WHAT A POLYGON CAN'T TELL YOU: SHIFTING FROM AREA-LEVEL TO POINT-LEVEL INVESTIGATION OF RESIDENTIAL SEGREGATION PATTERNS

Ross William Fineman

A DISSERTATION

in

Demography

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2017

Supervisor of Dissertation

____

Jason Schnittker, Professor of Sociology

Graduate Group Chairperson

____

Hans-Peter Kohler, Frederick J. Warren Professor of Demography

Dissertation Committee

Jason Schnittker, Professor of Sociology

Douglas Ewbank, Research Professor of Sociology

C. Dana Tomlin, Professor of Landscape Architecture

WHAT A POLYGON CAN'T TELL YOU: SHIFTING FROM AREA-LEVEL TO POINT-LEVEL INVESTIGATION OF RESIDENTIAL SEGREGATION PATTERNS

DEDICATION

To taking the open three…

ACKNOWLEDGEMENTS

I could not have done this work alone and therefore would like to give acknowledgement to those who helped me along the way.

I would like to thank my faculty advisors Jason Schnittker, Douglas Ewbank, and C. Dana Tomlin for all their academic guidance throughout this process. I would also like to thank John Logan and Hans-Peter Kohler for their assistance.

I would like to thank my parents Ed and Nancy Fineman and brother Grant Fineman for their support.

I would like to thank Sherelle Ferguson for being there for me every step of the way. There is no limit to how much I love you.

I would like to thank Nikkil Sudharsanan for holding the light for me when I could not see the end of the tunnel.

And lastly, I would like to thank anyone who was foolish enough to ask me what I study and bold enough to ask a follow up question.

ABSTRACT

WHAT A POLYGON CAN'T TELL YOU: SHIFTING FROM AREA-LEVEL TO POINT-LEVEL INVESTIGATION OF RESIDENTIAL SEGREGATION PATTERNS

Ross William Fineman

Jason Schnittker

The study of segregation is essential for understanding how place influences life outcomes. However, traditional segregation indices rely heavily on the use of areal units for calculation, which risks introducing both measurement and interpretation error. Researchers suggest that individual-level data avoids many of the problems facing traditional area-level indices. In this Dissertation, I use the recent release of the complete 1940 Census to investigate the potential problems with measuring segregation with areal units and develop a new method for measuring segregation at the individual level. In Chapter 1, I investigate the potential impact the modifiable areal unit problem (MAUP) may have on accurately measuring segregation when using areal unit indices. In Chapter 2, I develop a new measure of segregation, the Shortest Path Isolation (SPI) index, which captures the degree of racial isolation from the perspective of what an individual would experience. Using the SPI index developed in Chapter 2, Chapter 3 investigates how individual-level racial isolation in 1940 West Philadelphia is associated with access to neighborhood resources by race. Given that our understanding is only as good as our measurement, it is imperative that our measures accurately reflect our perceptions of segregation.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

CHAPTER 1: CHOPPED AND SCREWED: THE IMPACT OF THE MODIFIABLE AREAL UNIT PROBLEM ON MEASURING SEGREGATION

Introduction:

Accurate measurement is essential for rigorous science (Diez-Roux & Mair, 2010; Taeuber & Taeuber, 1976). Embedded in segregation measures are normative assumptions about how society should operate, establishing a framework through which we attempt to understand underlying structures (Johnston, Poulsen, & Forrest, 2010; Peach, 2009; Sundstrom, 2004). However, most segregation analysis has relied on arbitrary areal units, such as census tracts, for calculation (Flowerdew, 2011; Holt, Steel, & Tranmer, 1996; Manley, Flowerdew, & Steel, 2006). This reliance potentially risks introducing measurement error, such as the modifiable areal unit problem (MAUP) (Openshaw, 1989). To what extent the use of arbitrary areal units affect calculation is up for debate (Cohn & Jackman, 2011; Flowerdew, 2011; Fotheringham, 1998; Openshaw, 1984). Therefore, before studying the causes or consequences of segregation, it is first imperative to ensure that our indices accurately measure segregation.

Literature Review:

The modifiable areal unit problem (MAUP) is one of the largest challenges to accurately measuring segregation. Closely related to the ecological fallacy problem, the MAUP refers to the measurement error that arises from aggregating individual-level processes using arbitrarily defined boundaries (Butkiewicz et al., 2010; Fogarty, 2010; Kramer, Cooper, Drews-Botsch, Waller, & Hogue, 2010; B. A. Lee et al., 2008; Openshaw, 1989; Wong, 2009). The MAUP consists of two separate sub-problems, one statistical, the scaling effect, and one geographic problem, the zoning effect (Amrhein & Flowerdew, 1992; Dark & Bram, 2007; Fogarty, 2010; Openshaw, 1989; Openshaw & Taylor, 1979;

Tobler, 1990; Wong, 2004). The scaling effect refers to the fact that results change as the size, and consequently the number, of areal units change (Dumedah, Schuurman, & Yang, 2008; Gehlke & Biehl, 1934; Openshaw, 1989; Pietrzak, 2014a; Wong, 2004). The zoning effect, sometimes known as the aggregation effect, refers to the fact that the results change as the areal unit shape changes, regardless if the total area of the unit changes or not (Dumedah et al., 2008; Goodchild, 2011; Holt, Steel, Tranmer, & Wrigley, 1996; Pietrzak, 2014b; Wong, 2004). One example of the zoning effect can be seen in gerrymandering processes that can make the population and racial composition of an area appear differently depending on how boundaries are drawn (Duncan & Duncan, 1955). While researchers have debated over whether changing the size or changing the shape has a larger impact on measurement error, both are potentially significant (Briant, Combes, & Lafourcade, 2010; Cockings & Martin, 2005; Dumedah et al., 2008; Fotheringham & Wong, 1991; Goodchild, 2011; Guo & Bhat, 2004; Openshaw & Rao, 1995; Tobler, 1990). Together, the MAUP poses a considerable problem for segregation analysis by potentially generating different conclusions depending on the areal unit. Given its importance and potential impact on segregation analysis, this article describes how the MAUP may impact the measurement of segregation.

The MAUP is pervasive in almost all types of analysis involving spatial units (Openshaw, 1996). Researchers have found that the MAUP can lead to inconsistent results in univariate, bivariate, and multivariate analysis, as well as spatial interaction modeling and location-allocation modeling (Charlton, 2008; Fotheringham & Wong, 1991; Guo & Bhat, 2004). The MAUP is also prevalent in segregation analysis. Segregation scholars have known for a long time that results are sensitive to the choice of areal unit (Duncan &

Duncan, 1955; Jahn, Schmid, & Schrag, 1947; Jakubs, 1981; B. A. Lee et al., 2008; Massey & Denton, 1993; Östh, Clark, & Malmberg, 2015; Reardon et al., 2008; Taeuber & Taeuber, 1965; White, 1983; Wong, 1997, 2002b, 2003a, 2003b, 2004, 2009, 1993; Wong, Lasus, & Falk, 1999).

While researchers from many different fields have proposed solutions to the MAUP, none are universal (Clark & Avery, 1976; Dark & Bram, 2007; Fotheringham, 1989, 1998; Fotheringham & Wong, 1991; Guo & Bhat, 2004; Openshaw, 1984, 1989, 1996; Páez & Scott, 2004; Zhang & Kukadia, 2005). Segregation scholars have also attempted to deal with the MAUP in a several different ways.

One way is to ignore the problem and trust that results are insensitive to areal unit choice. This tactic has probably been the most common approach in all disciplines, including segregation studies (Openshaw, 1989). However, it would be naïve to continue this assumption given that enough empirical studies have shown that segregation values are not the same at all scales (Östh et al., 2015; Reardon et al., 2009; Wong, 1997, 2004; Wong et al., 1999). Therefore, this approach is not a viable option.

Another way to deal with the MAUP is to add contextual meaning and sociological theory to the choice of areal unit. The MAUP only exists because the units of aggregation are arbitrary; if the areal units employed are genuinely meaningful, then the MAUP goes away (Openshaw, 1989). Many scholars argue that segregation operates at different scales (B. A. Lee et al., 2008; J. R. Logan & Bellman, 2016; Manley et al., 2006; Östh et al., 2015; Reardon et al., 2009; Reardon & O'Sullivan, 2004). For example, Logan & Bellman (2016) measured segregation at multiple levels, from the household to the ward, to investigate the range of segregation scales. Therefore, many researchers treat the census tract as a contextually meaningful area, operationalizing it to represent a "neighborhood" (Farrell, 2008; Fischer, Stockmayer, Stiles, & Hout, 2004; Kramer & Hogue, 2009; Kwan, 2012; Massey & Hajnal, 1995; Massey, Rothwell, & Domina, 2009). While this is a reasonable argument, and certainly a better approach than pretending the MAUP does not exist, many researchers would argue that it is inappropriate to assume that administrative boundaries and "neighborhoods" are synonymous (Guo & Bhat, 2004; Manley et al., 2006; Martin, 1998; Oka & Wong, 2015). While enumerators may have tried to consider socio-economic distributions and natural barriers when delineating census tracts, census tracts were primarily created for statistical and financial purposes (Dumedah et al., 2008; Flowerdew, 2011; Martin, 1998; Openshaw, 1989; Steel, Holt, & Tranmer, 1994; Truesdell, 1942). Furthermore, administrative boundaries rarely represent residents' perceived neighborhoods (Coulton, 2012; Coulton, Korbin, Chan, & Su, 2001; Haynes, Daras, Reading, & Jones, 2007; Kwan, 2012). Investigating residents' perceptions of neighborhood boundaries, Coulton et al. (2001) found that while census tracts may be of approximately the same size as an individual's perceived neighborhood, they seldom share the same boundaries. Therefore, even if census units are the appropriate size, they are rarely the appropriate shape, making them susceptible to the zoning effect.

Another approach to solving the MAUP is to use better measurement techniques. Tobler (1990) argues that blame lies not with the choice of areal units but with the choice of analytical method. When researchers use appropriate measurements, errors associated with the MAUP can be mitigated (Briant et al., 2010; Fogarty, 2010). While not explicitly stated, it seems that some segregation scholars believe that spatial indices can avoid the MAUP, arguing that incorporating how census units relate to one another in space removes the absolute boundaries between arbitrary units (Cohn & Jackman, 2011; Jakubs, 1981; Morgan, 1982; Oka & Wong, 2015; Reardon et al., 2008). However, Wong (2004), who invented many of the spatial indices used, finds that spatial indices are also susceptible to the MAUP and therefore may not solve the problem either.

One certain way to avoid the MAUP in segregation analysis is to use individual-level data. Tobler (1990) argues that spatial analysis needs to be frame-independent, or reported in terms of its most elementary units, to prevent the MAUP. Many researchers claim that the only way for spatial analysis to truly be frame independent is for it to use individual-level data (Charlton, 2008; Fotheringham, 1989; Fotheringham & Wong, 1991; Horner & Murray, 2002; Páez & Scott, 2004; Reardon & O'Sullivan, 2004; Zhang & Kukadia, 2005). Given that individual-level data are the most basic building blocks, researchers can aggregate the information into non-arbitrary, socially contextual zones (Cockings & Martin, 2005; Martin, 1998; Openshaw, 1996). Unfortunately, while ideal, demographic data at this level of granularity is rarely available due to privacy and confidentiality concerns (Blalock, 1971; L. A. Brown & Chung, 2006; Butkiewicz et al., 2010; Cockings & Martin, 2005; Coulton, 2012; Fotheringham, 1989; Martin, 1998; Openshaw, 1996; Openshaw & Rao, 1995; Páez & Scott, 2004; Steel et al., 1994; Steel, Tranmer, & Holt, 2003). Therefore, conducting all segregation analysis using individual-level data is more of a dream than a potential reality.

Some researchers argue that solving the MAUP is an unrealistic and unobtainable goal (Clark & Avery, 1976; Dark & Bram, 2007; Fotheringham, 1989; Openshaw, 1989, 1996; Zhang & Kukadia, 2005). Given that the MAUP is almost inevitable, instead of developing universal solutions to the problem, researchers should work to perform robustness tests to better detail the problem for their specific analyses (Flowerdew, 2011; Fotheringham, 1989; Fotheringham & Wong, 1991; Kramer et al., 2010; Páez & Scott, 2004; Zhang & Kukadia, 2005). By detailing the MAUP, researchers can better understand how much confidence they should have in their conclusions; if results are consistent at multiple scales, there may be more confidence in the findings (Páez & Scott, 2004). For this reason, more work must be done to characterize the impact of the MAUP.

One study by Wong, Lasus, & Falk (1999) directly investigated how the MAUP impacted the measure of segregation by investigating how location, scale, and orientation affected the index of dissimilarity (ID) values for thirty cities. They found that the ID values varied with size and did so differently by city (Wong et al., 1999). While this work has helped the study of MAUP and segregation, it has some limitations that are worth noting.

First, Wong, Lasus, & Falk (1999) did not have access to individual-level data and had to resort to using the centroids of census enumeration districts (EDs) as their elementary units of analysis. The use of EDs is somewhat problematic. Openshaw (1984) argues that area-level analysis depends heavily on the first level of aggregation used. Therefore, the analysis done by Wong, Lasus, & Falk (1999) is heavily dependent on census blocks representing individual-level data. However, this assumption has been disproved by many scholars (Cockings & Martin, 2005; Cohn & Jackman, 2011; Manley et al., 2006; Martin, 1998; Openshaw, 1996; Tranmer & Steel, 1998). The individuals living in census blocks are not randomly distributed, and while they may be homogeneous for one characteristic, they are not for all relevant socio-economic characteristics (Manley et al., 2006; Tranmer & Steel, 1998). Additionally, Openshaw (1996) argues that while census blocks may be a good representation of the local environment in larger towns or cities, they may be a poor representation in small towns. However, blocks may mask important spatial relationships in the major cities too. Duncan & Duncan (1955) point out that if the non-White population lived in alleyways and the White population lived on main streets, then measuring segregation at the block-level would fail; this pattern has been empirically shown to exist (J. R. Logan & Bellman, 2016). Therefore, using arbitrary units, like census blocks, to test the MAUP is not the best approach.

Second, Wong, Lasus, & Falk (1999) focus on the index of dissimilarity. While the index of dissimilarity is arguably the most commonly used measure of segregation, it represents one dimension of segregation— evenness (Massey & Denton, 1988; Massey, White, & Phua, 1996; Taeuber & Taeuber, 1976). However, evenness is only one of five segregation, and the MAUP may impact the calculation of exposure, concentration, clustering, and centralization differently than evenness (Massey & Denton, 1988; Massey et al., 1996). Therefore, there needs to be more work to investigate how the MAUP impacts the other four dimensions as well.

Third, this article investigates only aspatial segregation indices. As many segregation scholars have shown, aspatial indices have major deficiencies that spatial indices overcome (Hong, O'Sullivan, & Sadahiro, 2014; Jakubs, 1981; Morgan, 1982, 1983b; Morrill, 1991; Oka & Wong, 2015; Östh et al., 2015; Reardon et al., 2008; Wong, 2003a, 2003b, 2004, 1993, 2002b). While Wong (2004) finds some evidence that spatial indices are just as impacted by the MAUP as aspatial indices, spatial indices deserve the same systematic investigation as the index of dissimilarity.

Lastly, Wong, Lasus, & Falk (1999) used square grids to partition the study areas into different scales. While this form of analysis is helpful in investigating scaling effects, it is less useful for investigating zoning effects, which would require looking at various shaped units.

This project performs an analysis similar to that of Wong, Lasus, & Falk (1999), but addressing many of the limitations mentioned above. Instead of using block-level data, it uses individual-level data to analyze how both size and shape impact a large set of segregation indices, both aspatial and spatial. By better understanding how the spatial structure of arbitrary units influences segregation measures, researchers can gain a better understanding of how their research may be affected by the MAUP.

Methodology:

Data Source:

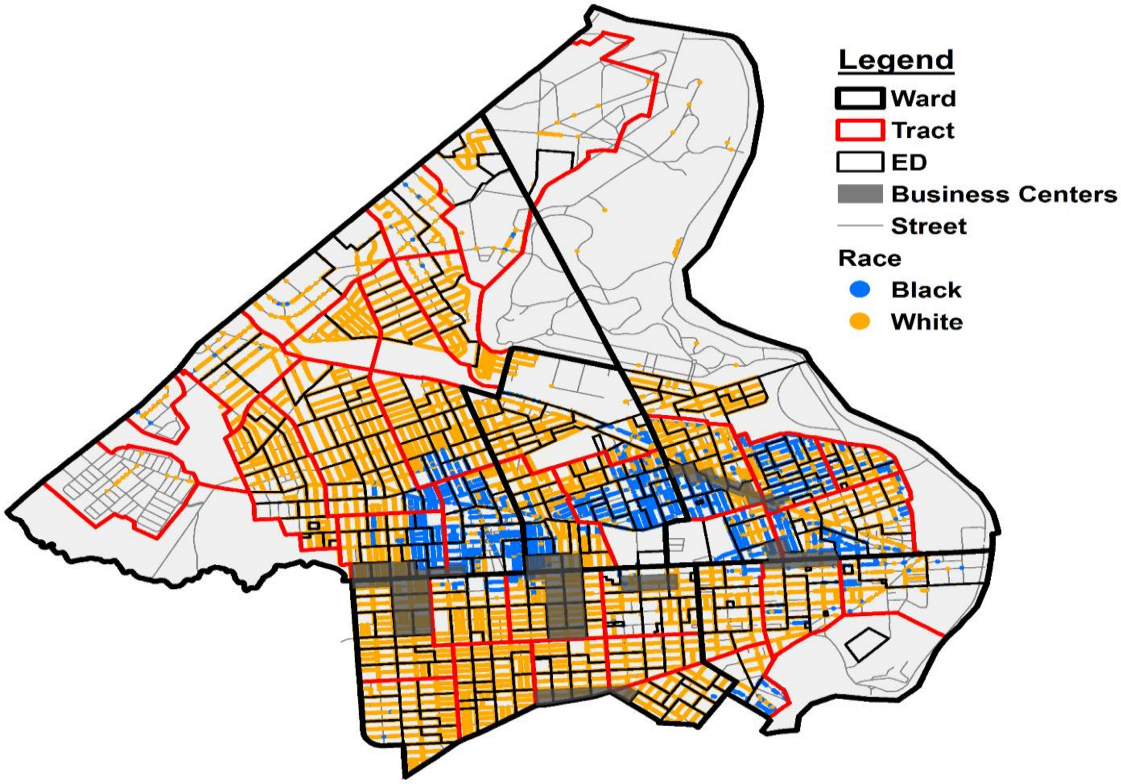
This project attempts to improve on the work done by Wong, Lasus, & Falk (1999) by shifting from EDs to individuals as the elementary units of observation. The population information for this analysis comes from the 100% count of the 1940 Census, provided by the Minnesota Population Center (MPC) (Ruggles, Genadek, Goeken, Grover, & Sobek, 2015). Before 2008, gaining access to individual-level data was prohibitively difficult. The 1940 Census is the most recent census that publicly provides individual-level data due to the "72-Year Rule," which states that all information from records, including house addresses, becomes accessible to the public only after 72 years (95th Congress, 1978). By utilizing the house number and street name information, this project could geocode all individuals to their exact house address instead of their census unit. Therefore, this project circumvents one of the largest limitations of the work done by Wong, Lasus, & Falk (1999).

