Major races	Three or more major races
New York, NY-Northeastern NJ	**$ 40.877**	$ 47.490	$ 25.668	$ 23.512	$ 37.321	$ 57.358	$ 43.167	$ 18.907	$ 33.259	$ 37.980
Los Angeles-Long Beach, CA	**$ 33.990**	$ 39.535	$ 25.409	$ 25.383	$ 38.682	$ 45.289	$ 35.779	$ 18.081	$ 36.093	$ 36.041
Chicago, IL	**$ 38.258**	$ 43.338	$ 22.763	$ 28.620	$ 43.509	$ 47.778	$ 41.934	$ 20.204	$ 38.299	$ 31.425
Dallas-Fort Worth, TX	**$ 37.523**	$ 41.892	$ 23.973	$ 33.972	$ 49.045	$ 48.968	$ 40.553	$ 18.319	$ 33.389	$ 32.113
Washington, DC/MD/VA	**$ 50.003**	$ 57.886	$ 34.035	$ 34.936	$ 53.022	$ 55.811	$ 46.957	$ 26.815	$ 48.034	$ 44.521
San Francisco-Oakland-Vallejo, CA	**$ 43.605**	$ 48.341	$ 27.717	$ 29.966	$ 42.627	$ 48.616	$ 41.450	$ 24.165	$ 45.994	$ 49.775
Philadelphia, PA/NJ	**$ 38.813**	$ 42.594	$ 23.282	$ 25.202	$ 42.078	$ 46.678	$ 39.641	$ 18.243	$ 34.598	$ 33.577
Houston-Brazoria, TX	**$ 36.209**	$ 41.300	$ 24.151	$ 27.211	$ 46.981	$ 44.394	$ 40.595	$ 18.660	$ 34.739	$ 38.131
Boston, MA/NH	**$ 45.709**	$ 48.244	$ 26.970	$ 33.058	$ 46.126	$ 42.059	$ 44.680	$ 20.448	$ 34.807	$ 31.896
Detroit, MI	**$ 34.543**	$ 37.514	$ 22.858	$ 24.093	$ 50.309	$ 49.972	$ 43.624	$ 26.048	$ 29.951	$ 24.281
Total	**$ 39.953**	**$ 44.813**	**$ 25.683**	**$ 28.595**	**$ 44.970**	**$ 48.692**	**$ 41.838**	**$ 20.989**	**$ 36.916**	**$ 35.974**

At the same time, we must not forget that same still not displays the score of the index that follows in the next overview. However, we can see the annual HEI amongst the subgroups which gives an indication of the segregation. For the final scoring with the "Income Ethnicity Segregation Index" not only the average incomes are offset but also multiplied by their weighting regarding total population. The differences get even more vivid when looking at the bar diagrams Figure 1 in black showing the total average for all ten Metropolitan Areas.

Clearly, some ethnicities such as "White", "Chinese", "Japanese" and "other Asian" exceed the average annual HEI that is at approximately 40.000$. However, others such as "Black", "American Indian", "Other Race" and the "Mixed Races" clearly underperform. Moreover, this general trend can be observed as well for the poorest (Figure 2 in orange) as for the richest MA (Figure 3 in green). Even though we see high discrepancies in the dimensions of the average annual HEI, the general constellation of ethnicities being richer and others poorer scatters throughout all the regions.