Study Area:

This study includes all non-institutional residents living in West Philadelphia in 1940. West Philadelphia, is contextually bounded, being bordered by the Schuylkill River on the east and north, Cobbs Creek and the county line on the west, and Baltimore Avenue to the south where the Baltimore Railroad tracks historically separated the residential area of the north from the industrial area of the south (Weaver, 1930). With close to half of the boundaries defined by natural resources and the other half defined by railroads and county limits, West Philadelphia is relatively self-contained. West Philadelphia encompasses five wards, 41 tracts, and 344 EDs. In 1940, there were 54,711 Black individuals, 242,347 White individuals, and 133 Asian/Pacific Islander and Native American individuals residing in West Philadelphia. Given the small size of the Asian/Pacific Islander and Native American population, this study only includes the Black and White populations, making for a total of 297,058 individuals.

Figure 1.1 shows the spatial distribution of individuals by race, as well as the different Census administrative boundaries for West Philadelphia for 1940.

Figure 1.1: West Philadelphia Population with Administrative Boundaries, 1940



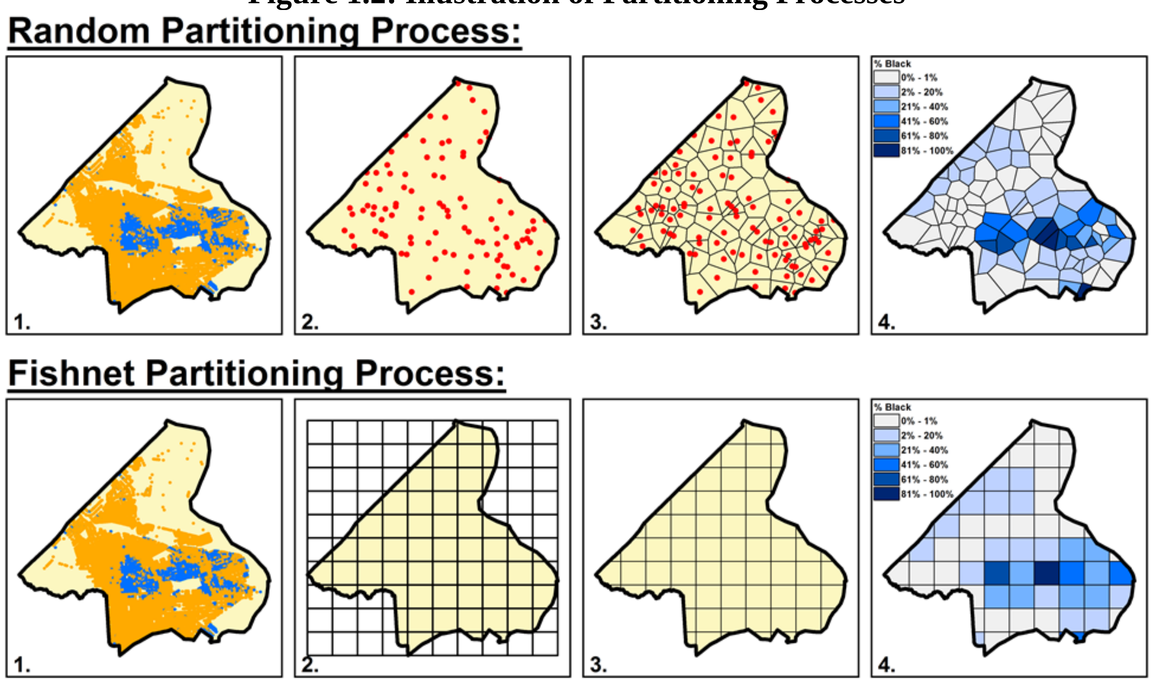
Note here that most Black individuals are concentrated in a few census tracts, while White individuals are spread throughout the area. To effectively capture this pattern using areal units is challenging. The Black neighborhoods do not fit well into any one set of administrative boundaries but instead occupy different census units to varying degrees. Additionally, it is evident that there are other ways to partition census tracts and EDs that would more accurately reflect this spatial pattern. Except for respecting the North/South Market Street divide, the boundaries employed are largely arbitrary (Boyer, 1896; Roberts, 1980; Warner Jr., 1987). Many of the tracts and EDs are gridded in an L-shape when they could have just as easily been gridded in a rectangular shape.

Had the boundaries been partitioned differently, segregation values would undoubtedly change, reflecting the zoning effect. There is evidence of the scaling effect in this study, as well. Depending on the areal units chosen, the index of dissimilarity (ID) is either 0.3424 using wards, 0.6128 using tracts, or 0.7857 using EDs. However, given the arbitrary nature of the areal units, the actual level of dissimilarity may be underestimated.

Analysis:

This project uses two different approaches to partition the study area. Each method provides different insights into how shape and size impact segregation calculation. Figure 1.2 provides a visual aid to understand the processes of the two approaches.

Figure 1.2: Illustration of Partitioning Processes



Random Partitioning Process: (1) start with study area (2) create N random points (3) create Thiessen polygons for each point (4) join population data to new units (5) repeat process 20 times for N between 2 and 400.
Fishnet Partitioning Process: (1) start with study area (2) create fishnet grid with specified dimensions (3) clip study area by fishnet grid (4) join population data to new units (5) increase fishnet height or width by 100 meters and repeat process until performed all combinations.

The first part of this analysis randomly partitions the study area into a specified number of units given as N and ranging between 2 and 400, then repeats this process 20 times to create 20 different zoning patterns for each N, leading to 7980 different configurations of West Philadelphia in total. The Census divides West Philadelphia into 5 wards, 41 census tracts, and 342 enumeration districts; therefore, partitioning the study area into units ranging between 2 and 400 is an appropriate range. This method allows for the simultaneous inspection of both the scaling and zoning effects. Differences in segregation values over the variation in N reflects the scaling effect, while differences in segregation values among the 20 simulations at any one N reflects the zoning effect.

However, randomly partitioning areas makes it difficult to inspect how shape impacts calculation. It is for this reason that the second part of this analysis uses a fishnet grid to partition the study area. Similar to the methodology employed by Wong, Lasus, & Falk (1999), this analysis starts with a 100-meter by 100-meter square grid and increases the size of the grid until the entire study area is encompassed inside only one square. However, unlike Wong, Lasus, & Falk (1999) who increased both the height and the width of the grid by the same amount to maintain a square shape, this analysis increases the height and width separately by 100-meters to produce both square and rectangular grids, both common shapes for census units. While this approach is somewhat basic, only creating two types of polygons, it allows for a more systematic investigation of how changing height and width differentially impact segregation calculations.

This project investigates the MAUP's impact on segregation indices from all five dimensions of segregation: evenness, exposure, concentration, clustering, and centralization (Massey & Denton, 1988). While West Philadelphia does not have a central business district, it contains six local business centers[1] which will act as the business district for the purpose of this analysis when calculating centralization. This project uses the Geo-Segregation Analyzer developed by Apparicio et al. (2014) to calculate the segregation indices. While the Geo-Segregation Analyzer can calculate as many as 43 different

---

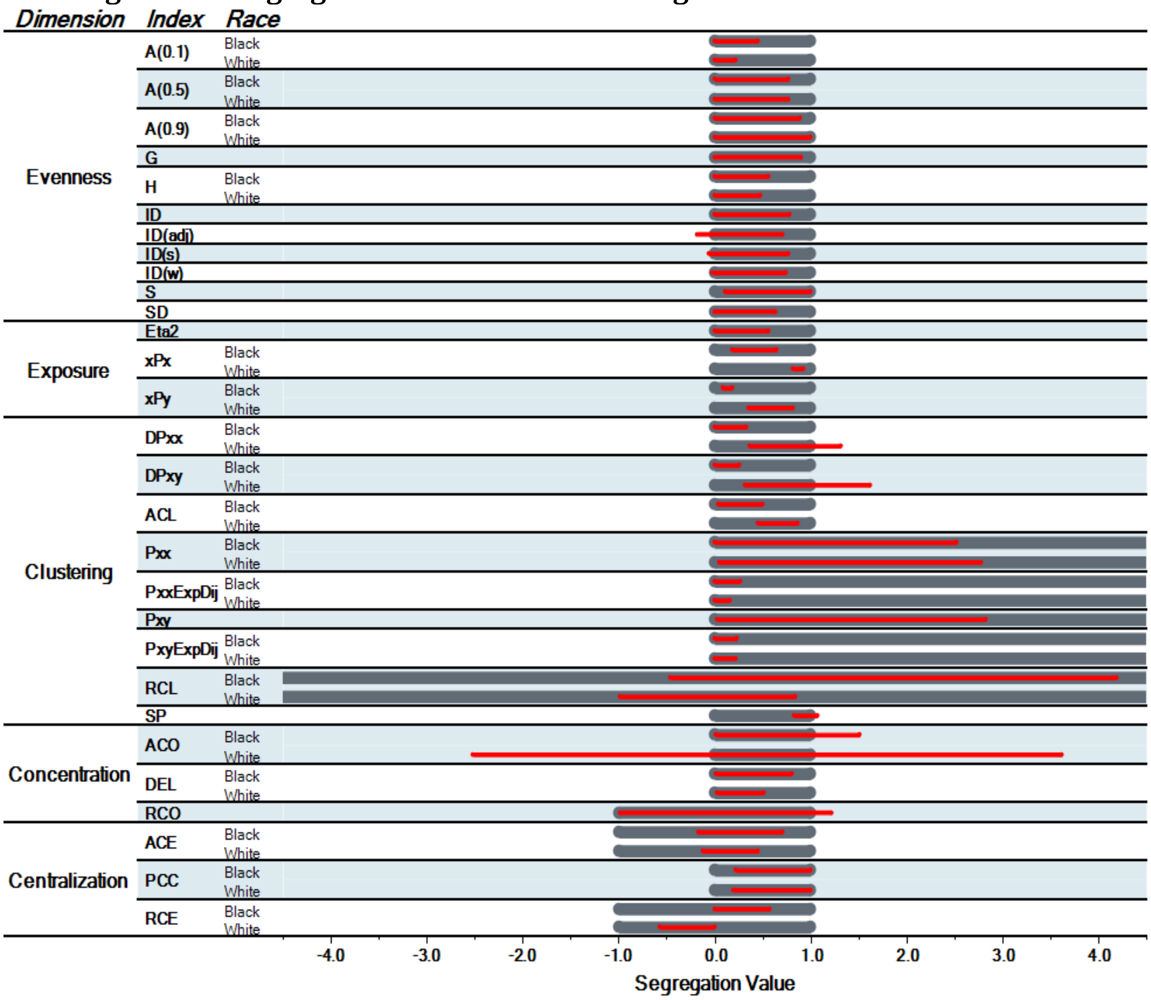
[1] Business center information comes from the 1935 Intra-City Business Census Statistics for Philadelphia, Pennsylvania (Proudfoot, 1937).

residential segregation indices, many of those indices duplicate one another when applied to only two racial categories. Additionally, while some spatial indices have two different options for the contiguity matrix, the method for identifying which adjacent units to include as neighbors for the calculation, it does not make a difference for this analysis (Apparicio et al., 2014; Morrill, 1991; Wong, 1993). Like moves on a chess board, areal units can have rook neighbors who share a line segment or queen neighbors who can share either a line or point segment (Wong, 2002a). However, as chance would have it, none of the simulations produced in this analysis through the random partitioning process had any units with queen neighbors, meaning that the two calculations are the same. Therefore, this project analyzes only 29 indices.

Results:

Figure 1.3 shows the range of values calculated by each segregation index from the units created using the random simulation process. Each segregation index has a different theoretical maximum and minimum, as shown by the wider gray band (Apparicio et al., 2014). For comparison, the segregation indices are grouped by their segregation dimension. If a segregation index has a large range of values, it means that the index produces very different results as the size and shape of the areal units change. Conversely, if a segregation index has a short range of values, it means that the index produces a consistent measure of segregation regardless of the shape or size of the areal unit. In general, researchers should want indices that produce the same value regardless of how the underlying pattern is aggregated.

Figure 1.3: Segregation Indices Values Using Random Simulation Process

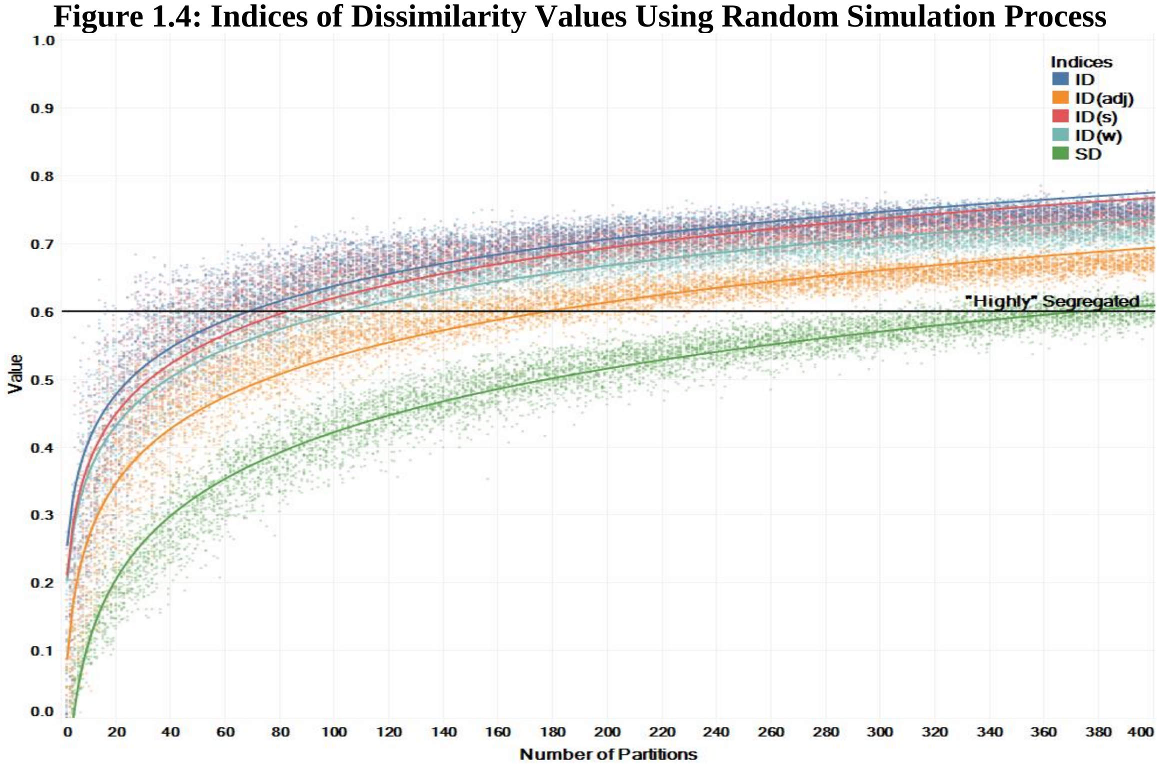


Note: A(0.1) = Atkinson Index with b=0.1; **A(0.5)** = Atkinson Index with b=0.5; **A(0.9)** = Atkinson Index with b=0.9; **G** = Gini Index; **H** = Entropy Index; **ID** = Index of Dissimilarity; **ID(adj)** = Index of Dissimilarity Adjusted for Tract Contiguity; **ID(w)** = Index of Dissimilarity Adjusted for Contiguous Tract Boundary Lengths; **ID(s)** = Index of Dissimilarity Adjusted for Contiguous Tract Boundary Lengths and Perimeter/Area Ratio; **S** = Deviation Ellipse Index; **SD** = Spatial Index of Dissimilarity; **Eta2** = Correlation Ratio; **$_xP_x$** = Isolation Index; **$_xP_y$** = Interaction Index; **DP_{xx}** = Distance-Decay Isolation Index; **DP_{xy}** = Distance-Decay Interaction Index; **ACL** = Absolute Clustering Index; **P_{xx}** = Mean Proximity between Members of Group X; **$P_{xx}ExpD_{ij}$** = Mean Proximity between Members of Group X with Distance Function=exp(-d_{ij}); **P_{xy}** = Mean Proximity between Members of Group X and Members of Group Y; **$P_{xy}ExpD_{ij}$** = Mean Proximity between Members of Group X and Members of Group Y with Distance Function=exp(-d_{ij}); **RCL** = Relative Clustering Index; **SP** = Spatial Proximity Index; **ACO** = Absolute Concentration Index; **DEL** = Delta Index; **RCO** = Relative Concentration Index; **ACE** = Absolute Centralization Index; **PCC** = Proportion in Central City; **RCE** = Relative Centralization Index

These results indicate that all segregation indices that rely on areal units are subject to the MAUP to some extent, regardless of segregation dimension. Despite taking the spatial relationship of areal units into consideration, either by contiguity matrices or distances functions, many spatial indices were just as impacted by the MAUP as their aspatial counterparts. However, some indices are more affected by the MAUP than others. The ID(adj), ID(s), RCO, ACO, DP_{xx}, DP_{xy}, and SP indices are potentially the most impacted by the MAUP given that simulations produced segregation values beyond the theoretical minimum and maximum value possible. While researchers have confirmed earlier that the RCO is in fact not bounded on the negative side, this analysis shows that the RCO is not bounded on the positive side either (Egan, Anderton, & Weber, 1998; Massey & Denton, 1998). This finding shows that the MAUP can be so impactful that it violates the basic criteria of these indices-- having a minimum and maximum (Jahn et al., 1947; D. R. James & Taeuber, 1985; Reardon & Firebaugh, 2002; Williams, 1948). The indices that had the shortest range of values, and therefore were the least impacted by scaling and zoning effects, were ${}_xP_x$, $P_{xx}ExpD_{ij}$, and $P_{xy}ExpD_{ij}$. These findings are interesting for two reasons. First, it is interesting to note that P_{xx} and P_{xy} have a large range of segregation values while $P_{xx}ExpD_{ij}$ and $P_{xy}ExpD_{ij}$ do not given that the only difference between these two sets of indices is which distance function they employ (Apparicio et al., 2014; Massey & Denton, 1988). Second, it is interesting to note that for the White population, the P* indices (${}_xP_x$ and ${}_xP_y$) have a smaller range of values than the DP* indices (DP_{xx} and DP_{xy}). While Massey & Denton (1988) initially classified the DP* indices as a measure of clustering, Massey, White, & Phua (1996) later argued that the DP* indices are nothing more than distance-weighted P* measures. These two findings are somewhat in contention with one another. On the one hand, the first finding would lead one to conclude that an exponential distance function is more effective than a linear distance function in avoiding the MAUP (White, 1983). On the other hand, the second finding would lead one to conclude that using an aspatial measure is more effective than an exponential distance function in avoiding the MAUP (Mitra, 1984). Together, the results of Figure 1.3 show that there is no absolute certainty when using areal units to calculate segregation, whether the indices are spatial or not.

Given its prominent use in the literature, many researchers have argued that modifying the index of dissimilarity (ID) should be a low priority (Duncan & Duncan, 1955; Taeuber & Taeuber, 1976). Despite this recommendation, many researchers have modified the ID (Cortese, Falk, & Cohen, 1976; Morgan, 1982, 1983b; Morrill, 1991; Wong, 1993, 2002b, 2003a; Wong & Chong, 1998). Figure 1.4 displays how the MAUP impacts the segregation values for the index of dissimilarity and its modified indices. A reference line at a segregation value of 0.6 is included to show the threshold where most scholars would consider values above that line as highly segregated (Massey & Denton, 1989, 1993; Sin, 2002).

Figure 1.4: Indices of Dissimilarity Values Using Random Simulation Process



Note: ID = Index of Dissimilarity; **ID(adj)** = Index of Dissimilarity Adjusted for Tract Contiguity; **ID(w)** = Index of Dissimilarity Adjusted for Contiguous Tract Boundary Lengths; **ID(s)** = Index of Dissimilarity Adjusted for Contiguous Tract Boundary Lengths and Perimeter/Area Ratio; **SD** = Spatial Index of Dissimilarity

All the indices of dissimilarity are affected by both the scaling effect and, to a lesser extent, the zoning effect of the MAUP. If the indices were not impacted by the MAUP, we would expect to see a straight line, meaning that the indices calculate the same value regardless of how the population is aggregated. However, the results show the opposite. As the number of areal units increases, and the average size of the areal units decrease, the segregation value increases. These changes are not trivial. These indices can characterize West Philadelphia as having either high levels of segregation or low levels of segregation depending on the scale of aggregation.

The West Philadelphia results also show evidence of the zoning effect. At any one scale, there is variability in segregation values amongst the different configurations. While the overall impact of the zoning effect is smaller than that of the scaling effect, it is not at all trivial at certain scales. At larger scales, meaning few partitions, there is greater variability among the different simulated configurations. However, as the number of units increases, the within-variability declines as well.

Figure 1.5 displays the results from part two of the analysis which used fishnet methodology to grid West Philadelphia. Each dot represents a different fishnet grid with its height and width indicated by the axes. The color of the dot represents its ID value.

Figure 1.5: Index of Dissimilarity Values Using Fishnet Gridding Method

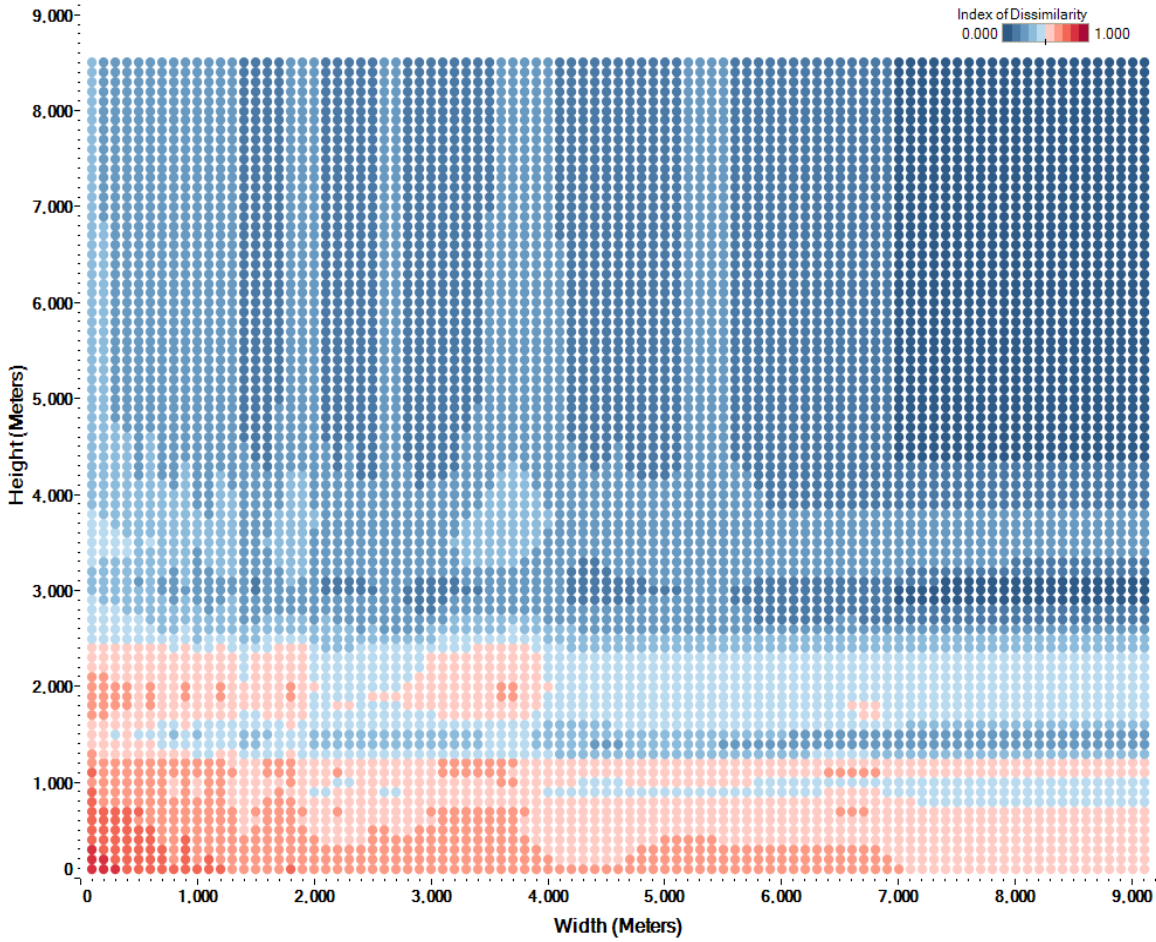


Figure 1.5 reconfirms the claim that the smaller the unit, the larger the ID value. However, Figure 1.5 also shows that increasing the unit area does not decrease the ID value uniformly. Increasing height often had a larger impact on the ID value than increasing width. Gridding the study area with heights above 2,700 meters results in low ID values, regardless of the width. Correspondingly, heights below 1,200 meters always result in high ID values. This stark difference in ID values between extremely tall and extremely wide areal units arises because of where the Black population lives in West Philadelphia. Most of the Black population live between Market Street and Girard Avenue and live all the way west from 30th Street to all the way east to 60th Street. Therefore, when there are wide units, it is likely that each unit may only include one race, resulting in a high ID value.

However, when the units are tall, it is more likely that each unit will include both races, resulting in a lower ID value. This finding shows that unit dimensions can present different conclusions about the level of segregation, even if units have the same total area.

One of the most interesting patterns that emerge from Figure 1.5 is how the ID changes when adjusting the width of the units with the height of the units between 1,700 meters and 2,600 meters. There are four components of this pattern: 1) a high ID value when width is below 2,000 meters; 2) a low ID value when width is between 2,000 meters and 2,500 meters; 3) a high ID value when width is between 2,500 meters and 4,000 meters; and a low ID value when width is greater than 4,000 meters. The first component has a high ID value because of scaling effect. When there are many small areal units, it is more likely that any one unit will be more racially homogenous, leading to high ID values. The second component has a low ID value because of a zoning effect. At these widths, the unit boundaries rest in the middle of the Black neighborhoods, splitting the Black population among many units which also include some White individuals. Therefore, these units appear heterogeneous, leading to a low ID value. The third component has a high ID value because of a zoning effect. At these dimensions, the Black neighborhoods fit entirely into one unit, making the area appear homogeneous, resulting in a high ID value. The fourth component has a low ID value because of scaling effect. Once the widths are a certain length, they are so wide that they must include both White and Black individuals, making the area appear integrated. This figure shows that there is an interplay between the scaling and zoning effects. Depending on the unit dimensions, which effect explains more of the variation in segregation values changes. Figure 1.5 shows that while the scaling effect is a major factor at the extremes, the zoning effect may matter more when the height and width are closer to even. However, without the ability to perform robustness checks, like the ones done in this analysis, it would be difficult to know which effect is acting at each moment.

Discussion:

The results from this article expand on the work of Wong, Lasus, & Falk (1999) in several important ways. First, this article's findings show that the MAUP not only impacts segregation values calculated with the index of dissimilarity but is a problem to a certain degree for all 29 segregation indices calculated in this study, regardless of dimension. Second, the findings providing support for Wong's (2004) claim that, like their aspatial counterparts, spatial indices are also impacted by the MAUP. Third, this article more systematically investigates how both the shape and size of areal units impact measurement, finding evidence of both the scaling and zoning effect for the study area. The scaling effect of the MAUP is potentially very dangerous when measuring segregation because of its ability to classify the same underlying pattern as both highly segregated and highly integrated depending on the size of the unit. Additionally, while the zoning effect may be less significant than the scaling effect in general, it can be quite impactful at higher levels of aggregation.

Given this article's findings, we as researchers can no longer avoid pretending that the MAUP does not exist or that it does not impact our results. Instead, we need to acknowledge its existence, and there are several ways to do so.

We need to move away from arbitrary definitions. The MAUP only exists because the areal units used are arbitrarily bound, lacking any socio-economic meaning (Openshaw, 1989). By incorporating historical and ethnographic information, we can start to identify more meaningful boundaries and thereby mitigate measurement errors arising from the MAUP. For example, if someone wants to measure centralization, then identifying the exact delineations of the different concentric zones would allow for more accurate measurement (Burgess, 1928). On a similar note, we need to move away from arbitrary threshold values of segregation. If we acknowledge that the segregation values change depending on the scale, then blindly following arbitrary cutoff points, like 0.6 for the index of dissimilarity, can lead to inaccurate conclusions (Sin, 2002). Instead, we need to shift to scale-dependent definitions of what it means to be highly segregated. While scale-dependent definitions do not fully eliminate the problem, it shifts the problem from a scaling effect to a zoning effect, which the findings of this article find to have less variability.

We need to perform more robustness checks in our analysis. The MAUP is primarily a form of measurement error, no different from any other measurement error encountered in quantitative research. Sometimes we can address such errors by improving the data; other times we can use more precise measurement. However, when those options are not available, we need to apply error bounds to our estimates and provide more conservative interpretations of those results. One approach is to develop standard error measures that are associated with the segregation indices to create confidence intervals. An alternative approach is to repeat the analysis at multiple scales. If the results are similar at multiple scales, it does not necessarily mean the finding is correct but does indicate that something is occurring that is worth investigating further. Alternatively, if the results are different, and there is no theory to explain why they might differ by scale, then there may be an error from the scaling effect. These types of robustness checks will allow for more conservative conclusions even in the presence of the MAUP.

Lastly, we need to make a push for more publicly available individual point-level data. This geographical level of specificity addresses many of the concerns mentioned above. Disaggregated, individual-level point data avoids the MAUP because they represent the individual or the household, an area of contextual value (Ellis, Holloway, Wright, & Fowler, 2012). Individual-level data can also be aggregated into contextually-defined areal units as well. Additionally, individual-level point data can be aggregated in many different arrangements, allowing for robustness checks by creating empirical bounds of reasonable estimates. Given the potential improvements brought on by point-level data, we should work harder to make this type of data publicly available while still protecting privacy and confidentiality.[2]

Limitations:

Despite this project's improvements on the work done by Wong, Lasus, & Falk (1999), it is not without its limitations. The main limitation of this study is that it only investigates one space and time—West Philadelphia in 1940. Wong, Lasus, & Falk (1999) found that the ID not only varied by scale in each city but that it did so differently in different cities. This analysis should, therefore, be repeated using other cities and time

---

[2] Interesting side note: the choice of 72 years to wait for Census data to become public is to some extent an arbitrary one. While many would believe that the 72-Year Rule was established to reflect the life expectancy of the U.S. population at the time, this is incorrect (Weintraub, 2008). In 1921, a fire damaged the U.S. Commerce building and destroyed the 1890 Census information. This event prompted the formation of the National Archives to house federal data (Weintraub, 2008). When the Census data were transferred to National Archives in 1942, the information became public to researchers (Weintraub, 2008). At that time, the latest Census record transferred was the 1870 Census, a difference of 72 years (Weintraub, 2008). In fact, in 1950, Congress passed a law that required a 50 year limit, but allowed the Archivist to set a longer waiting period (Weintraub, 2008). It was only later during the 1970 Census that concern about privacy and confidentiality arose, leading to the 1978 bill that included the 72-Year Rule (95th Congress, 1978; Weintraub, 2008). Therefore, the origin of the current provision is largely a consequence of historical events, not a desire to protect confidentiality and privacy.

periods to investigate how these factors may interact with administrative boundary formation and the MAUP. Another limitation of this study is that it only uses two groups, not allowing an investigation of how the MAUP impacts multigroup segregation indices. As indicated by the work of Iceland (2004), there is a need for segregation analysis to push away from the Black-White divide and focus more on the multiethnic context. A third limitation is that this study only systematically investigated how changing the dimensions of square and rectangular units impact measurement. While many of the areal units in West Philadelphia were squares or rectangles, these two shapes represent a small fraction of all possible geographic shapes; each shape has a potentially different impact on measurement (Boyce & Clark, 1964; Taylor, 1971). Therefore, future research should also investigate other unit shapes, particularly the L-shaped unit which commonly appear at the ED-level.

Conclusion:

Acknowledging the existence of the modifiable areal unit problem is the first step toward improving the measurement of segregation. Through improving the accuracy of our indices by reducing measurement error, our indices start to get closer to our perceptions of how segregation operates. However, as long as segregation analysis continues to rely on areal units for calculation, there will always remain the risk of introducing the MAUP. For this reason, we need to develop a better language when talking about the MAUP in our research. When we discuss how segregation values change at different scales, we are in fact talking about two separate phenomena: 1) the change in value due to the picking up signals from socio-economic structures that operate at different scales and 2) the change in value due to the measurement error associated with the MAUP. However, researchers often lack the vocabulary to distinguish these two phenomena. Given that researchers will undoubtedly continue to use areal units in their research, having the language to discern socially relevant changes and measurement error will be vital for continuing rigorous analysis.

In addition to developing methods to scrutinize area-level data better, researchers should continue the effort to ascertain more individual-level data while still maintaining privacy and confidentiality concerns. Being the most basic building block in an area, using individual-level data to aggregate information into contextually relevant boundaries is one of the best ways to eliminate the MAUP (Reardon & O'Sullivan, 2004). However, individual-level data also presents a new set of possibilities. While most ecological models on segregation focus on the individual, area-level indices provide a bird's eye perspective of the entire population (Alba & Logan, 1992; Tomlin, 2017). However, with individual-level data, we can shift the point of view to a worm's eye perspective and investigate what it means for any one individual to experience segregation (Tomlin, 2017). As this happens, there will be a need to develop new indices that can utilize individual-level data to measure the individual-level experience of segregation. Through these processes, we will have more accurate tools to study the causes or consequences of segregation at all relevant scales (Ellis et al., 2012).

CHAPTER 2: THE SHORTEST PATH ISOLATION INDEX: A NEW MEASURE FOR INDIVIDUAL-LEVEL RESIDENTIAL SEGREGATION

Introduction:

Grannis (1998) notes that "while it is no revelation that segregation begins with one's neighbor, what is often overlooked is how these neighborly interactions, like links in a chain, form the backbone of racial housing segregation." (p. 1560). While social science has measured American residential segregation for almost a century, little research focuses on this most basic unit— the individual. Despite researcher's desire to so, many have resolved to study segregation at a larger level of aggregation due to the lack of available data. With the recent releases of public complete count Census data, researchers now can investigate residential segregation at the individual-level. However, few individual-level indices exist. In this absence of capable indices, this article proposes the Shortest Path Isolation (SPI) index and highlights its potential for studying this scale of segregation.

Literature Review:

Residential segregation is a complex social phenomenon that refers to different spatial patterns and operates at different geographic scales (Kaplan & Woodhouse, 2005; Massey & Denton, 1988; Reardon et al., 2008; Reardon & O'Sullivan, 2004). Given its complexity, researchers argue that there is no one index that can accurately summarize all aspects of segregation (Duncan & Duncan, 1955; R. Harrison, 2001; Jahn et al., 1947; Massey et al., 1996; Taeuber & Taeuber, 1965, 1976). Therefore, instead of developing such a summary index, researchers should use different indices to study different dimensions and scales (Massey et al., 1996).

There has been a constant push in the residential segregation literature to refine segregation indices to better detail the complexities of residential segregation (Duncan & Duncan, 1955; Massey & Denton, 1988). In the beginning, segregation indices were largely aspatial, not accounting for the locational relationship among residential units (Massey & Denton, 1988; Reardon & O'Sullivan, 2004). Many of these measures were later modified to include spatial components (Jakubs, 1981; Morgan, 1982, 1983a; Morrill, 1991). As mapping software became more common, more spatial indices arose (Kaplan & Woodhouse, 2005; Reardon & O'Sullivan, 2004; Wong, 2004; Wong & Shaw, 2011). As the Asian and Latino populations grew in the United States, researchers again modified the indices to account for multiple groups (Iceland, 2004; F. J. James, 1986; K. Jones, Johnston, Manley, Owen, & Charlton, 2015; Maly, 2000; Reardon & Firebaugh, 2002). Now, as the availability of more detailed geographic information is growing, researchers have a new task— measuring segregation at different scales, particularly the scale of the individual (Grigoryeva & Ruef, 2015; B. A. Lee et al., 2008; J. R. Logan & Bellman, 2016; T. D. Logan & Parman, 2017; Massey & Hajnal, 1995; Massey et al., 2009; Reardon et al., 2008).