FIGURE 1: INCOME SEGREGATION – ANNUAL HEI FOR ALL 10 MA

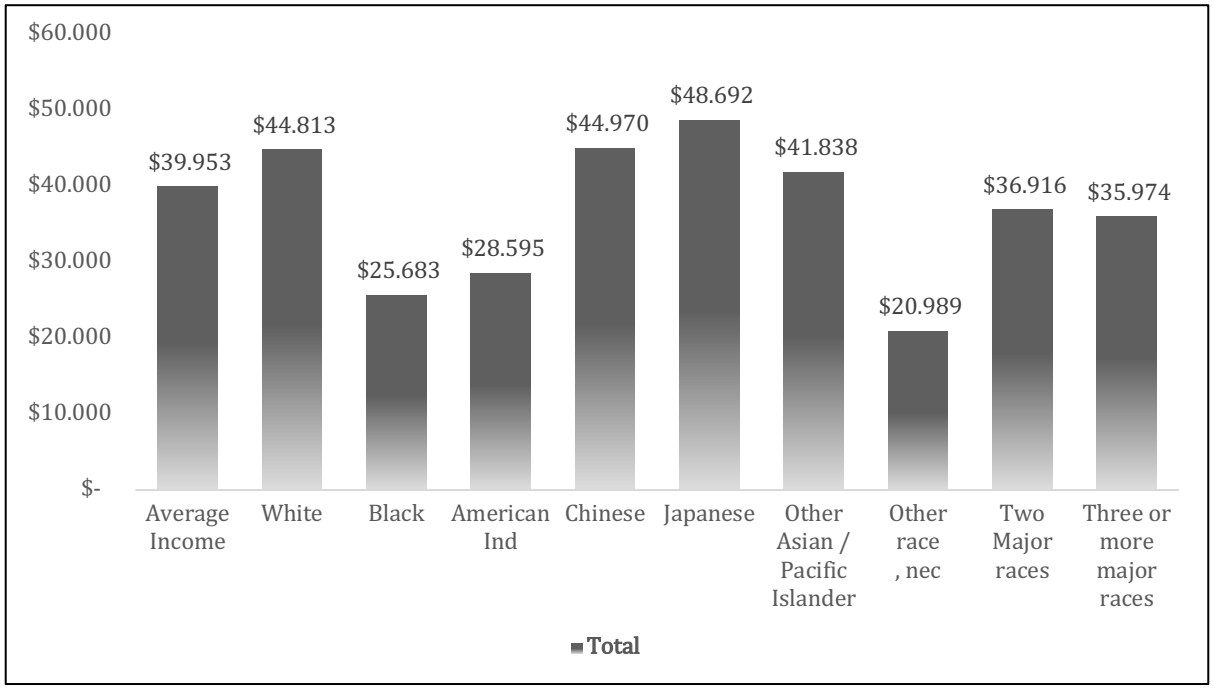


FIGURE 2: INCOME SEGREGATION – ANNUAL HEI FOR POOREST MA

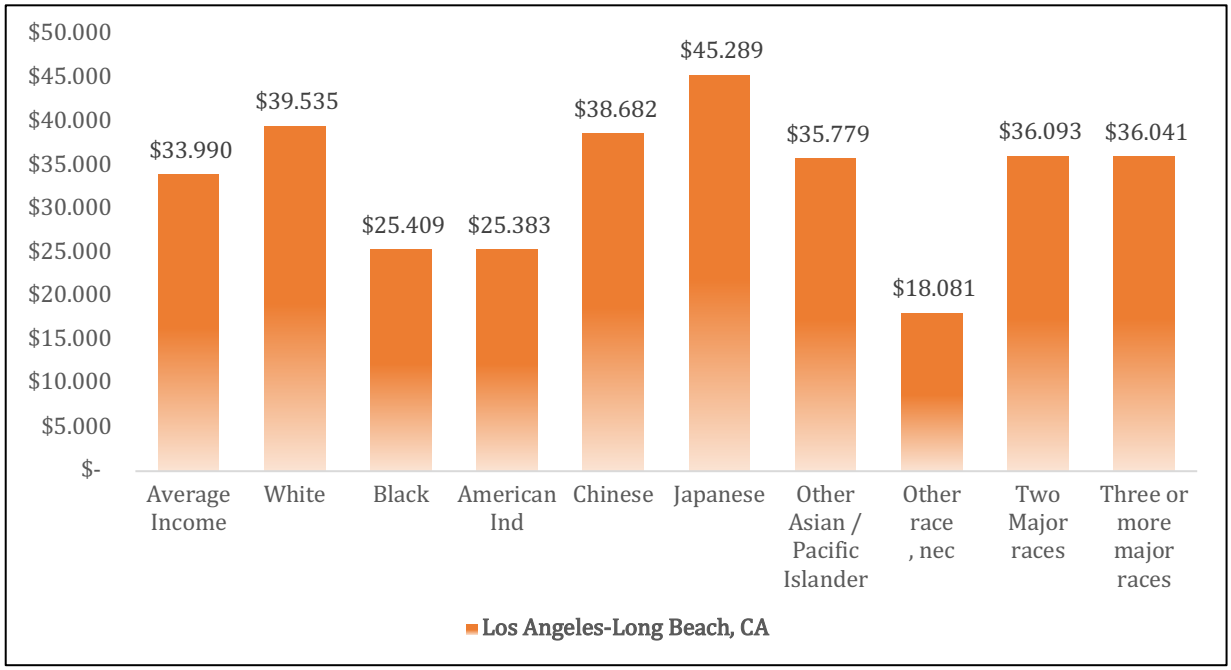
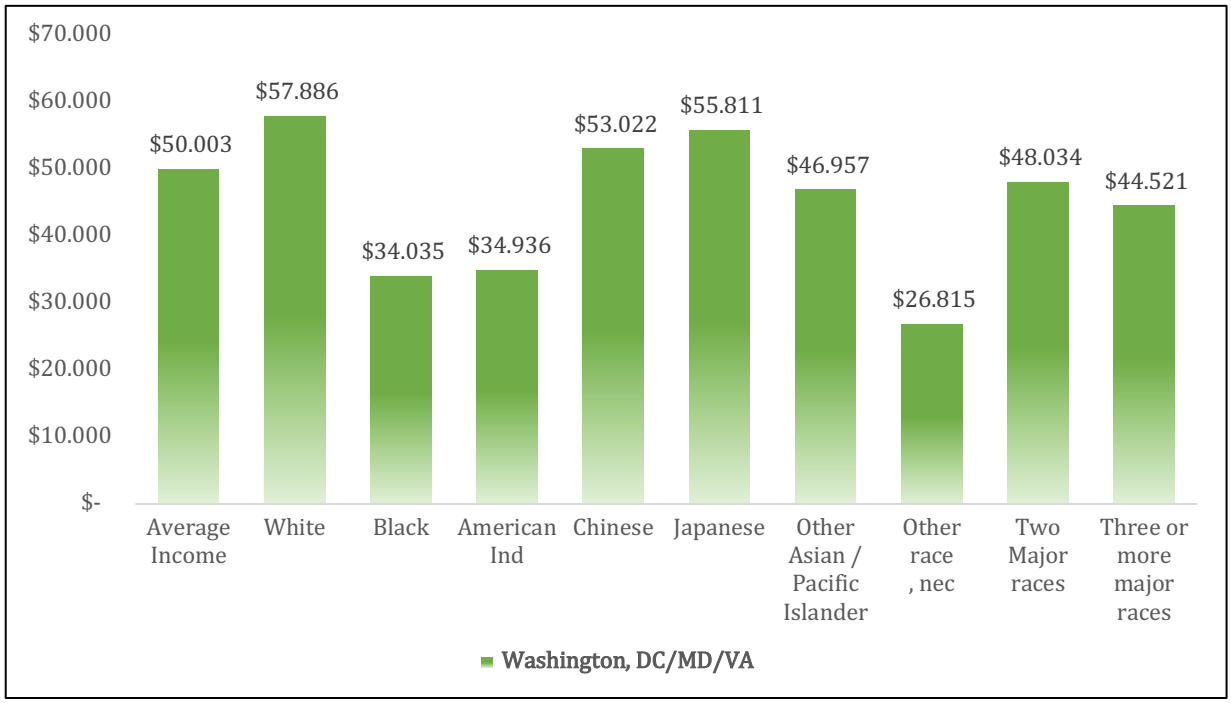