Studying segregation at different scales is important to investigate the different possible mechanisms which may contribute to segregation. For example, migration patterns and federal policies may contribute to macro-level segregation patterns at the state or metropolitan area while discriminatory housing and lending policies may contribute to micro-level segregation patterns at the neighborhood or street segment (J. R. Logan & Bellman, 2016; Massey et al., 2009). However, when studied at the incorrect scale, segregation becomes invisible and therefore impossible to understand the mechanisms at play (J. R. Logan & Bellman, 2016). Therefore, specifying the appropriate geographic scale to match with specific theory is important for studying how segregation operates.

Along with studying segregation at larger scales, researchers are interested in how the individual, as well as the household, fit into the continuum of scales at which segregation operates (Ellis et al., 2012; Reardon et al., 2008). In addition to institutional practices that operate at larger scales, individual attitudes and behaviors also play a major role in segregation (Sundstrom, 2004). Consequently, many of the underlying theories specify how segregation works at the level of individuals (Alba & Logan, 1992). Hipp (2012) argues that much of the segregation processes operate at the micro-neighborhood level, particularly at the level of the household, and that many of these processes may not be detected using aggregate-level data. By investigating the individual, the researcher shifts the perspective from place to people, investigating how individual behaviors and actions influenced and are influenced by broader segregation patterns (Kwan, 2009). Despite this desire for to study the individual, few datasets have the appropriate information for such analysis, forcing researchers to make conclusions at the aggregate-level and extrapolate about individual-level (Alba & Logan, 1992, 1993; Lieberson, 1963; Robinson, 1950; White, 1983). However, this reliance on aggregate-level data has introduced both conceptual and technical problems.

One of the biggest problems using aggregate place-based data is that the results are sensitive to the choice of the areal unit (Jahn et al., 1947; Reardon et al., 2008). Using different areal units, like census blocks or tracts, to measure segregation will produce different results despite the fact that nothing about the underlying residential pattern has changed (Jakubs, 1981; Massey & Denton, 1993; Taeuber & Taeuber, 1965). The literature refers to this phenomenon as the modifiable areal unit problem (MAUP) (Dark & Bram, 2007; Flowerdew, Manley, & Sabel, 2008; Gehlke & Biehl, 1934; Wong, 2004). This MAUP can manifest itself in a few ways. First, using units of different area sizes, or scales, can produce different results (Wong, 2004). This enigma makes sense when considering the assumptions necessary for area-level calculations. Due to data constraints, area-level indices often hold the assumption that there is an equal chance of meeting any individual within an areal unit (Bell, 1954). As the size of the unit used increases, say from using a block to a tract, each unit will incorporate a larger population. Given that population distributions are rarely uniform, it is likely that the racial composition of each unit changed, which will, in turn, affects the overall calculation (Dark & Bram, 2007). However, changing scale is not the only problem. Modifying the shape of the unit, while maintaining the same total area, may also produce inconsistent results (Wong, 2004). A real-life example of this effect is gerrymandering. Through gerrymandering, the population and racial composition of an area may look very different depending on how the boundaries are drawn (Duncan & Duncan, 1955). For these reasons, the MAUP is a real problem that brings into question the ability of area-level measures to produces consistent results.

Another issue with using area-level place-based data is that the areal units often do not represent an area of contextual relevance. Given that most traditional segregation indices treat areal units as separate and independent units, individuals who live close to one another but are across unit boundaries are dealt with as more distant than those who live further away but are within the same areal unit (Reardon & O'Sullivan, 2004). Even spatial segregation indices do not alleviate this problem because they weight individuals living in other tracts as less than those living in the same tract (Morrill, 1991). In addition to the potential for measurement error, not using contextually relevant boundaries makes it difficult to investigate potential causal pathways (Kwan, 2012). By not using areal units that represent neighborhoods, it is almost impossible to test the role of social and economic processes in forming segregation patterns (Charles, 2003; Kramer et al., 2010; Kwan, 2012). Therefore, even if segregation measures are reliable, they do not always provide the most pertinent information.

Unfortunately, there has been no easy solution to alleviate the MAUP while still relying on aggregate place-based data. Reducing the size of the areal units does not resolve the problem. As Duncan & Duncan (1955) point out, even if the analysis switched from census tracts to blocks, if the non-White population lived in alleyways and the White population lived on main streets, then measuring segregation at the block-level would still fail. Additionally, increasing the number of sample units in an attempt to have a more normal distribution of units also would not solve the problem (Carrington & Troske, 1997). Furthermore, spatial segregation indices do not solve the MAUP and in certain situations may result in an even larger degree of inconsistency than the aspatial measures (Wong, 2004). The fact remains, using areal units will almost always present the risk of introducing the MAUP.

Using individual-level geographic data is the only proposed solution to address areal unit problems. Reardon & O'Sullivan (2004) argue that exact locations of individuals eliminate the MAUP entirely. Additionally, reducing the geographic unit of analysis to the individual, or household, allows researchers to create contextually relevant ego-centric neighborhoods centered around individuals (Chaix, Merlo, Evans, Leal, & Havard, 2009; Spielman & Logan, 2013). By reducing the scale to the individual, researchers can shift from a study of place to a people-based representation, focusing more on the experiences of the individual (Kwan, 2009). Therefore, creating segregation indices that use individual-level data has the potential to avoid many of the problems that have plagued prior segregation indices.

Thus far, there have been two proposed indices that empirically use individual-level data: the Sequence Index of Segregation (SIS) and the neighbor-based measure (Grigoryeva & Ruef, 2015; T. D. Logan & Parman, 2017). Using the household order in the Census enumeration forms from the 100% 1880 Census data, Grigoryeva & Ruef (2015)'s Sequence Index of Segregation (SIS) calculates the number of sequences of racially alike neighbors and compares the observed number to expected number based on the racial composition and population size of the area. Similarly, Logan & Parman (2017) neighbor-based measure uses the household ordering of the 100% 1940 Census data to compare the number of individuals who have a neighbor of a different race to the expected numbers under complete segregation and integration. These indices have significantly advanced the study of segregation patterns. In addition to the new substantial knowledge provided, these indices avoid many of the technical and conceptual limitations that befell prior indices, particularly the MAUP (Grigoryeva & Ruef, 2015; T. D. Logan & Parman, 2017). However, despite these major improvements, these indices are not without limitations.

The major limitation of both indices is that while the calculations are at the individual-level, the produced results are at the area-level for interpretation (Grigoryeva & Ruef, 2015; T. D. Logan & Parman, 2017). This process shifts the indices from a people-based representation back to a place-based representation (Kwan, 2009). While this may not necessarily introduce the MAUP, it likely introduces a contextual problem given that the aggregation is often at the enumeration or tract scale (Kwan, 2012). Therefore, it becomes more difficult to investigate individuals' exposure given that the information is aggregated at a scale that is not contextually relevant (Kramer et al., 2010; Kwan, 2012). By reducing the data back to areal units, these indices risk falling into the same traps when interpreting findings as that prior area-level indices.

Another limitation of both indices is their reliance on Agresti's (1980) assumption to reduce the cost associated with geocoding individuals. Agresti (1980) argues that since Census takers walked along a typical pedestrian path, recording household information along the way, that it was reasonable to assume that "most households adjacent in the census listings were also adjacent in their neighborhoods" (Agresti, 1980, p. 390; Austin & Hopkins, 1940). Grigoryeva & Ruef (2015) use Agresti's (1980) assumption to plot households to an approximate location based on the ordering in the census list, while Logan & Parman (2017) avoid geocoding altogether and simply construct their binary outcome based on the record above and below the individual on the enumeration form. Unfortunately, by relying on this assumption, both indices limit their full potential. This approach can only safely identify one's next-door neighbors, exclusively those living on the same side of the street. However, this definition of neighbor is not the only pattern of segregation worth measuring (Grannis, 1998, 2005; Grigoryeva & Ruef, 2015). In addition to studying one's next-door neighbor, the studying one's next-next-door neighbors, one's across-the-street neighbors, and one's "backyard" neighbors are also important. Unfortunately, without knowing exact house addresses, identifying any other form of neighbor based on one's relative position along an enumeration form is practically impossible. Therefore, by relying on the sequence in the enumeration form to determine proximity instead of truly geocoding individuals to their residences limits the extent to which both indices can measure isolation in an area.

Despite the major advances brought by these indices, there is still room for improvement. This article presents a new individual-level people-based measure, the Shortest Path Isolation index, which utilizes individuals geocoded to their exact residence to avoid the limitations that have befallen both area-level and other individual-level indices.

Methodology:

The Measure:

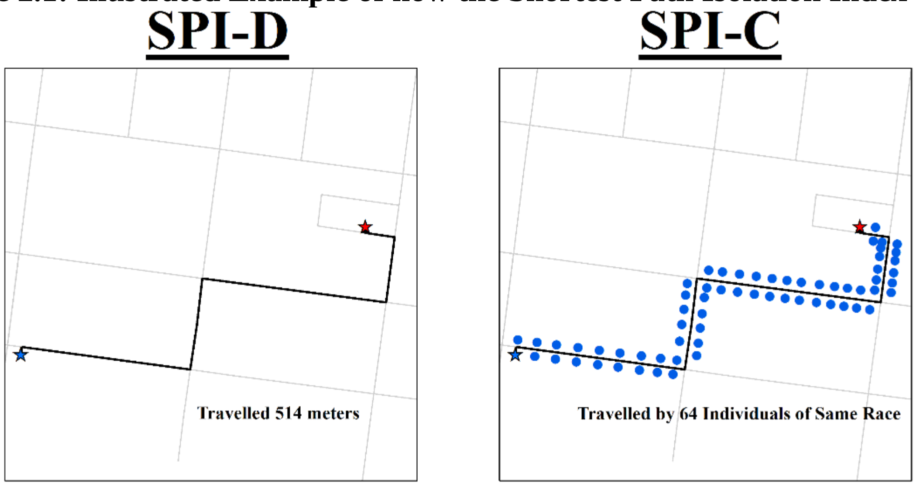
The term segregation has many meanings and patterns (Kaplan & Woodhouse, 2005). Therefore, before detailing how the Shortest Path Isolation Index operates, it is necessary to clarify which form of segregation the measure addresses. In their survey of the literature, Massey & Denton (1988) categorize segregation into five distinct dimensions: evenness, exposure, concentration, centralization, and clustering. However, Reardon & O'Sullivan (2004) argue that the Massey & Denton (1988) classification of segregation into five dimensions is largely due to a reliance on areal units. Reardon & O'Sullivan (2004) posit that if one knows the exact location of individuals, there are only two distinct dimensions of segregation: 1) spatial exposure, the "extent that members of one group encounter members of another group… in their local spatial environments" and 2) spatial evenness, the "extent to which groups are similarly distributed in residential space" (Reardon & O'Sullivan, 2004, p. 125). In either of these classifications, exposure, or the degree of potential contact between groups, benefits most from an individual-level measure (Lieberson, 1980; Massey & Denton, 1988; Reardon & O'Sullivan, 2004). While the other segregation dimensions pertain to the geographic location of groups, the exposure dimension focuses on capturing the how the average person experiences segregation (Bell, 1954; Massey & Denton, 1988; Reardon & O'Sullivan, 2004). In this sense, an individual-level measure of the exposure dimension creates a people-based representation, focused on people's exposures, compared to a place-based representation focused on the spatial distribution (Kwan, 2009).

Therefore, the SPI is an individual-level, people-based representation of exposure, more specifically, isolation.

Individuals potentially face racial isolation in more than one way (Hipp & Perrin, 2009; White, 1983). If a person must travel a far distance before encountering a person of another race, that individual is isolated by distance. Alternatively, if a person must encounter a large number of individuals of the same race before coming into contact with someone of a different race, that individual is racially insulated, regardless of the amount of distance traveled. While these two forms are related, they are not identical, and it is possible for an individual to experience one form of isolation without the other, for example in a sparsely populated rural area (White, 1983). Therefore, the SPI is comprised of two sub-measures to distinguish these two aspects, one that measures isolation by distance, the SPI-D, and the other by interpersonal contact, the SPI-C. In the rest of the article, the term SPI will refer to the overall measure, and SPI-C or SPI-D will refer to the specific sub-measure.

Figure 2.1 illustrates how the SPI works conceptually. To calculate the SPI, all individuals walk the shortest path along the street network from their residence to the residence of the first person of a different race. Utilizing street-networks allows for a more accurate measurement of exposure, as opposed to using Euclidean distance, which is known to overestimate coverage (Q. Brown et al., 2016; Gutiérrez & García-Palomares, 2008; O'Neill, Ramsey, & Chou, 1992). In studying segregation patterns in Philadelphia in 1880, Logan & Bellman (2016) note that "even using point data, a measure that is based on Euclidean distance rather than the actual street network would not reveal the clear spatial separation shown" (p. 699-700). Additionally, using street-networks allows for a more detailed picture of how individuals may behave and move around their neighborhood. Outliers were removed to prevent individuals from traveling to an "atypical" individual of a different race. Outliers are defined as individuals who live entirely separate from people of their own race. Through empirical testing, calculating the average distance to the ten nearest neighbors of the same race and removing all individuals whose average distance was more than three standard deviations above the mean proved to capture these individuals. Therefore, any individual that matched this criterion were considered outliers and exclude as potential points of contact. In the case of an interracial household, the closest destination is one's household. The SPI-D is how far that individual traveled to get to the destination and the SPI-C is the number of people of the same race that individual walked past before arriving at the destination. While this article focuses on the SPI's ability to measure racial residential isolation, it is worth noting that the SPI measures any point-level isolation. For example, the SPI can calculate the isolation among people of different socio-economic status or people of different nativity status. Furthermore, the SPI is limited to measuring people. The SPI could as easily measure the isolation among various types of criminal offenses or properties by land use. To automate the SPI calculation process, I wrote a Python script using ESRI's ArcMap with the Network Analyst Extension.

Figure 2.1: Illustrated Example of how the Shortest Path Isolation Index Works



Unlike other segregation measures, the SPI purposefully reports an absolute value, ranging from 0 to ∞. Given that the SPI is a people-based representation of isolation, the goal of the SPI is to provide a measure that captures the experience of an individual. Individuals often do not think in proportions or percentages but in raw numbers. While walking, individuals do not conceptualize the length of their current route in relation to the combination of all possible street segments in the city, but in terms of the actual distance traveled. Additionally, individuals do not think about the proportional share of the total population they have encountered on that route, but rather the sheer number of people they have walked past. An added benefit of having an absolute measure is that it allows researchers to determine their context-specific cutoff values for defining isolation based on their understanding of the local environment (Sin, 2002). By being an absolute measure, the SPI does not depend on any specific level at which segregation occurs, meaning that it abides by the arbitrary boundary independence criteria detailed by Reardon & O'Sullivan (2004). While absolute measures violate the composition invariance criteria which states that group composition should not alter the measure, researchers have largely debated if this condition is necessary in all cases (Jahn et al., 1947; D. R. James & Taeuber, 1985; Reardon & Firebaugh, 2002). Furthermore, violating the composition invariance criteria is not outside the tradition of exposure indices (Lieberson, 1980; Morrill, 1991). The asymmetrical quality of Lieberson's (1980) isolation index provides a more realistic interpretation because it takes into account that the probability of interaction or isolation differs by group size (Farley, 1984; Lieberson & Carter, 1982). Therefore, it is valuable for the SPI measure to reflect the fact that isolation may appear differently by group composition.

Before using the SPI, it is important to understand the assumptions implicitly held. First, the SPI assumes that individuals do not prefer where they travel other than that it minimizes distance. While, in reality, individuals travel multiple paths in their daily life, the SPI-D only calculates the hypothetically shortest path. In this sense, the SPI-D should be considered the minimum level of isolation by distance. Second, the SPI-C assumes that walking past someone's residence is the same as encountering those living there. Therefore, the SPI-C should be considered the maximum potential number of people one might encounter near one's residence.

Data Source:

Until recently, the ability to use individual-level data for geographic analysis was more of a dream than reality due to the lack of available data. However, in 2008, the Minnesota Population Center (MPC) in collaboration with the Church of Latter-Day Saints released harmonized complete data from the 1880 Census; and in 2012, released the first public complete dataset of the 1940 Census (Duke, 2012; IPUMS-USA, 2014; Ruggles et al., 2015). According to the "72-Year Rule" of the Census Bureau, all records are publicly available after 72 years, including house addresses (95th Congress, 1978; Weintraub, 2008). The 1940 Census is, therefore, the most recent census data that would allow for geocoding exact house addresses. Therefore, the data for this analysis comes from the 100% count of the 1940 Census. By geocoding exact addresses, the SPI avoids the potential pitfalls brought by relying on Agresti's (1980) assumption when analyzing individual-level data.

Study Area:

This study includes all non-institutional residents living in West Philadelphia in 1940, making for a total of 297,191 individuals and 82,141 households. West Philadelphia, encompassing an area of approximately 14.2 square miles, is defined as being bordered by the Schuylkill River on the east and north, Cobbs Creek and the county line on the west, and Baltimore Avenue to the south where the Baltimore Railroad tracks historically separated the residential area of the north from the industrial area of the south (Weaver, 1930). With close to half of the boundaries defined by natural resources and the other half defined by railroads and county limits, West Philadelphia is relatively self-contained. However, to avoid potential spurious errors resulting from boundary effects, the individuals living in the census tracts just south of Baltimore Avenue were included to perform the SPI calculation. These individuals do not receive an SPI score and are excluded in the subsequent analysis. Individuals along the other boundaries are excluded under the assumption that people do not cross county boundaries or across rivers.

Table 2.1 compares the descriptive statistics for West Philadelphia compared to the entire city of Philadelphia.

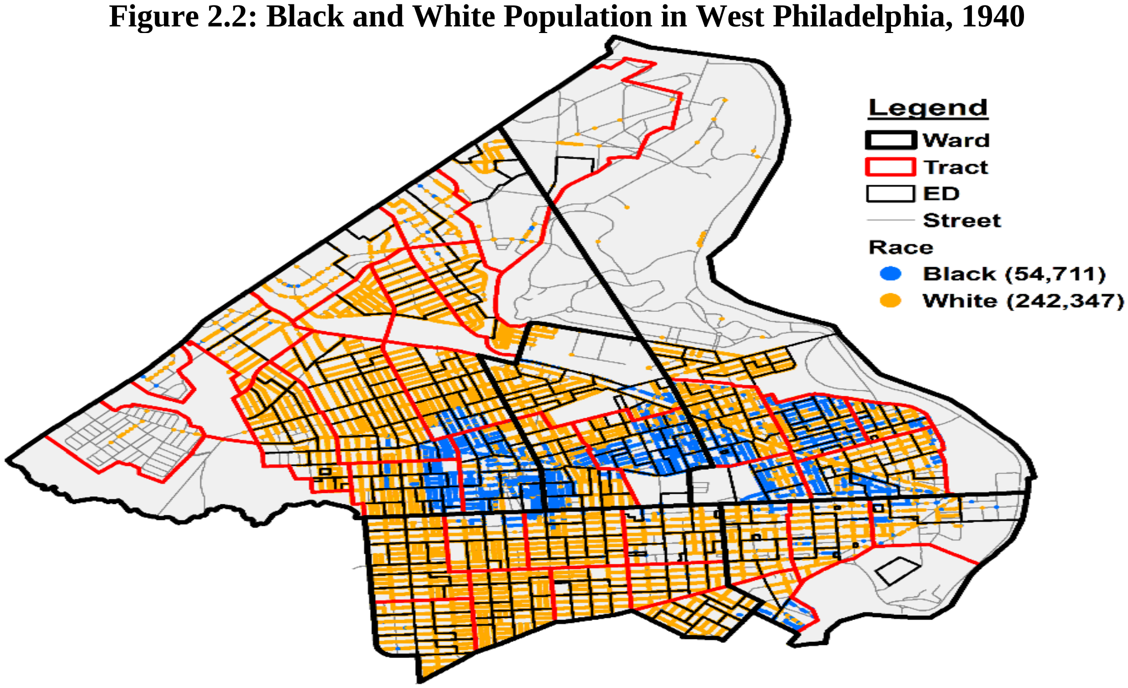
Table 2.1: Descriptive Statistics for Philadelphia, 1940

Variables	*West Philadelphia*	*Total Philadelphia*
White	81.5%	87.2%
Black	18.4%	12.7%
Male	46.9%	48.6%
Age (Median)	32	31
Currently Married	45.9%	46.0%
U.S. Born	85.5%	84.8%
Dwelling Size	4.6	4.7
In Labor Force	46.3%	45.8%
Homeownership	33.4%	38.9%
Years of Schooling (Median)	9	9
Family Income (Median)	$1,200	$1,200
Individuals	297,191	1,892,209
Households	82,141	510,999

Note: Only includes non-institutionalized population

West Philadelphia is similar to the city average in most sociodemographic variables, barring the notable exceptions of the racial breakdown and homeownership rate. Compared to the rest of the Philadelphia, West Philadelphia has lower homeownership rates and a larger Black population. Figure 2.2 shows the spatial distribution of individuals by race in West Philadelphia for 1940.

Figure 2.2: Black and White Population in West Philadelphia, 1940



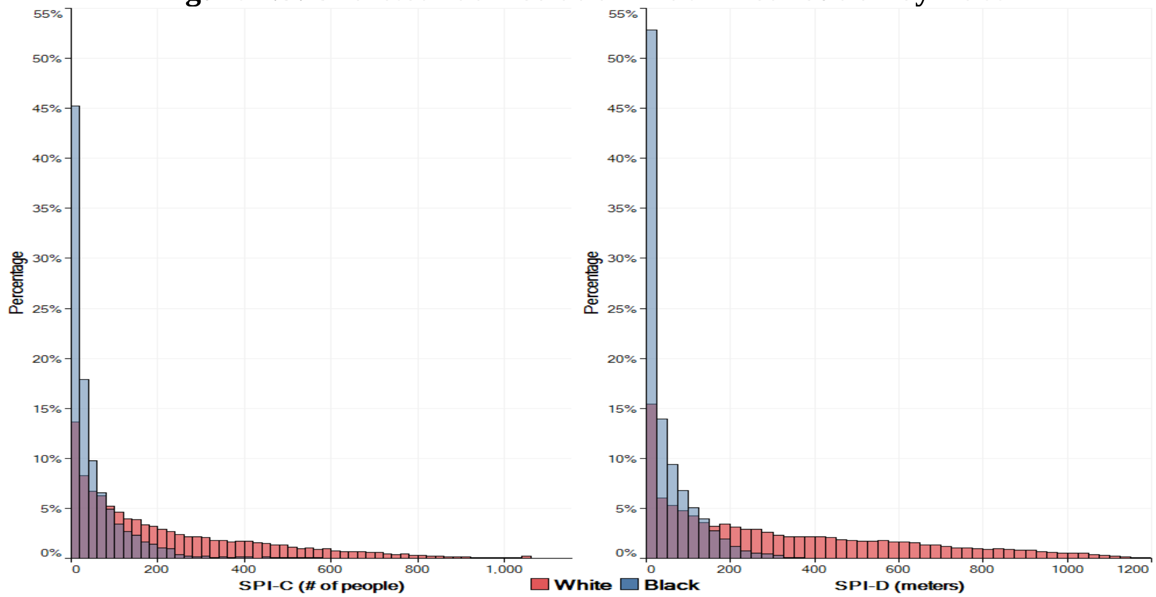
As the map in Figure 2.2 shows, most Black individuals are concentrated in a few census tracts, while White individuals are spread across the area. Using the census tract as the base unit, the White isolation index in West Philadelphia is 0.875, the Black isolation index is 0.447, and the index of dissimilarity is 0.613; traditional segregation indices would qualify this area as highly segregated (Massey & Denton, 1989, 1993; Sin, 2002). However, these measures may underestimate the true level of Black isolation. As shown in Figure 2.2, the census tract boundaries divide the Black neighborhoods into many census tracts. Had the tract boundaries been partitioned differently to contain that neighborhood in only one tract, the isolation index would undoubtedly be larger. Still, despite being racially isolated, Figure 2.2 shows that there are almost always a few people of the non-predominate race sprinkled throughout each tract. For these reasons, West Philadelphia is an interesting study area to study in analyze individual-level isolation.

The following section will provide descriptive statistics and detail the different ways in which to spatially display the SPI at the aggregate-level, highlighting the benefits and limitations of each approach.

Results:

Figure 2.3 shows the distribution of the SPI measures by race.

Figure 2.3: Shortest Path Isolation Index Distribution by Race



While the distribution of both populations is right-skewed, the tail of the distribution is much longer for the White population compared to the Black population. As the distribution shows, a small proportion of the Black population is racially isolated. Over 95% of the Black population have an SPI-C value under 100 people, with approximately 8% having an SPI-C value of exactly 0, meaning one's nearest neighbor in any direction is non-Black. Additionally, around 85% of the Black population have SPI-D value under 108 meters, which approximately equals the average length of a Philadelphia street segment. Conversely, a large proportion of the White population is racially isolated, both by distance and interpersonal contact. Approximately 60% of the White population has an SPI-C value over 100 people and 40% with values over 200 people. Additionally, 67% of the White population have an SPI-D value over 108 meters, with 25% having SPI-D values over half a kilometer. The SPI measures show that overall, the White population is more isolated, both by distance and interpersonal contact, than the Black population.

However, how does this distribution compare to that of a randomly distributed population? Table 2.2 compares the SPI by race between the observed population and a randomly distributed population where each household was randomly assigned to a new residence among the original set of dwellings. This approach randomizes the spatial racial distribution of households while maintaining the original residential zoning structures of the area.

Table 2.2: Comparison of Shortest Path Isolation Index Distributions between Observed and Random Population

	White					*Black*			
Variable	*Observed*	*Random*	*Ratio*	*t-Value*	*Variable*	*Observed*	*Random*	*Ratio*	*t-Value*
SPI-C	219.29 (214.80)	13.69 (26.24)	16.02	467.73***	SPI-C	50.59 (70.36)	2.66 (7.11)	19.02	158.53***
SPI-D	319.05 (296.45)	13.41 (45.92)	23.79	501.57***	SPI-D	48.47 (60.85)	5.60 (45.91)	8.66	131.55***
Count	242,347				Count	54,711			

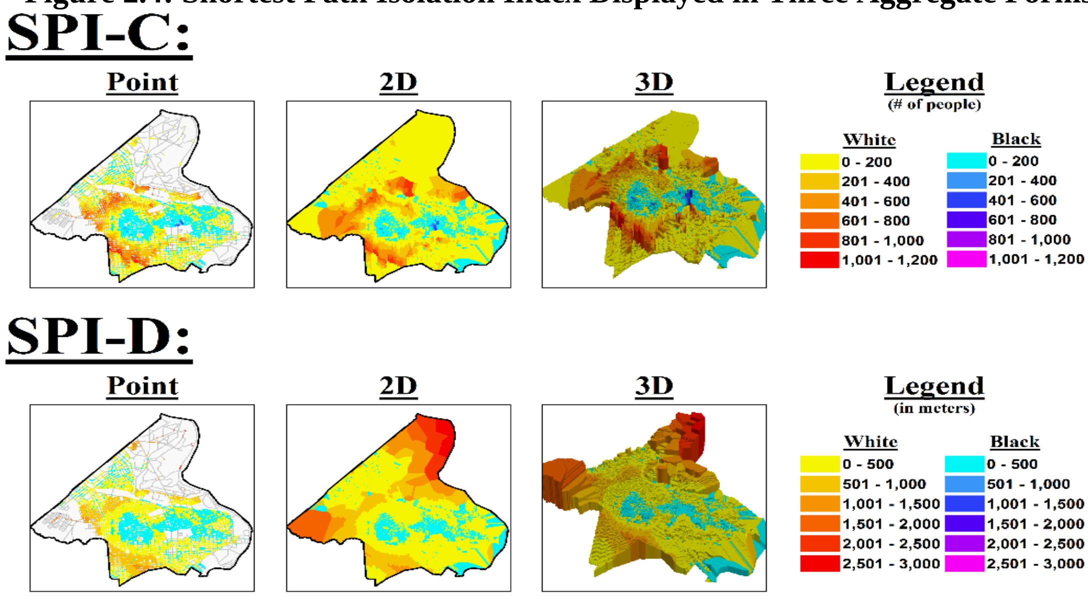
Standard Deviation in Parenthesis
* p<0.05 ** p<0.01 *** p<0.001

As shown in Table 2.2, the SPI measures for the observed population is dramatically larger than the random population. Compared to the random population, the mean value SPI-C for the observed population is around 16 times greater for the White population and 19 times greater for the Black population, while the average value SPI-D for the observed population is around 24 times greater for the White population and around nine times greater for the Black population. These differences in means between the observed and random population are all exceptionally highly significant, having t-values over 100 for the Black population and near 500 for the White population. This difference in SPI measures between the two populations indicates that the observed population is significantly more isolated than one could expect by simple random assortment.

Spatial Representation:

The SPI can be visually represented in three distinct ways. As shown in Figure 2.4, researchers can represent SPI data as points, point-extrapolate Thiessen polygons, or 3D polygons. While each of these representations depicts the same information, they each do so in slightly different ways that are worth noting.

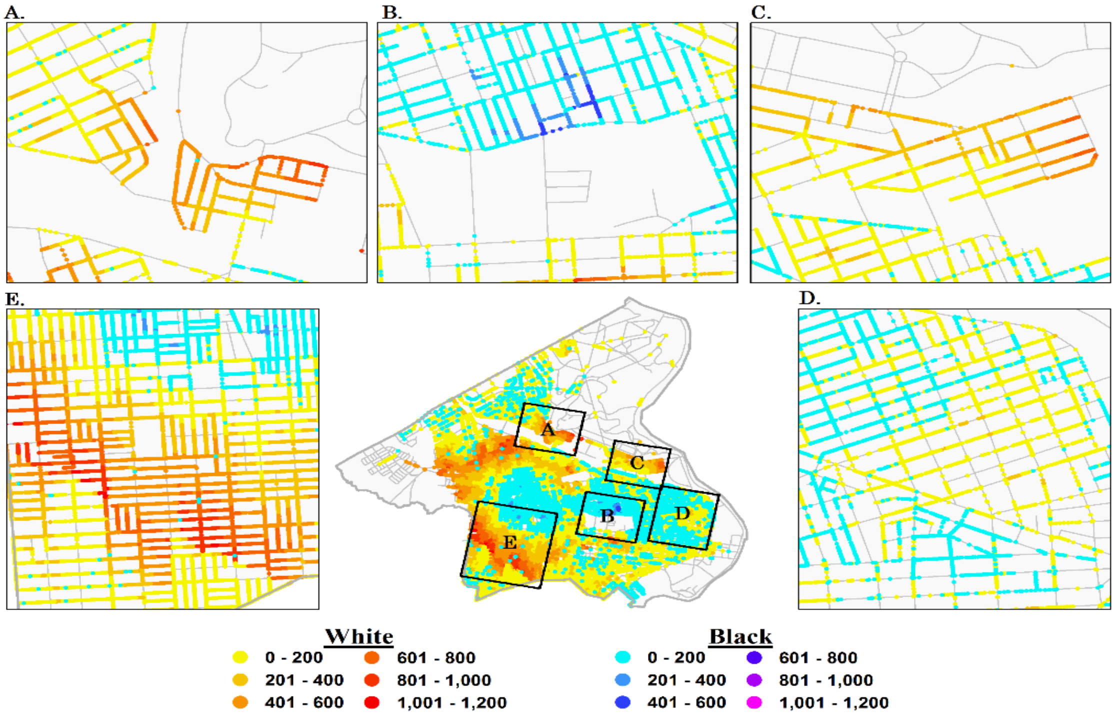
Figure 2.4: Shortest Path Isolation Index Displayed in Three Aggregate Forms



Point Data to Investigate Detailed Patterns:

Point data is the most accurate representation of the three forms because it shows the exact locations of individuals. Point data can display small incremental change from point to point. Therefore, the SPI point data is helpful in studying small sample areas and investigating possible micro-level explanations for patterns. Figure 2.5 shows the SPI-C for both White and Black individuals, highlighting five sample sites within West Philadelphia for closer examination. Each of these sites provides examples of how the residential pattern, coupled with the street network formation, can lead to varying levels of interpersonal isolation among select individuals.

Figure 2.5: SPI-C as Point Data



Area A illustrates an example of how the street networks affect isolation. While there are three locations where non-White individuals reside, there is no straight Euclidean path ("as the crow flies") to get there. For example, people living in the center of Area A must traverse along a single street to connect to where the nearest non-White person lives. In doing so, these individuals encounter more people than had they been able to traverse a straight path as if there were no streets or buildings in the way. This example provides further support for the need to consider the street networks when measuring isolation (Grannis, 1998). If the SPI had only measured Euclidean distance, like some other indices, it would have underestimated the level of isolation for this area.

Area B shows a predominately Black neighborhood. As one might expect, those living at the center of this neighborhood are more isolated than those on the periphery because the periphery acts as a buffer from non-Black individuals. Additionally, those at the center are blocked off on one side by the Pennsylvania Hospital, restricting the number of streets one could travel to encounter someone of a different race.

Area C illustrates an example of a “natural" street boundary. To the north of Area C is a park and to the east is a zoo followed by a river. This structure results in making those living closer to the park and zoo more isolated because they can only move in one direction to encounter a non-White resident, encountering more White individuals along the way.

Area D shows the importance of investigating isolation at different scales. On the street-level, this area is highly segregated, since most streets are almost either entirely White or Black. This finding matches similar findings to Logan & Bellman (2016) that Black and White population lived in close proximity but not on the same streets. However, while individuals may live on a racially isolated street, they do not need to walk by that many people before encountering someone of a different race. So, while street-level isolation may be high, individual-level isolation is low due to the intersection of these streets.

Area E shows a densely populated White neighborhood sandwiched between two Black neighborhoods. Those living closer to the non-White neighborhood have low levels of SPI-C. Living an additional block further from the non-White neighborhood puts more White individuals between them and the non-White residents, which is shown by the change in SPI-C from under 200 people to over 1,000. This high level of SPI-C is both due to the physical distance one would have to travel, as well as, the population density of the area that allows an individual to encounter a large number of people in a short amount of distance.

Visualizing the SPI-C as points provides evidence of the potential mechanisms which racially isolate people. In addition to living in racially separated areas, population density and the streets network patterns can also lead to high levels of isolation for individuals. This finding provides additional support for the role street networks have in orchestrating isolation (Grannis, 1998, 2005; Grigoryeva & Ruef, 2015). While point data allows for these type relationships to become more evident, this level of detail is not as helpful when investigating general trends for the entire study area. To do this, one should visual the data using point-extrapolate Thiessen polygons.

Thiessen Polygons Data to Compare Isolation Patterns:

Unlike point data, which can depict the small incremental change in isolation, the purpose of point-extrapolate polygonal data is to more easily display trends for the entire study area, making it useful for cross comparisons. While there are a few ways to extrapolate the point data to cover the study area, Thiessen polygons may be the best approach. Thiessen polygons— sometimes referred to as Voronoi polygons—partition an area based on the a set of points such that "each polygon bounds the region that is closer to one point than to any adjacent points" (Kennedy, 2000, p. 99). Therefore, it is better to think of Thiessen polygons as demarcations of a point's space rather than as areal units. Thiessen polygons are helpful because they cover the entire study area without having to interpolate values. Additionally, the size of the Thiessen polygon provides insights into the population density of the area since small Thiessen polygons would mean points are near each other (Brassel & Reif, 1979; Klein, 1988). For these reasons, Thiessen polygons are a helpful approach to representing spatial trends over a large study area.

Figure 2.6 displays the SPI measures as Thiessen polygons for White and Black individuals for West Philadelphia.