FIGURE 3: INCOME SEGREGATION – ANNUAL HEI FOR RICHEST MA



5.3 The Income Ethnical Segregation Index

In this subsection we will be analyzing the most important unit which is the income segregation score for each subgroup over time. Thus, in *Table 5* can be observed the development:

TABLE 5: INCOME ETHNICAL SEGREGATION INDEX

INDEXING All 10 Metropolitan Areas		Income Ethnical Segregation			
Metropolitan area		**1990-2010**	**1990**	**2000**	**2010**
	New York, NY-Northeastern NJ	4,8%	4,6%	5,5%	5,1%
	Los Angeles-Long Beach, CA	4,2%	3,8%	5,4%	3,9%
	Chicago, IL	4,3%	4,1%	4,3%	4,7%
	Dallas-Fort Worth, TX	3,7%	4,2%	4,4%	3,9%
	Washington, DC/MD/VA	3,1%	3,0%	3,3%	3,9%
	San Francisco-Oakland-Vallejo, CA	2,4%	1,9%	3,2%	2,7%
	Philadelphia, PA/NJ	4,1%	3,9%	4,1%	4,8%
	Houston-Brazoria, TX	3,6%	4,7%	4,3%	2,7%
	Boston, MA/NH	2,1%	2,0%	2,2%	3,0%
	Detroit, MI	2,4%	2,8%	2,1%	2,9%
Total		**3,8%**	**3,7%**	**4,4%**	**3,9%**

New York, NY-North-eastern NJ showing the poorest score with an average score of 4,8% of Income Ethnical Segregation. On the other end, Boston is showing the best score with the lowest segregation of only 2,1%. Remarkable is furthermore also the fact that neither of both extremes for the richest and poorest MA (Washington/ LA) show severe scores for the segregation. However, we see that richer areas tend to show better results regarding the segregation score and vice versa with New York being the exception since it is at the 4th place of the 10 biggest richest MA areas. The total "Income Ethnical Segregation" score for the whole 10 biggest US-Metropolitan Areas was 3,8%. If now we consider the trend and the development that the score is showing, we must notice a slight decrease over the whole period. The heaviest increase was observed for the decade of 1990 until 2000 where segregation increased by 0,7%. In the following years a decrease was recorded to 3,9% however never reaching the 1990 level. Unfortunately, more recent data was not available from the data source making future

predictions at this state impossible. As mentioned earlier this score can be seen as the "Between" as it is stating the segregation of inequality amongst subgroups for the MA. If we now want to integrate that score into the whole Inequality Ranking of the Entropy Index, we must consequently also take the "Within" into account.

5.4 The Total Income Ethnical Inequality – Entropy Index

In the *Table 6* both, the "Between" Segregation as well as the "Within" inequality are shown:

TABLE 6: TOTAL INCOME ETHNICAL INEQUALITY

INDEXING All 10 Metropolitan Areas		Entropy Index				Total Population Share/ "Within" + "Between"
	Ethnicity	**1990-2010**	**1990**	**2000**	**2010**	**Percent (1990-2010)**
	White	**0,563**	**0,446**	**0,537**	0,619	66,85
	Black/African American	**0,801**	0,725	0,811	0,775	13,79
	American Indian, Alaska Native or Latino	0,674	0,547	0,758	**0,611**	0,38
	Chinese	0,782	0,566	0,723	**0,867**	2,40
	Japanese	**0,943**	**0,879**	**1,023**	0,799	0,46
	Other Asian or Pacific Islander	0,700	0,673	0,663	0,665	5,32
	Other race, nec	0,738	0,728	0,733	0,689	8,56
	Two major races	0,732	X	0,765	0,659	2,10
	Three or more major races	0,643	X	0,715	**0,544**	0,14
						100%
1990-2010	Within	**X**	X	X	X	0,641
	Between	**X**	X	X	X	**0,038**
	Population	**0,675**	X	X	X	0,675
1990	Within	**X**	X	X	X	0,535
	Between	**X**	X	X	X	**0,037**
	Population	**0,572**	X	X	X	0,572
2000	Within	**X**	X	X	X	0,625
	Between	**X**	X	X	X	**0,044**
	Population	**0,668**	X	X	X	0,668
2010	Within	**X**	X	X	X	0,665
	Between	**X**	X	X	X	**0,039**
	Population	**0,699**	X	X	X	0,699

To understand all the contents of the table, we will begin with the column on the right. In the top, the subgroup weights for the total population are represented. In other words, 67% of the population was represented by the "white" subgroup as mentioned earlier etc. There below, we see the most important three scores for the whole period as well as the three years 1990, 2000 and 2010. The first one the "Within" that states the sum of Inequality within the nine subgroups. The bold written figures below are the "Between" Segregation scores that we saw already in the previous table. The "Total Income Ethnical Inequality" is as explained earlier the sum of the "Within" and the "Between" that can be observed on each bottom for the various periods as well as on the left-hand side as it is the most important stat that can now be reviewed. We see that for "Total Income Ethnical Inequality" the score is constantly rising from 0,572 in 1990 to 0,668 in 2000 and finally remaining more less at same level at 0,699 in 2010. Same goes hand in hand with the observation that we made earlier for the gap between rich and poor that is over the years significantly increasing and the "Between" score that we saw increasing from 3,7% to 3,9%. For our empirical evidence clearly inequality and segregation show that both criterions follow the same trend observed over time. If we have a further look on the subgroups/ ethnicities and corresponding scores for the Entropy Index, we can make further assumptions. "Total Income Ethnical Inequality" for "Whites" is with 0,563 much lower than for "Blacks" with 0,801. However, statistics show that for subgroups with small relative weight especially visible for "Japanese" the EI scores tend to be much larger. Same can be explained with the mathematical phenomenon that variance flattens out the greater the sample size as estimators gets more precise with more data. Therefore, the "Entropy Index" can better be trusted for the whole population as the scores are multiplied by their respective weighting.

6. Conclusion

We saw throughout the thesis that with the construction of the "Income Ethnical Segregation Index" the critical variable "Income" can indeed be used as primary unit and subcategorized with the secondary unit "race", which allowed for a profound empirical time series analysis. Clearly, similarities on trends for segregation can be observed when compared with developments for overall inequality. The index constructed could furthermore be used to the entire US, other countries or even continents as the principle remains the same and allows for different constellations of subgroups. Same must on the other hand imply that variables are consistently available and time data is required to put numbers into perspective, which for the most part is extremely difficult and exclusive. Moreover, a common comparison basis of the economical unit, which was the HEI is of crucial importance and contributes to a more consistent estimation of welfare conclusions drawn. A drastic reduction of observations was necessary not to completely spread the scope of this framework and the population overview, financial and racial inequality, as well as results for the segregation highlighted for the 10 biggest US-Metropolitan Areas throughout from 1990 until 2010. Furthermore, the population was ethnically divided into nine subgroups that represented a respective unit for the "between" segregation that was applied with the help of the "Income Ethnical Segregation Index" constructed. Thus, consequently results for different subgroups as well as regional distinctions could be used for policy making by focusing especially on MA or ethnicities that show high tendencies for segregation and inequality. We saw that indeed a correlation of poorer MA based on average annual HEI can but not must indicate a higher inequality and segregation, displaying the old problem of the increasing gap between rich and poor that especially showed in the MA of Los Angeles-Long Beach, CA in the data. However, in order to best improve welfare for the whole economy, poor individuals should be tried to be subsidized regardless of ethnical affiliation or regional borders. For the largest MA over 80.000 individuals representing 1,5% of total population were with effectively zero "Household Equivalent Income", which clearly

contributes to the inequality observed. Even though overall segregation increased from 1990 until 2010 by 0,2%, a clear improvement was observed in the last decade considered as "Income Ethnical Segregation" decreased by 0,5%. Also looking at the "Total Income Ethnical Inequality" expressed by the combined sum of the "Within" and "Between" component showed a clear stagnation of the previous negative trend giving hope for future positive developments.