Figure 2.6: Shortest Path Isolation Index as Thiessen Polygons

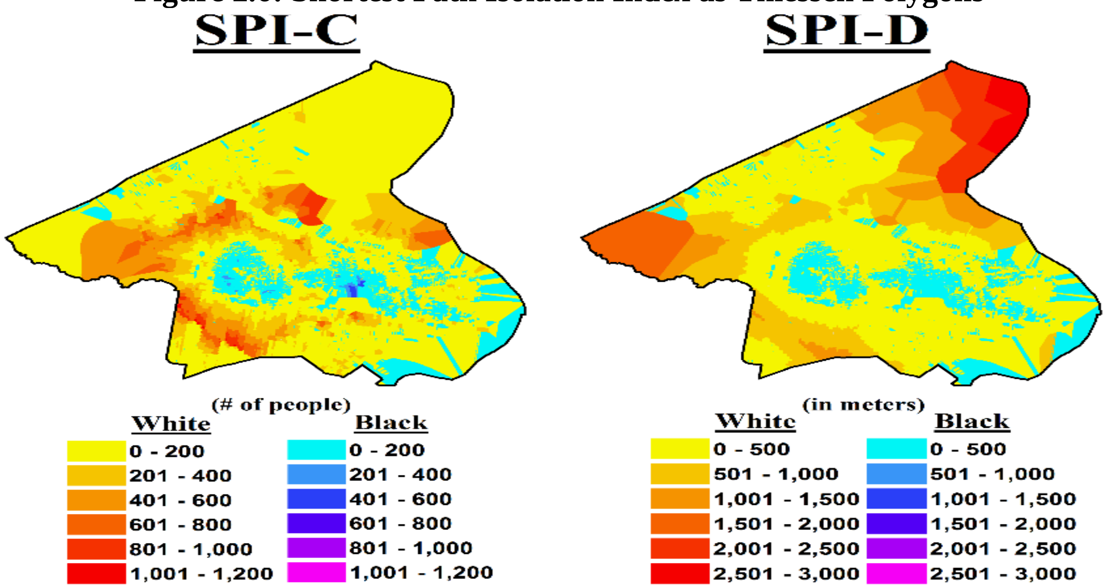


Figure 2.6 shows that most places with residents with high SPI-D values also have high SPI-C values. This finding makes sense; those that live a further distance from people of another race most likely encounter more people of the same race as well. However, this is not always true. Those with the highest SPI-D values are a small set of White individuals residing in the north corner of the study area known as Fairmount Park. Population density is sparse in this area, meaning that these individuals can travel a far distance without coming into contact with that many people on the way. It is important to point out that Fairmount Park is one of the few open spaces within West Philadelphia, which otherwise is quite residentially dense. While the SPI-C and SPI-D may not show dramatically different trends in an urban setting, this example of Fairmount Park may provide some

insights as to how the SPI measures may look differently investigating rural or semi-rural areas, which are underexamined areas often due to their large areal unit size (Butkiewicz et al., 2010; Krupka, 2007; Lichter, Parisi, Grice, & Taquino, 2007; Murdock, Hwang, & Hoque, 1994; Openshaw, 1996).

Looking by race, both SPI measures show that the White population is more isolated than Black individuals. Despite living in predominately Black areas, other than a few exceptions, no one Black person had to travel that far or encounter that many Black people before encountering a non-Black individual. There are only a few small areas where Black individuals have an SPI-C value over 200, whereas there are many areas where White individuals have an SPI-C value over 500. This finding matches the earlier findings when comparing the SPI distributions by race graphically.

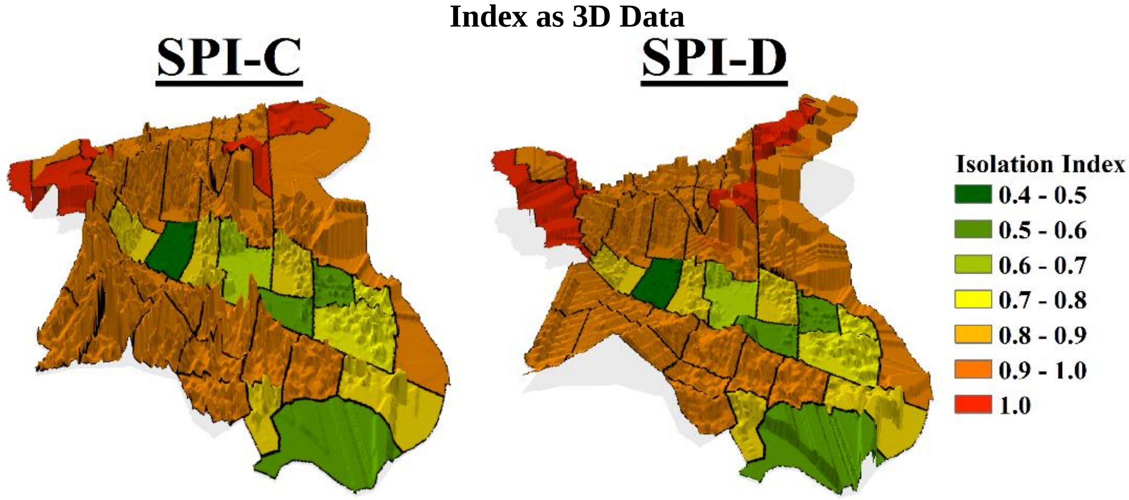
Visualizing the SPI data as Thiessen polygons shows that while some Black individuals are isolated, overall, the Black population is not isolated at the same level as White individuals. While this form of representation is helpful for comparing the SPI between different groups, it is less helpful for investigating bivariate relationships between the SPI with area-level characteristics that may be relevant for understanding isolation. For this analysis, a 3D representation is more useful.

3D Representation to Investigate Area-Level Characteristics:

One can easily turn the 2D representation into a 3D form by assigning the SPI value as the elevation of the study. This process topographically reimagines the study area as a series of "mountain ranges" of high isolation and "valleys" of close contact. Transforming the data into a 3D imagery has many benefits for studying area-level characteristics.

First, visually displaying bivariate relationships on a map is difficult. By assigning the SPI value as the elevation of the study area and assigning another variable as the color symbology, one can simultaneously inspect the relationship between two variables. Figure 2.7 shows the relationship between the SPI index and the isolation index. In this case, the elevation of the study area corresponds to the level of SPI of individuals living at that location, while the symbology corresponds to the traditional isolation index, ${}_xP_x$, popularized by Lieberson (1980) using the enumeration districts as the base unit so that each tract has a ${}_xP_x$ value ranging between 0 and 1. For visual assistance, the elevation for SPI-C is multiplied by a factor of 1.5 and the elevation for SPI-D is in meters.

Figure 2.7: Comparing White Shortest Path Isolation Index to Lieberson's Isolation Index as 3D Data



There is a visually clear relationship between both aspects of SPI index and the traditional isolation index. Individuals with the highest levels of SPI-C tend to live in the most isolated tracts and conversely. Additionally, those living in the most isolated tracts have slightly higher SPI-D than the other tracts. While the SPI and the isolation index are positively correlated, there is variability both within and across tracts. Figure 2.7 also shows that the relative decrease in average SPI-C values by tract is much greater than the relative decrease in $_xP_x$ values. Tracts with $_xP_x$ values of 1 have an SPI-C average of 308 people, which steadily decreases to 199 people for tracts with $_xP_x$ values between 0.8 and 0.9 and 60 people for tracts with $_xP_x$ values between 0.6 and 0.7. This difference points to the difference between a place-based and people-based representation of isolation. Even if an area is isolated, every individual that lives there may not necessarily be racially insulated. Visualizing data in 3D allows the researcher to visually inspect the relationship between the SPI measure and area-level characteristics, a task that could not be easily done using 2D representations. This 3D representation helps show that while the SPI measure picks up similar signals as the traditional isolation index, the two measures are not identical and the SPI can provide new insights about an area that the traditional measures cannot.

Second, using 3D representations helps provide a new conceptualization of boundaries and neighborhoods. Some of the variability within census tracts shown in Figure 2.7 comes from the fact that the SPI "mountain ranges" span across multiple tracts and stop without respect to tract boundaries. This finding highlights how tracts, or any area-level unit, may be arbitrarily dividing true boundaries. Given that the data is now reimagined as a mountain range, one can use the ArcMap's Surface Toolset to investigate the terrain and identify neighborhood boundaries and characteristics. By taking the slope, or first derivative, of the 3D representation, one could investigate how quickly isolation changes in an area. Additionally, one could identify the neighborhood boundaries of an area by taking the curvature, or the second derivative, of the terrain. In a sense, this process would delineate homogenous neighborhoods by identifying "where the mountain range meets the valley." Furthermore, if one has a sense of a cutoff value for the study area, one can create contour lines and investigate which parts of the study area of the study area have

SPI values above that cutoff. Using the 3D representation allows for more flexibility in identifying neighborhoods than relying on administrative boundaries.

Third, by using focal statistics, one could use the 3D representation for statistical testing. The focal statistics process performs statistical operations for a local neighborhood around each location, in a sense creating ego-centric neighborhoods (Chaix et al., 2009; Wade & Sommer, 2006). One can therefore calculate the average level of isolation for each local environment by calculating the focal mean of the SPI terrain. Then, one can calculate the local variance in isolation by subtracting each individual's SPI value from the average value calculated at that location. This process identifies those whose level of isolation significantly deviates from their local area. Additionally, one could repeat this process and subtract the focal mean from the global mean to identify which areas are significantly more or less isolated.

Reimagining isolation as 3D elevation expands the possibilities for how to investigate isolation. The ability to investigate bivariate relationships, identify contextually relevant neighborhood boundaries, and spatially identify the local variance in isolation are all valuable assets for investigation isolation.

All three of these forms of representation provide unique insights that are helpful for investigating aggregate trends and spatial patterns while simultaneously avoiding the MAUP. Each form acts as a different lens, allowing researchers to zoom in and zoom out of an area depending on the scale of interest. The beauty of relying on point data is that the SPI index is not limited to detailing space as a series of colored polygons as area-level measures are, but can detail space in a more fluid fashion covering the entire study area.

Discussion:

The Shortest Path Isolation index is a versatile tool for measuring isolation. The ability to represent data on the aggregate-level without introducing a modifiable areal unit problem is a valuable quality when studying segregation. Additionally, the capacity to measure individual-level isolation as either a function of distance or interpersonal contact makes the SPI useful in a variety of situations. Moreover, by using geocoded individuals and street networks for calculations, the SPI provides a more realistic measure for what an individual may experience while avoiding the need for any area-level assumptions. These factors make the SPI a strong candidate for measuring individual-level isolation.

While the SPI index is a people-based representation of isolation, given that most traditional segregation measures are place-based, it would be beneficial to have a place-based representation of the SPI index as well. There are a few ways in which to transform the SPI index into a place-based measure. The simplest approach is to calculate the mean SPI for a population residing in a specified area. As long as the area has contextual meaning, there is no risk of introducing the MAUP. However, this approach requires the researcher to have a sense of the contextually relevant boundaries. Another approach is to use the 3D representation of the SPI as a guide to designate empirically defined neighborhood boundaries. In either case, one can take the neighborhood average and compare it to the global average of the study area to identify how the level of isolation in the smaller local areas compare to the overall level of isolation. This method allows the researcher to gain a sense of the degree of isolation based on place.

Limitations:

Notwithstanding the advantages highlighted throughout this article, like all indices, the Shortest Path Isolation index is not without its limitations. First, the data requirements to use the SPI are currently limited. At the moment, there are only a few public datasets with exact house addresses and those that are available are often historical data. Additionally, the task of geocoding individuals from historical data is often labor intensive. Hopefully, as more individual-level data becomes more easily accessible, this limitation will become less of a concern. In the meantime, researchers should weigh the importance of using the SPI for their research question with the cost associated with calculating the SPI. Researchers may be able to perform analysis on a weighted population sample as long as the point data refers to individual exact addresses. Second, the SPI does not capture the full extent of an individual's isolation. The SPI calculates one path, while in reality, most people traverse multiple different paths in their daily lives. Therefore, the SPI is an underestimate of the level of isolation and represents the minimum absolute isolation an individual could experience. Future research may benefit from modifying the SPI to account for the shortest path to the n^{th} nearest person of a different race; however, the results may not change much due to the outlier correction performed by the SPI index.

Despite these limitations, the SPI remains an extremely powerful tool. With that said, the SPI does not stand alone. The SPI cannot replace the findings done by the area-level place-based indices that are already well established in the segregation literature. Nor does it intend to. Segregation, and isolation more specifically, does not occur at one scale, but rather works differently at multiple different unit-scales (Reardon et al., 2008). It is, therefore, important to be able to measure isolation at both the area-level and individual-level, the place-based and the people-based representation. In addition to developing new individual-level segregation indices, there still needs to be a push to modify traditional area-level segregation indices to avoid MAUP. In this sense, the SPI is not a segregation measure separate from the rest, but rather a tool that is a part of an arsenal of tools, each equipped to address different questions and segregation patterns ultimately providing new insights of segregation dynamics at all unit-scales.

Conclusions:

The SPI index is a powerful tool for detailing how an individual experiences racial isolation in his or her surroundings. By simulating individuals to walk to their nearest person of a different race, this measure gets closer to modelling people's behaviors and actions. However, this article has only highlighted one actualization of how to use the SPI index. While the destination calculated in this article was the first person of a different race, the SPI index can be expanded to measure the racial exposure along the paths people take while walking to different destinations in their neighborhood. For example, researchers may be interested in understanding the racial breakdown of the path to one's place of worship, or to the grocery store, or the bus stop. In adapting the measure in this fashion, this measure can get closer to capturing the full extent of exposures people experience in their daily life.

CHAPTER 3: IS AN INDIVIDUAL A GHETTO OR AN ENCLAVE?: HOW INDIVIDUAL-LEVEL RACIAL ISOLATION RELATES TO ACCESS TO NEIGHBORHOOD RESOURCES

Introduction:

From a sociological standpoint, the study of segregation is relevant for understanding how residential patterns relate to access to opportunities and resources (Charles, 2003; Morrill, 1991; White, 2015). Much of the segregation literature is rooted in a social ecological framework which argues that areas differentiate through a process of competition over scarce resources (Burgess, 1928; Park, Burgess, Mckenzie, & Wirth, 1925). As the minority population grows, segregation may occur as either a "threat response," where the majority group uses its relative advantage to enforce boundaries through discrimination, or as a "threshold response," where the minority group has reached a critical mass to form its own neighborhood (Alba & Logan, 1993; Boyd, 2001; Charles, 2003; Cloutier, 1984; Glazer & Moynihan, 1963; Lieberson, 1963; Marcuse, 1997; Massey & Denton, 1993). These responses create different types of segregated areas, such as the resource-deprived Ghetto, the resource-abundant Citadel, or the ethnically-congregated Enclave (Marcuse, 1997; Massey & Denton, 1993; Peach, 2009). How places are segregated spatially also matters for access to resources (Massey & Denton, 1988; Massey et al., 1996; Sundstrom, 2004). Whether the pattern is evenness, concentration, clustering, centralization, or exposure, Massey & Denton (1989) point out that a "high level of segregation on any one of these dimensions is problematic because it isolated a minority group from amenities, opportunities, and resources that affect social and economic well-being" (Massey & Denton, 1989, p. 373).

While segregation scholars continue to debate over the specific causes of segregation, the common fact remains that the consequences of segregation have to do with differential access to resources. As Taeuber & Taeuber (1976) argue, even if the unequal allocation of resources or population was a result of random processes, any future consequences would result from the current spatial pattern. Therefore, while understanding the processes that lead to residential segregation is important, understanding how the level of segregation relates to access to resources is important as well.

Literature Review:

Despite being deemed the "City of Neighborhoods," most Philadelphia researchers have focused on living conditions as it relates to the urban center, often overlooking the study of "the neighborhood" (Bauman, 1981; Savery, 1980). However, studying neighborhoods, separate from their relationship to the urban center, may also shed insights about residential segregation that would otherwise be missed. In this sense, West Philadelphia in 1940 is an interesting time and place to study residential segregation as it pertains to access to neighborhood resources.

During the pre-industrial period in the late nineteenth-century, much of Philadelphia's ethnic segregation could be explained by occupational segregation and workplace accessibility, with the major exception of the Black population (Ericksen & Yancey, 1979; Greenberg, 1980; Hershberg, Burstein, Ericksen, Greenberg, & Yancey, 1979; Hershberg, Cox, Light Jr., & Greenfield, 1981). During this period, the poor transit system made it difficult to commute, meaning that most people lived within walking distance from where they worked (Greenberg, 1980; Hershberg et al., 1979). Therefore, ethnic segregation was closely connected to occupational segregation (Hershberg et al., 1979). However, with the introduction of the electric trolley system in the 1890s and the elevated train in the early 1900s, people were able to move away from their place of work (Cheape, 1980; Cutler III, 1980; Gin & Sonstelie, 1992; Hershberg et al., 1979; Marsh, 1980). This technological advancement created new neighborhoods, known as "streetcar suburbs," where individuals could separate their work and residence (Adams et al., 1993; Warner Jr., 1987). West Philadelphia was one of those neighborhoods.

Over the next 30 years, West Philadelphia experienced rapid population growth. Being directly west of Center City, the Central Business District, West Philadelphia became one of the more desirable and fastest growing streetcar suburbs in Philadelphia (Cutler III, 1980; Marsh, 1980). The population tripled from around 45,000 in 1870 to around 130,000 in 1900 and then more than doubled again by 1930 with a population over 300,000 (Pierson, 1994). The Black population also benefited from moving to streetcar suburbs (Adams et al., 1993). With weaker ties between workplace and home and the availability of land, streetcar suburbs were less stable and established, meaning that Black individuals who had the earning power could move away from the poor living conditions of their previous residences into better living conditions in streetcar suburbs (Adams et al., 1993; Marsh, 1980; F. Miller, 1984; Morgan, 1983b; Warner Jr., 1987; Weaver, 1930). The result was a massive influx of Black people into West Philadelphia, starting with a population of under 2,000 in 1870 growing to a population of nearly 11,000 in 1910, doubling to almost 20,000 by 1920 and doubling again to 44,000 by 1930 (Pierson, 1994). The era before 1930 in West Philadelphia was one of abundance and growth.

However, while housing availability in West Philadelphia was abundant in the first few decades of the twentieth-century, as shown by the substantial growth in both the Black and White populations, the 1930s started showing signs of stagnation (Pierson, 1994). During this period, the most common architectural housing-type in Philadelphia was the row house, which comprised close to 90% of all houses in 1915 (Hershberg et al., 1979; Weaver, 1930). Row houses packed narrow streets, often sharing common walls, making it difficult for any future vertical growth (Hershberg et al., 1979; Marsh, 1973; R. Miller & Siry, 1980). So, by 1930, when most of the area in West Philadelphia was developed, there was little possibility for future development (Warner Jr., 1987). The Philadelphia City Planning Commission (1956) reported that by 1944, West Philadelphia only had 344 acres of unused open land left, approximately 4% of the land. With most of the available land taken, population growth stagnated, with the overall population only increasing by 706 people between 1930 and 1940 (Pierson, 1994). While the overall population failed to grow, the relative racial share of the population continued to change. Between 1930 and 1940, the Black population grew from 14.5% to close to 19% of the West Philadelphia population, often settling in already established Black areas, resulting in a pattern of increased racial concentration (Adams et al., 1993; F. Miller, 1984; Pierson, 1994). This pattern of change matches the theory of racial residential succession where the Black population replaces the White population in an area through a process of penetration, invasion, consolidation, and finally piling up (Burgess, 1928; Duncan & Duncan, 1957). Compositional change often leads to competition over resources (Burgess, 1928; Duncan & Duncan, 1957).