Appendix

List of Appendices

Appendix 1: Formula for Equivalent Scale

Formula 4: Equivalent Scale

Step 1: "Find out number of Household Members e.g. 3."

Step 2: "Create Variable that displays the square root."

$$e.g.SR = \sqrt{3}$$

Step 3: "Divide each Households Total Income by the SR."

$$\boldsymbol{HEI} = \frac{\boldsymbol{Household\ Total\ Income\ (FTI)}}{\sqrt{\boldsymbol{Number\ of\ Household\ members\ e.g.\ 3}}}$$

→ Same is the unit selected for the further analysis of Economic Income Segregation

Source: Modified by Author

Appendix 2: Formula 4: School Separation Index

$$\boldsymbol{SSI(x)} = \sum \frac{n_c}{n_x} * ln\left(\frac{\mu_x}{\mu_c}\right)$$

Where x is the whole population of the respective area; c is each subcategory, for instance the "school 1", "school 2" et cetera. Furthermore, n_x is the total sum of population in the corresponding area and μ_x its total mean income. On the other hand, μ_c is the mean income of the respective subcategory e.g., "school 1". For instance, if we analyse 50 different subgroups/ schools we will have, n_1: pupils of "school 1", …, until we have the last n_{50}: pupils of "school 50" using respective sample weights. Furthermore, μ_1: is the mean income of "school 1" individuals, et cetera.

Source: Casilda Lasso de la Vega & Oscar Volij (2020)

The below table shows the Index results for the richest US-Metropolitan area that is Washington, DC.

APPENDIX 3: INDEX FOR WASHINGTON, DC

INDEXING Washington, DC RICHEST AREA		Entropy Index				Total Population Share / "Within" + "Between" Percent (1990-2010)
	Ethnicity	**1990-2010**	**1990**	**2000**	**2010**	
	White	0,452	0,331	0,407	**0,513**	**62,07**
	Black/African American	0,609	0,532	0,599	**0,585**	**24,46**
	American Indian, Alaska Native or Latino	0,540	0,278	0,426	0,703	**0,34**
	Chinese	0,466	0,408	0,398	0,472	**1,56**
	Japanese	1,015	0,839	1,785	0,295	**0,24**
	Other Asian or Pacific Islander	0,620	0,515	0,557	0,636	**5,63**
	Other race, nec	0,533	0,621	0,523	0,509	**3,33**
	Two major races	0,660	X	0,653	0,643	**2,17**
	Three or more major races	0,326	X	0,657	0,213	**0,19**
						100%
1990- 2010	Within	**X**	X	X	X	0,516
	Between	**X**	X	X	X	**0,031**
	Population	**0,543**	X	X	X	0,543
1990	Within	**X**	X	X	X	0,404
	Between	**X**	X	X	X	**0,030**
	Population	**0,433**	X	X	X	0,433
2000	Within	**X**	X	X	X	0,482
	Between	**X**	X	X	X	**0,033**
	Population	**0,514**	X	X	X	0,514
2010	Within	**X**	X	X	X	0,545
	Between	**X**	X	X	X	**0,039**
	Population	**0,580**	X	X	X	0,580

Source: Modified by Author

Lowest Total Income Ethnical Inequality.

However, not the lowest score in the Segregation (Boston is still better).

The below table shows the Index results for the poorest US-Metropolitan area that is Los Angeles-Long Beach, CA.

APPENDIX 4: INDEX FOR LOS ANGELES

INDEXING Los Angeles-Long Beach, CA POOREST AREA		Entropy Index				Total Population Share/ "Within" + "Between" Percent (1990-2010)
	Ethnicity	**1990-2010**	**1990**	**2000**	**2010**	
	White	**0,683**	**0,590**	0,669	0,695	56,83
	Black/African American	**0,795**	0,709	0,856	0,673	6,92
	American Indian, Alaska Native or Latino	0,683	0,448	0,805	0,643	0,58
	Chinese	0,859	0,684	0,683	1,036	3,42
	Japanese	**0,728**	**0,527**	0,693	0,908	1,32
	Other Asian or Pacific Islander	0,746	0,681	0,692	0,736	8,77
	Other race, nec	0,707	0,648	0,710	0,672	19,1
	Two major races	0,691	X	0,771	0,536	2,87
	Three or more major races	0,728	X	1,001	0,520	0,19
						100%
1990- 2010	Within	**X**	X	X	X	0,709
	Between	**X**	X	X	X	**0,042**
	Population	**0,745**	X	X	X	0,745
1990	Within	**X**	X	X	X	0,619
	Between	**X**	X	X	X	**0,038**
	Population	**0,656**	X	X	X	0,656
2000	Within	**X**	X	X	X	0,703
	Between	**X**	X	X	X	**0,054**
	Population	**0,754**	X	X	X	0,754
2010	Within	**X**	X	X	X	0,700
	Between	**X**	X	X	X	**0,039**
	Population	**0,733**	X	X	X	0,733