Consequently, 1940 represents the first time in which we can study West Philadelphia as it pertains to competition over resources. West Philadelphia is especially interesting because its residential pattern was not dictated by the local availability of jobs

(Marsh, 1973). Therefore, in this space, we can investigate how residential segregation is related to access to resources without having to consider work availability.

Neighborhood Resources:

In addition to access to jobs, there are other resources and institutions people may want to live near and therefore relate to segregation. Schafer Jr. (2014) states that "from the workplace to entertainment and leisure to shopping, different social classes had distinct patterns of participation and consumption" (Schafer Jr., 2014, p. 141) These differential individual patterns of participation and consumption may lead to differential residential patterns as well. Given their historical relevance to West Philadelphia, there are four resources that may be interesting to study: public transit; places of worship; leisure activities; and physician's offices.

West Philadelphians cared about is access to transit. Whereas living near jobs may not have been as important for residents in a commuter suburb like West Philadelphia, access to transit was undoubtedly important as it related to access to jobs. It is highly possible in West Philadelphia that there was unequal access to transit stops. While by 1940, the transportation system in Philadelphia was public, most of the original transit lines were constructed in the early 1900s by a private syndicate known as the Philadelphia Rapid Transit Company (PRT), later reorganized as the Philadelphia Transportation Company (PTC) (Cheape, 1980; Warner Jr., 1987). Just like the earlier private horsecar lines of the 1850s and private omnibus lines of the 1830s, the trolley and elevated track lines were designed for profitable gains and in many cases replaced where the horsecars and omnibuses once went (Cheape, 1980; Marsh, 1980; Warner Jr., 1987). As is the case with most privatized decision, this potentially means that transit stops were not distributed evenly throughout West Philadelphia and potentially favored those with purchasing power.

Places of worship and leisure activities also played a major role in Philadelphian society (Cutler & Gillette Jr., 1980; Marsh, 1980; Warner Jr., 1987). Considered essential for maintaining social cohesion and fostering local civic engagement, these religious and social institutions helped unify West Philadelphia residents (M. R. Lee & Ousey, 2005; Marsh, 1973, 1980). However, many White institutions were often exclusionary (Cutler & Gillette Jr., 1980; Marsh, 1973; Warner Jr., 1987). As a response, many of Black institutions self-segregated as well (Bauman, 1974). Lane (1992) noted that in the post-Civil War era, the Black population, regardless of class, "were joined together by an astonishingly thick network of churches, mutual-benefit associations, bands, teams, and social and political clubs" (p. 43). These practices resulted in segregated institutions, each catering to only one socioeconomic and racial/ethnic group (Warner Jr., 1987). In fact, Thomas (1980) argues that given the historical relationship between social class and church affiliation and that most church parsons lived near their church, one could conduct "journey-to-church" analysis, similar to the "journey-to-work" analysis, to identify ethnically and socially distinct neighborhoods. This claim may also be supported in the contemporary setting, as researchers have found evidence that certain churches, depending on their religious denomination, may play a role in segregating areas (Blanchard, 2007; Merino, 2011). Given this context, studying the accessibility to different leisure activities and places of worship by religious denominations may point to which types of institutions were the most inclusive and exclusionary.

Understanding access to health professionals is important as well. There has been a long-established connection between segregation and health (Kramer & Hogue, 2009). Residential segregation has been associated with higher mortality rates, poor pregnancy outcomes, higher rates of infectious and chronic illnesses, and poor self-rated health (Acevedo-Garcia, 2000; Culhane & Elo, 2005; Diez-Roux & Mair, 2010; Kramer & Hogue, 2009; Subramanian, Acevedo-Garcia, & Osypuk, 2005; T.-C. Yang & Matthews, 2015). While there are many proposed pathways for how segregation relates to health, one pathway is access to healthcare (Acevedo-Garcia, 2000). This may be because residential segregation is related to the demand for primary care services, creating “medical deserts,” areas with little to no medical resources (Gaskin, Dinwiddie, Chan, & McCleary, 2012a, 2012b). It is likely that West Philadelphians had differential access to physician’s offices by race. Like everyone else, many physicians moved out to West Philadelphia (Hershberg et al., 1981). As a result, these physicians moved their offices from the Central Business District to their neighborhood to practice medicine (Hershberg et al., 1981). With over 95% of the physicians in Philadelphia in 1940 being White, it is likely that most physician’s offices were situated in White neighborhoods (Schafer Jr., 2014). However, it is possible that Black neighborhoods had access to doctors as well. As the Southern Black population moved into the city center, many of the Black elites, including physicians, moved into West Philadelphia (Bauman, 1974). Therefore, while it is most likely the case that residential segregation is related to doctor availability, there is a chance it was not as large a problem as one might expect.

Studying the level of accessibility of all four of these resources, transit locations, places of worship, leisure activities, and physician’s offices, in West Philadelphia has the potential to illuminate interesting residential patterns that may not necessarily match that of access to employment. However, before one can start to study these patterns, one must first identify the relevant unit-scale of analysis (B. A. Lee et al., 2008).

Unit-Scale of Analysis:

Despite the fact that most ecological models are specified at the individual-level, most research has used aggregate-level place-level data for their unit-scale of analysis (Alba & Logan, 1992). However, using aggregate-level data has potential pitfalls. In addition to the modifiable areal unit problem (MAUP), which is discussed at length in the previous chapters, aggregate-level data analysis faces interpretation problems as well (Charles, 2003; Openshaw, 1989).

When studying the impact of segregation, the desire is to know how the location-specific forces relate to individuals' outcomes above and beyond their individual-level characteristics (Macintyre, Ellaway, & Cummins, 2002). However, as Cloutier (1984) argues, the link between the theoretical models and empirical findings are weak. While researchers want to capture the larger structural forces of an area, most are constrained to rely on compositional measures, aggregate measures of individual-level characteristics, treating them as proxy measures for neighborhood attributes (Diez-Roux, 2001, 2004; Diez-Roux & Mair, 2010; Pickett & Pearl, 2001). This forces researchers to move away from strong causal claims and move toward weaker inferences of how the presence of different sociodemographic groups may help to explain segregation patterns (Bayer, McMillan, & Rueben, 2004; Carrington & Troske, 1997; Cloutier, 1984). However, this approach poses problems. From a theoretical standpoint, using compositional measures implies that the collective socioeconomic characteristics of an area influence individuals' outcomes separately from their individual characteristics, which most researchers do not believe (Kawachi & Berkman, 2003; Macintyre et al., 2002). Therefore, this approach leads to a disconnect between the empirical findings and sociological reasoning.

Another common approach is to assume that all residents living in the same areal unit have the same average socioeconomic characteristics and experiences (Bell, 1954; Kramer et al., 2010). However, given that aggregate correlations are not substitutes for individual correlations, this type of assumption could lead to an ecological fallacy, erroneous conclusions that individual behaviors are the same as aggregate results (Robinson, 1950; Wong, 2004). For example, Ericksen & Yancey (1979) discovered that their aggregate-level analysis of manufacturers journey-to-work underestimated the true distances individuals had to travel for work. While some researchers have argued that aggregate-level data can provide an unbiased estimate of individual-level relationships under the right conditions, Alba & Logan (1992) find that this claim is not often true in practice (Firebaugh, 1978). While Alba & Logan (1992, 1993) proposed incorporating microdata samples with the aggregate data to resolve this problem, this approach is not plausible in all circumstances and may introduce problems of its own. Methodologically, it is difficult, if not impossible, to separate area-level and individual-level effects when both measures are derived from the same individual-level characteristics (Diez-Roux & Mair, 2010). If one is not careful, including individual-level data with aggregate-level data risks introducing multicollinearity among the explanatory measures (Alba & Logan, 1992; Charles, 2003).

Therefore, while it is theoretically possible to use area-level place-level data to investigate the relationship between segregation and individual-level outcomes, there are many analytical landmines along the way that make the process difficult.

One solution to this problem is to reduce the segregation unit-scale to the level of the individual household (Ellis et al., 2012). This process shifts the unit of analysis from place-based to people-based representation (Kwan, 2009). By focusing on the individual household, researchers can define ego-centric neighborhoods, individual-specific neighborhoods centered around the individual's local area of exposure, meaning that researchers do not need to assume that every individual in an area has the same set of exposures (Chaix et al., 2009). Additionally, as long as the neighborhood characteristics included in the analysis are not derived from the specific individual's characteristics, this process would allow researchers to control for individual-level characteristics separately from area-level characteristics, avoiding both ecological fallacy and multicollinearity concerns while simultaneously being better aligned with ecological segregation theory (Alba & Logan, 1992; Charles, 2003). Therefore, by shifting from a place-based representation to a people-based representation, analysis has the potential to more accurately measure the level of exposure experienced by individuals, thereby providing a better understanding of the impact of segregation (Kwan, 2009).

Methodology:

Study Area:

This study focuses on West Philadelphia in 1940. West Philadelphia, encompassing an area of approximately 14.2 square miles, is defined as being bordered by the Schuylkill River on the east and north, Cobbs Creek and the county line on the west, and Baltimore Avenue to the south where the Baltimore Railroad tracks historically separated the residential area of the north from the industrial area of the south (Weaver, 1930). With close to half of the boundaries defined by natural resources and the other half defined by railroads and county limits, West Philadelphia is relatively self-contained.

Data Source:

The population information for this analysis comes from the 100% count of the 1940 Census, provided by the Minnesota Population Center (MPC) (Ruggles et al., 2015). According to the "72-Year Rule" of the Census Bureau, all records become publicly available after 72 years, including house addresses, meaning that all dwellings in this study could be geocoded to their exact house address (95th Congress, 1978; Weintraub, 2008). This study includes all 82,141 non-institutional household dwellings residing in West Philadelphia in 1940.

Table 3.1 shows the descriptive statistics of household heads by race.

Table 3.1: Descriptive Statistics of Household Head by Race

Variables	White	Black	Other*
Home Owner	36.5%	17.1%	12.0%
Married	73.4%	77.4%	50.0%
US Born	70.9%	98.2%	34.0%
In Labor Force	78.2%	78.7%	86.0%
Per Capita Income	$492.40	$303.69	$172.38
Total	68,921	13,170	50

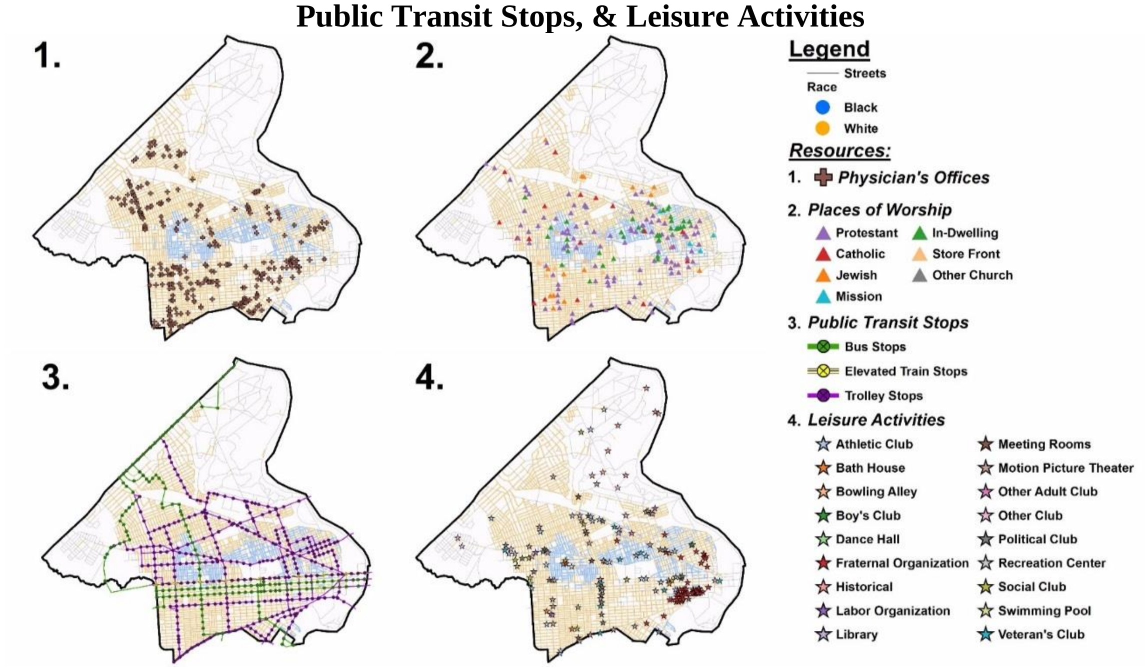
*Other Race: Asian/Pacific Islander and American Indian/Alaska Native

On average, White and Black household heads had similar levels of being married and being in the labor force. White household heads were more likely to be homeowners and have a higher annual per capita income than Black household head, while Black household heads were likely to be U.S. born than their White counterparts.

The locational information comes from three sources of data. The places of worship, leisure activities, and physician's offices come from the 1939 Social Base Map of Philadelphia (Works Progress Administration (WPA), 1939). The Social Base Map is a land use survey conducted by the Works Progress Administration (WPA) in combination with the Council of Social Agencies and the Philadelphia Board of Education to detail the locations of relevant institutions and buildings in the Philadelphia area (Works Progress Administration (WPA), 1939). The transit line information come from two maps: the 1932 Philadelphia Rapid Transit Company map and the 1944 Philadelphia Transportation Company map (Philadelphia Rapid Transit Company, 1932; Philadelphia Transportation Company, 1944).

Figure 3.1 shows the locations of the physician's offices; places of worship, by religious denomination; transit lines, by mode of transportation; and leisure activities, by activity in West Philadelphia, as well as the residential pattern by race.

Figure 3.1: Map of West Philadelphia's Physician's Offices, Places of Worship, Public Transit Stops, & Leisure Activities



As Figure 3.1 shows, each resource has a different spatial pattern in West Philadelphia. Most of the physician's offices are situated in White areas, with very few offices in the predominately Black areas. Most leisure activities appear to concentrate on a few select streets; the Philadelphia City Planning Commission (1956) classified these streets as commercial land use. While there are not many leisure activities directly in the Black area, many sit just outside of Black areas. A large portion of West Philadelphians live within close proximity to a transit stop, however, which mode of transportation differs by race. Black individuals appear to live near trolley lines or the Market elevated track, whereas some White individuals are near trolley tracks and others near bus lines. While the Black areas appear to lack the other three resources, there is a large concentration of places of worship, whereas there are some White areas that lack a place of worship. Most churches in Black areas appear to be Protestant, in-dwelling, or store front. Given that these spatial patterns differ, it may be possible that the processes that created them differ as well.

While these patterns are clear to see as point-level data, the pattern would be much more difficult to identify had the information been provided as a count within the typical administrative boundaries. As shown in Figure 3.1, there are many enumeration districts and census tracts that do not have physician's offices, places of worship, or a leisure activity. Additionally, many of the transit routes run along the shared border of many areal units, making it difficult to discern which unit to assign to them. Moreover, when there are resources in an areal unit, they are rarely distributed evenly throughout the space, often clustered along major streets. Therefore, without this detailed level of information, it is likely that our understanding of the true level of access would be obscured using areal units.

Had this analysis been conducted two decades earlier, the results would have been even more obscured since census tract information was not available, only ward information, where 5 wards covered all of West Philadelphia (Pierson, 1994).

Analysis:

For this project, I conduct a series of logistic regressions to investigate the relationship between household-level isolation and the likelihood that the household's ego-centric neighborhood includes the select resource, controlling for a set of characteristics of the household head. A resource is within a household's ego-centric neighborhood if it is within a quarter of a mile walking along the street network. The literature commonly uses threshold distances between one fifth and one half of a mile to represent a "pedestrian neighborhood," with a quarter mile being the most common distance (Atash, 1994; Aultman-Hall, Roorda, & Baetz, 1997; Q. Brown et al., 2016; Hoehner, Brennan Ramirez, Elliott, Handy, & Brownson, 2005; B. A. Lee et al., 2008; McCormack, Giles-Corti, & Bulsara, 2008; Millward, Spinney, & Scott, 2013; Nagel, Carlson, Bosworth, & Michael, 2008; Wineman et al., 2014). While some authors have challenged the quarter mile cutoff point, Millward et al. (2013) argue that this distance is partially validated with empirical evidence (Larsen, El-Geneidy, & Yasmin, 2010; Y. Yang & Diez-Roux, 2012). Additionally, in dense urban areas, access to resources is greater, meaning that using distances much larger than a quarter mile may result in everyone having a resource in their ego-centric neighborhood, leading to problems with estimation (Q. Brown et al., 2016). I use street network-distance instead of Euclidean distance because the latter has been found to overestimate coverage compared to network-distance (Q. Brown et al., 2016; Gutiérrez & García-Palomares, 2008; O'Neill et al., 1992).

To investigate a people-based representation of isolation, this analysis needs a segregation measure that can use individual-level data and produce segregation values separately for each household. While there are some segregation indices that use individual-level data in their calculations, such as the Sequence Index of Segregation (SIS) and the k-nearest neighbor approach, the value calculated is only relevant when aggregated to an area-level (Grigoryeva & Ruef, 2015; Östh et al., 2015). Additionally, while there are other segregation indices that calculate a localized measure of segregation, such as the enclave classification typology and the local spatial segregation index , they often rely on area-level composition for calculation (Poulsen, Johnston, & Forrest, 2001, 2002; Wong, 2002a). Therefore, neither of these types of indices are sufficient to conduct this type of analysis. One potential measure for this analysis is the neighbor-based measure, developed by Logan & Parman (2017). As part of its calculation, the neighbor-based measure determines for each household if the immediate next-door neighbor is of a different race (T. D. Logan & Parman, 2017). Therefore, although not the author's initial intent, the neighbor-based measure could theoretically be used in individual-level regression analysis as a binary variable (T. D. Logan & Parman, 2017). However, this analysis may be limited because the neighbor-based measure can only determine if the immediate next-door neighbor is of a different race, which is a somewhat limited view of the full extent of isolation a household may experience. Therefore, the only measure capable of using individual-level data to produce individual-level segregation values and captures a more fuller extent of segregation is the Shortest Path Isolation (SPI) index. For this reason, this project uses the SPI index to measure household-level isolation.