Source: Modified by Author

The below table shows the Index results for the US-Metropolitan area:

APPENDIX 5: INDEX FOR NEW YORK, NY-NORTHEASTERN NJ

INDEXING New York, NY-Northeastern NJ		Entropy Index				Total Population Share/ "Within" + "Between"
	Ethnicity	**1990-2010**	**1990**	**2000**	**2010**	**Percent (1990-2010)**
	White	0,615	0,492	0,629	0,644	65,54
	Black/African American	0,828	0,720	0,920	0,761	16,1
	American Indian, Alaska Native or Latino	0,628	0,436	0,769	0,409	0,28
	Chinese	0,760	0,545	0,855	0,744	2,89
	Japanese	1,295	1,354	1,682	0,781	0,27
	Other Asian or Pacific Islander	0,786	0,699	0,756	0,779	4,84
	Other race, nec	0,856	0,881	0,856	0,772	7,85
	Two major races	0,860	X	0,861	0,772	2,1
	Three or more major races	0,653	X	0,642	0,596	0,13
						100%
1990- 2010	Within	X	X	X	X	0,705
	Between	X	X	X	X	0,048
	Population	**0,749**	X	X	X	0,749
1990	Within	X	X	X	X	0,587
	Between	X	X	X	X	0,046
	Population	0,632	X	X	X	0,632
2000	Within	X	X	X	X	0,736
	Between	X	X	X	X	0,055
	Population	0,791	X	X	X	0,791
2010	Within	X	X	X	X	0,699
	Between	X	X	X	X	0,051
	Population	0,434	X	X	X	0,743

Source: Modified by Author

Highest score for segregation and inequality for all 10 MA.

The below table shows the Index results for the US-Metropolitan area:

APPENDIX 6: INDEX FOR CHICAGO, IL

INDEXING Chicago, IL		Entropy Index				Total Population Share/ "Within" + "Between"
	Ethnicity	**1990-2010**	**1990**	**2000**	**2010**	**Percent (1990-2010)**
	White	0,486	0,368	0,448	0,559	70,81
	Black/African American	0,852	0,743	0,448	0,860	16,41
	American Indian, Alaska Native or Latino	0,550	0,547	0,660	0,365	0,26
	Chinese	0,777	0,522	0,778	0,809	0,84
	Japanese	0,830	0,719	0,693	0,996	0,2
	Other Asian or Pacific Islander	0,554	0,423	0,595	0,541	3,26
	Other race, nec	0,636	0,623	0,725	0,546	6,69
	Two major races	0,687	X	0,673	0,665	1,45
	Three or more major races	0,554	X	0,598	0,482	0,08
						100%
1990- 2010	Within	X	X	X	X	0,583
	Between	X	X	X	X	0,043
	Population	**0,623**	X	X	X	0,623
1990	Within	X	X	X	X	0,473
	Between	X	X	X	X	0,041
	Population	0,515	X	X	X	0,515
2000	Within	X	X	X	X	0,571
	Between	X	X	X	X	0,043
	Population	0,616	X	X	X	0,616
2010	Within	X	X	X	X	0,620
	Between	X	X	X	X	0,047
	Population	0,658	X	X	X	0,658

Source: Modified by Author

The below table shows the Index results for the US-Metropolitan area:

APPENDIX 7: INDEX FOR DALLAS-FORT WORTH, TX

INDEXING Dallas-Fort Worth, TX		Entropy Index				Total Population Share/ "Within" + "Between"
	Ethnicity	**1990-2010**	**1990**	**2000**	**2010**	**Percent (1990-2010)**
	White	0,499	0,388	0,537	0,619	72,59
	Black/African American	0,721	0,743	0,811	0,775	12,79
	American Indian, Alaska Native or Latino	0,633	0,637	0,758	0,611	0,54
	Chinese	0,809	0,559	0,723	0,867	0,62
	Japanese	0,299	0,265	1,023	0,799	0,1
	Other Asian or Pacific Islander	0,582	0,967	0,663	0,665	3,45
	Other race, nec	0,665	0,605	0,733	0,689	7,95
	Two major races	0,685	X	0,765	0,659	1,85
	Three or more major races	1,123	X	0,715	0,544	0,1
						100%
1990- 2010	Within	X	X	X	X	0,560
	Between	X	X	X	X	0,037
	Population	**0,593**	X	X	X	0,593
1990	Within	X	X	X	X	0,475
	Between	X	X	X	X	0,042
	Population	0,518	X	X	X	0,518
2000	Within	X	X	X	X	0,625
	Between	X	X	X	X	0,044
	Population	0,668	X	X	X	0,668
2010	Within	X	X	X	X	0,665
	Between	X	X	X	X	0,039
	Population	0,699	X	X	X	0,699

Source: Modified by Author

The below table shows the Index results for the US-Metropolitan area:

APPENDIX 8: INDEX FOR SAN FRANCISCO-OAKLAND-VALLEJO,CA

INDEXING San Francisco-Oakland-Vallejo, CA		Entropy Index				Total Population Share/ "Within" + "Between"
	Ethnicity	**1990-2010**	**1990**	**2000**	**2010**	**Percent (1990-2010)**
	White	0,591	0,424	0,526	0,631	60,97
	Black/African American	0,817	0,814	0,721	0,731	8,61
	American Indian, Alaska Native or Latino	0,599	0,432	0,684	0,560	0,54
	Chinese	0,677	0,482	0,548	0,792	7,9
	Japanese	0,805	0,569	0,829	0,782	1,03
	Other Asian or Pacific Islander	0,542	0,398	0,571	0,475	10,65
	Other race, nec	0,596	0,500	0,538	0,641	6,64
	Two major races	0,642	X	0,592	0,663	3,33
	Three or more major races	0,656	X	0,835	0,469	0,33
						100%
1990- 2010	Within	X	X	X	X	0,621
	Between	X	X	X	X	0,024
	Population	**0,645**	X	X	X	0,645
1990	Within	X	X	X	X	0,480
	Between	X	X	X	X	0,019
	Population	0,499	X	X	X	0,499
2000	Within	X	X	X	X	0,563
	Between	X	X	X	X	0,032
	Population	0,595	X	X	X	0,595
2010	Within	X	X	X	X	0,635
	Between	X	X	X	X	0,027
	Population	0,661	X	X	X	0,661

Source: Modified by Author

The below table shows the Index results for the US-Metropolitan area:

APPENDIX 9: INDEX FOR PHILADELPHIA, PA/NJ

INDEXING Philadelphia, PA/NJ		Entropy Index				Total Population Share/ "Within" + "Between"
	Ethnicity	**1990-2010**	**1990**	**2000**	**2010**	**Percent (1990-2010)**
	White	0,506	0,392	0,453	0,580	77,2
	Black/African American	0,830	0,713	0,707	0,966	15,78
	American Indian, Alaska Native or Latino	0,693	1,120	0,340	0,356	0,18
	Chinese	1,139	0,544	0,901	1,392	0,71
	Japanese	2,721	5,162	1,411	0,414	0,08
	Other Asian or Pacific Islander	0,899	1,355	0,693	0,806	2,63
	Other race, nec	1,124	1,499	0,861	0,979	2,14
	Two major races	0,708	X	0,669	0,717	1,2
	Three or more major races	0,647	X	0,317	0,600	0,09
						100%
1990- 2010	Within	X	X	X	X	0,625
	Between	X	X	X	X	0,041
	Population	**0,669**	X	X	X	0,669
1990	Within	X	X	X	X	0,520
	Between	X	X	X	X	0,039
	Population	0,563	X	X	X	0,563
2000	Within	X	X	X	X	0,537
	Between	X	X	X	X	0,041
	Population	0,582	X	X	X	0,582
2010	Within	X	X	X	X	0,714
	Between	X	X	X	X	0,048
	Population	0,764	X	X	X	0,764

Source: Modified by Author

The below table shows the Index results for the US-Metropolitan area:

APPENDIX 10: INDEX FOR HOUSTON-BRAZORIA, TX

INDEXING (THEIL-BETWEEN) Houston-Brazoria, TX		Entropy Index				Total Population Share
	Ethnicity	**1990-2010**	**1990**	**2000**	**2010**	**Percent (1990-2010)**
	White	0,597	0,453	0,579	0,653	66,39
	Black/African American	0,732	0,741	0,713	0,653	15,65
	American Indian, Alaska Native or Latino	0,928	1,074	0,850	0,925	0,47
	Chinese	0,740	0,520	0,795	0,697	1,23
	Japanese	0,442	0,180	0,335	0,618	0,1
	Other Asian or Pacific Islander	0,637	0,872	0,497	0,598	4,38
	Other race, nec	0,664	0,666	0,657	0,599	10,03
	Two major races	0,703	X	0,780	0,563	1,66
	Three or more major races	0,254	X	0,189	0,488	0,1
						100%
1990- 2010	Within	X	X	X	X	0,635
	Between	X	X	X	X	0,036
	Population	**0,665**	X	X	X	0,665
1990	Within	X	X	X	X	0,544
	Between	X	X	X	X	0,047
	Population	0,590	X	X	X	0,590
2000	Within	X	X	X	X	0,614
	Between	X	X	X	X	0,043
	Population	0,656	X	X	X	0,656
2010	Within	X	X	X	X	0,646
	Between	X	X	X	X	0,027
	Population	0,667	X	X	X	0,667

Source: Modified by Author

The below table shows the Index results for the US-Metropolitan area:

APPENDIX 11: BOSTON, MA/ NH

INDEXING (THEIL-BETWEEN) Boston, MA/NH		Entropy Index				Total Population Share
	Ethnicity	**1990-2010**	**1990**	**2000**	**2010**	**Percent (1990-2010)**
	White	0,541	0,416	0,503	0,609	83,95
	Black/African American	0,708	0,601	0,823	0,661	5,8
	American Indian, Alaska Native or Latino	1,161	1,328	0,490	1,779	0,19
	Chinese	1,006	0,636	0,488	1,348	1,97
	Japanese	2,844	4,046	2,701	1,242	0,21
	Other Asian or Pacific Islander	0,951	0,988	0,812	0,844	2,93
	Other race, nec	0,849	0,964	0,789	0,776	3,1
	Two major races	0,643	X	0,796	0,542	1,76
	Three or more major races	0,330	X	0,344	0,298	0,1
						100%
1990- 2010	Within	X	X	X	X	0,600
	Between	X	X	X	X	0,021
	Population	**0,620**	X	X	X	0,620
1990	Within	X	X	X	X	0,472
	Between	X	X	X	X	0,020
	Population	0,493	X	X	X	0,493
2000	Within	X	X	X	X	0,559
	Between	X	X	X	X	0,022
	Population	0,580	X	X	X	0,580
2010	Within	X	X	X	X	0,656
	Between	X	X	X	X	0,030
	Population	0,685	X	X	X	0,685

Source: Modified by Author

Lowest Segregation but not lowest Inequality.

The below table shows the Index results for the US-Metropolitan area:

APPENDIX 12: DETROIT, MI

INDEXING (THEIL-BETWEEN) Detroit, MI		Entropy Index				Total Population Share
	Ethnicity	**1990-2010**	**1990**	**2000**	**2010**	**Percent (1990-2010)**
	White	0,456	0,337	0,405	0,578	75,04
	Black/African American	0,831	0,785	0,770	0,856	19,73
	American Indian, Alaska Native or Latino	0,603	0,390	1,218	0,217	0,38
	Chinese	0,621	0,515	0,914	0,278	0,42
	Japanese	0,642	1,169	0,542	0,121	0,16
	Other Asian or Pacific Islander	0,577	0,588	0,699	0,487	1,92
	Other race, nec	0,899	0,390	1,295	0,803	0,9
	Two major races	0,596	X	0,613	0,557	1,37
	Three or more major races	0,487	X	0,327	0,514	0,09
						100%
1990- 2010	Within	X	X	X	X	0,559
	Between	X	X	X	X	0,024
	Population	**0,583**	X	X	X	0,583
1990	Within	X	X	X	X	0,451
	Between	X	X	X	X	0,028
	Population	0,479	X	X	X	0,479
2000	Within	X	X	X	X	0,513
	Between	X	X	X	X	0,021
	Population	0,535	X	X	X	0,535
2010	Within	X	X	X	X	0,651
	Between	X	X	X	X	0,029
	Population	0,676	X	X	X	0,676

Source: Modified by Author

References

Adams, John S., Barbara J. Van Drasek & Eric G. Phillips (1999) "Metropolitan Area Definition In The United States", Urban Geography, 20:8, 695-726

Deltas, George; The Small-Sample Bias of the Gini Coefficient: Results and Implications for Empirical Research. The Review of Economics and Statistics 2003; 85 (1): 226–234.

Gini, C. (1912) Variabilita e mutabilita. Reprinted in E. Pizetti & T. Salvemini, eds., Memorie di Metodologia Statistica. Rome: Libreria Erendi Virgilio Veschi, 1955.

Jargowsky, P. A., "Take the Money and Run: Economic Segregation in US Metropolitan Areas," American Sociological Review 61 (1996), 984–98.

Kim, J., and P. A. Jargowsky, "The Gini coefficient and segregation on a continuous variable," in Y. Flückiger, S. F. Reardon, and J. Silber, eds., Occupational and Residential Segregation (Research on Economic Inequality, Vol. 17) (Bingley, United Kingdom: Emerald Group Publishing Limited, 2005), 57–70.

Lasso de la Vega, C. and Volij, O, "The Measurement of Income Segregation," International Economic Review, 61(4): 1479-1500, 2020.

Lynch, John W., et al. "Income inequality and mortality in metropolitan areas of the United States." American journal of public health 88.7 (1998): 1074-1080.

Massey, D. S., and N. A. Denton, "The Dimensions of Residential Segregation," Social Forces 67 (1988), 281–315.

Internet Sources/ Websites

OECD on Inequality. 2015. "Income Inequality (Accessed July 27, 2021) https://data.oecd.org/inequality/income-inequality.htm

Office of Financial Management USA - Unemployment Rates Info USA - (Accessed August 4th, 2021)

https://ofm.wa.gov/washington-data-research/statewide-data/washington trends/economic-trends/unemployment-rates

Stanford Center on Poverty and Inequality. 2021. *"Income Segregation in the United States' Largest Metropolitan Areas"* (Accessed July 25, 2021). https://inequality.stanford.edu/income-segregation-maps

Trading Economics - Los Angeles Unemployment Rates Info (Accessed August 30th, 2021)

https://tradingeconomics.com/united-states/unemployment-rate-in-los-angeles-county-ca-percent-m-nsa-fed-data.html