The SPI index, which is described in detail in the previous chapter, is somewhat unique of a measure in that it can be used to investigate isolation on the aggregate-level as well as be used in individual-level regression analysis. The SPI index measures isolation in two separate ways: distance and interpersonal contact. The SPI-D measures isolation as a function of the distance one would have to travel along the street network before encountering someone of a different race. The SPI-C measures isolation as a function of interpersonal contact, or the number of individuals of the same race one would have encountered along that path. One major advantage of the SPI index over area-level indices and other proposed individual-level indices is its use of the street-network to calculate isolation. By utilizing individual-level location data and the street-network, the SPI index provides a more realistic measure of each specific household's level of racial isolation. In this way, the SPI index provides a contextually relevant measure of an individual household's exposure.

Each logistic regression controls for a set of household characteristics including whether the household head was: a homeowner, married, U.S. born, and in the labor force participation and the total household income divided by the number of household members (henceforth referred to as per capita income).

Logistic regressions are run separately by race of the household head. This analysis is repeated for all four resources: physician's offices, places of worship, leisure activities, and transit stops. This analysis is then repeated by place of worship denomination, mode of transportation, and type of leisure activity to see if there were any difference by subcategory. The Social Base Map did not provide any additional information about the type of physician's office, so there is no easy way of disaggregating that information. For visual inspection, predicted margins are plotted at the entire range of SPI-C and SPI-D values, controlling the other household characteristics at their race-specific means.

While using a quarter mile to define the ego-centric neighborhood is consistent with prior literature, this decision is somewhat arbitrary (Larsen et al., 2010; Watson, Carlson, Humbert-Rico, Carroll, & Fulton, 2015; Y. Yang & Diez-Roux, 2012). Individuals may differ in how far they consider their neighborhood and may be more or less willing to travel certain distances depending on the specific resource (Y. Yang & Diez-Roux, 2012). Misspecification of the ego-centric neighborhood distance can lead to erroneous conclusions about the relationship between isolation and access. Therefore, as a robustness test, I repeat this analysis at lengths ranging from 0.01 to 1 mile to investigate how one's choice of distance impacts this study's findings.

Results:

Table 3.2 shows the percentage of Black and White households that are within a quarter mile walking distance from the set of neighborhood resources.

Table 3.2: Percentage of Households Living within a Quarter Mile of Select Resources, by Race

Resource	Black	White	t-value
Physician's Offices	76.1%	88.0%	-36.71
Places of Worship	98.2%	85.2%	41.42
Protestant	84.6%	70.5%	33.45
Catholic	24.4%	21.1%	8.52
Jewish	6.6%	18.0%	-32.84
Mission	8.2%	5.7%	11.03
In-Dwelling	71.7%	23.4%	118.74
Store Front	31.7%	12.6%	56.22
Other Church	7.9%	3.3%	24.81
Public Transit Stops	96.9%	97.7%	-5.53
Bus Stops	36.8%	52.0%	-32.29
Elevated Train Stops	16.5%	9.4%	24.61
Trolley Stops	96.9%	90.6%	23.69
Leisure Activities	82.6%	72.7%	23.95
Athletic Club	23.8%	17.5%	16.99
Bath House	1.0%	6.2%	-24.51
Bowling Alley	0.8%	1.2%	-4.47
Boy's Club	3.6%	4.1%	-2.68
Dance Hall	2.5%	2.3%	0.85
Franternal Organization	27.9%	15.3%	35.44
Historical	0.0%	0.0%	-1.24
Labor Organization	0.5%	3.5%	-18.17
Library	0.9%	6.0%	-24.25
Meeting Rooms	11.6%	9.0%	9.48
Motion Picture Theatre	45.4%	41.8%	7.63
Other Adult Club	3.1%	5.2%	-10.33
Other Club	2.3%	4.3%	-10.92
Political Club	17.3%	20.6%	-8.68
Recreation Center	6.2%	4.9%	6.08
Social Club	12.1%	10.2%	6.84
Swimming Pool	6.4%	0.4%	55.68
Veteran's Club	20.4%	10.6%	31.91

As shown in Table 3.2, except for a few select leisure activities, Black and White households have significantly different levels of access to neighborhood resources. Given this differential access between Black and White households, it would be interesting to investigate if more isolated households have more or less access to these resources by race.

Table 3.3 shows the results of the logistic regressions for individual households, by race, controlling for household characteristics.

Table 3.3: Logistic Regression of SPI on being within a Quarter Mile of Resources

Variables	***Physician's Offices***			
	White		**Black**	
SPI-C, per 25 people	0.976***		0.705***	
	(-8.74)		(-11.03)	
SPI-D, per 100 feet		0.940***		0.527***
		(-29.10)		(-19.16)
Observations	68,875		13,153	
Pseudo R-Squared	0.013	0.034	0.011	0.027

Variables	***Places of Worship***			
	White		**Black**	
SPI-C, per 25 people	0.891***		6.320***	
	(-56.75)		(6.60)	
SPI-D, per 100 feet		0.846***		2.335*
		(-91.06)		(2.41)
Observations	68,875		13,153	
Pseudo R-Squared	0.068	0.178	0.038	0.015

Variables	***Leisure Activities***			
	White		**Black**	
SPI-C, per 25 people	0.861***		1.391***	
	(-73.33)		(7.70)	
SPI-D, per 100 feet		0.881***		1.392***
		(-82.53)		(7.36)
Observations	68,875		13,153	
Pseudo R-Squared	0.092	0.099	0.011	0.011

Variables	***Public Transit Stops***			
	White		**Black**	
SPI-C, per 25 people	0.924***		0.556***	
	(-17.95)		(-11.49)	
SPI-D, per 100 feet		0.900***		1.180
		(-28.89)		(1.32)
Observations	68,875		13,153	
Pseudo R-Squared	0.039	0.083	0.047	0.015

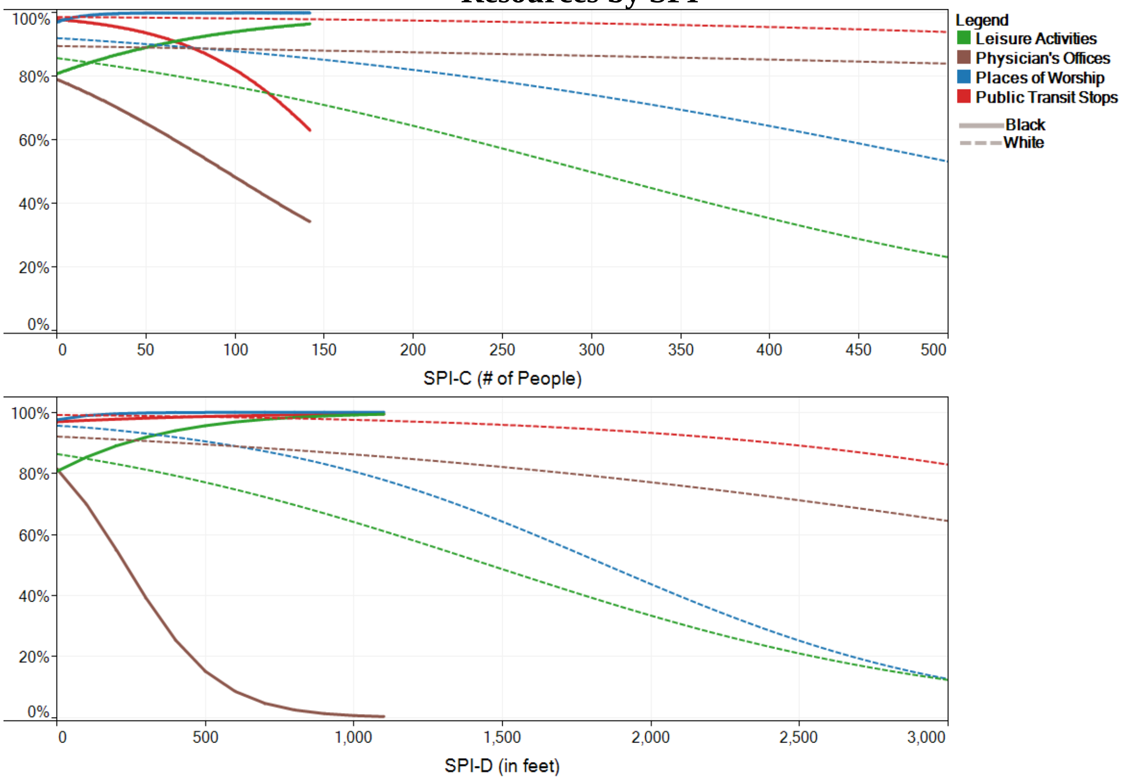
Note: Models controls for whether household head is a homeowner, married, US born, in labor force, and household per capita income
Coefficients expressed as Odds Ratios; t-statistics in parentheses
* $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$

The relationship between SPI and access is statistically significant for most resources, only Black SPI-D for public transit is non-significant. In the case of White households, an increase in isolation is associated with a decrease in access to all four resources. For Black households, an increase in isolation is negatively associated with access to physician's offices and public transit stops, at least for SPI-C, but is positively associated with access to places of worship and leisure activities. These findings show that isolation is related to access, but not necessarily in the same way for every resource.

Figure 3.2 through Figure 3.5 display the predicted margin plots of the change of SPI on being within a quarter mile of the select resource, holding the other covariates at their race-specific means [Figure 3.7 through Figure 3.10 display the standardized predicted margins plots in the Appendix]. The margin plots allow for inspection of how isolation relates to access throughout the full spectrum of isolation values found in this study.

Figure 3.2 shows the predicted margins for access to the four resources by race and SPI sub-measure.

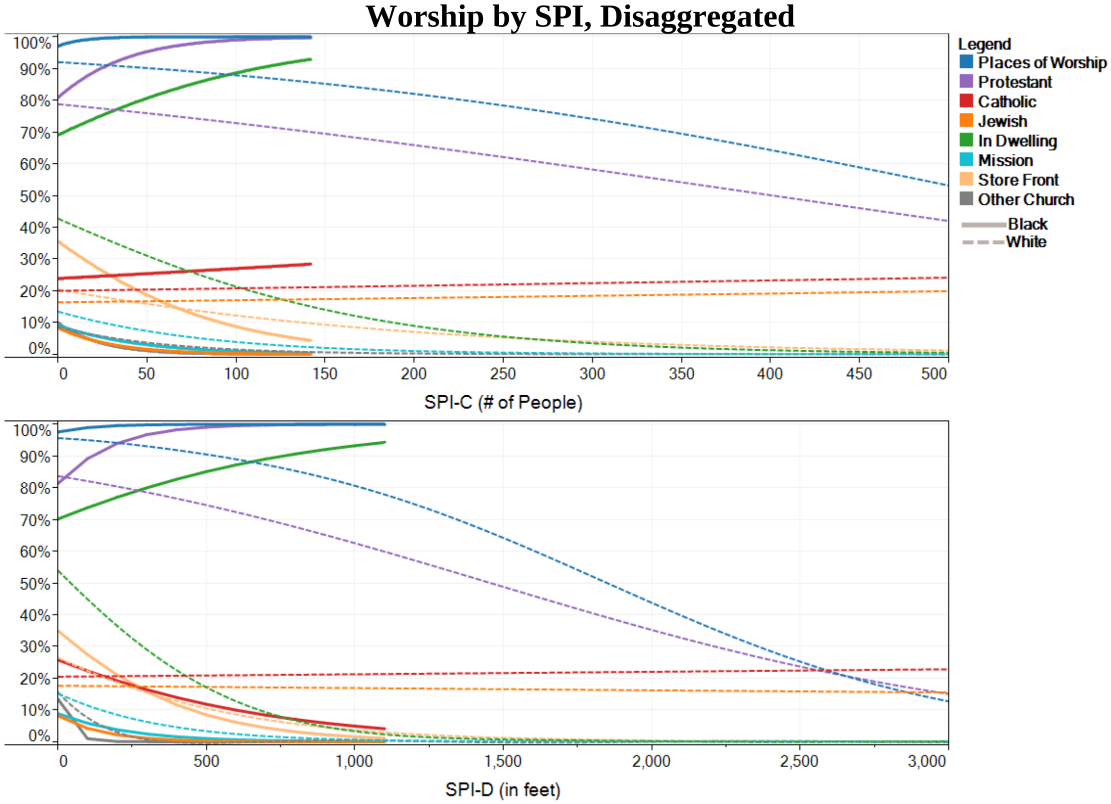
Figure 3.2: Predicted Probabilities of being within a Quarter Mile of Select Resources by SPI



For White households, the predicted probability of being near a place of worship declines by nearly one third for those with an SPI-C value above 400 people or an SPI-D value above 1,500 feet, while the predicted probabilities for Black households increase to 100% of the households having access. For White households, the predicted probability of being near a leisure activity also decrease by nearly one half for those with an SPI-C value above 350 people or an SPI-D value above 1,750 feet, while the predicted probabilities increase by around 10% for Black households. This finding may indicate that isolated Black households were more likely to live near commercial and local business areas than isolated White households, matching early social ecological theory that Black households live closer to the urban center (Burgess, 1928; Park et al., 1925). Interestingly, an increase in isolation is associated with a decreased probability of being near a physician's office for both White and Black households, however, the probability starts about 10% lower for Black households and decreases faster as well. The overall trends show that those who are more isolated have a much different likelihood of access than the mean household. Figures Figure 3.3 through Figure 3.5 further investigate these relationships by looking at how these relationships hold once the resources are disaggregated.

Figure 3.3 shows the predicted margins for access to a place of worship disaggregated by religious denomination.

Figure 3.3: Predicted Probabilities of being within a Quarter Mile of a Place of Worship by SPI, Disaggregated

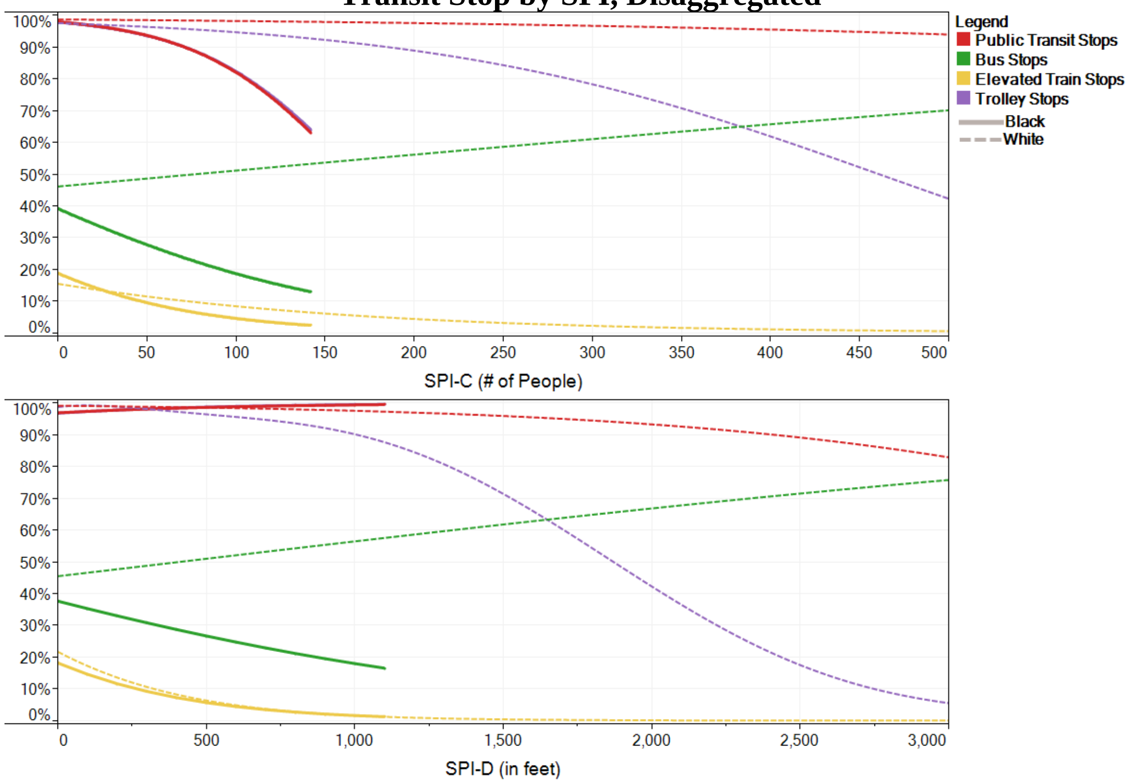


The predicted margins show that there is differential access to places of worship by race depending on the denomination. As one might expect, given their dominance in the area, Protestant churches are the main determining factor if one has access to place of worship, where an increase in isolation is positively associated with access for Black households and negatively associated for White households. Access to a Jewish synagogue increases slightly as White households increase SPI-C, but remains at a constant level for SPI-D. This finding may point to a threshold effect where isolated Jews were large enough to create their own enclave with a synagogue. Russian-born White households, who were often Jewish, had a much higher access to synagogues, around 43% compared to 18% for all White households (Marsh, 1980; F. Miller, 1984). Additionally, while the likelihood of access to a Catholic church changes slightly as isolation increases, it remains at relatively similar levels for both White and Black households. This finding may match contemporary findings which state that Catholic churches are more accepting of racial diversity than certain Protestant churches and while there may be Black and White Catholic churches, there are also Catholic churches with multiracial congregations as well (Blanchard, 2007; Merino, 2011). Probably one of the most striking differences relates to access to in-dwelling churches. Practically no White households with an SPI-C value above 200 people or SPI-D value above 1,000 feet have an in-dwelling church in their ego-centric neighborhood. On the contrary, over 70% of Black households have an in-dwelling church in their ego-centric neighborhood at an SPI values of 0 people or 0 feet; that percentage increases as isolation increases. This difference most likely arises from the fact that in-dwelling churches have a historical significance in Black communities (Cnaan, Boddie, McGrew, & Kang, 2006; Kostarelos, 1995). What is interesting is that the same pattern is not found for store front churches, which declines as isolation increases, which also have a historical significance in Black communities (Cnaan et al., 2006; Collins, 1970; I. E.

Harrison, 1966). However, this finding matches more contemporary findings that while Philadelphian storefront churches are comprised mainly of Black patrons, they were not located in predominantly Black areas, but rather on major streets (Cnaan et al., 2006). These findings show that, in most cases, the composition of places of worship appear to change as isolation increases.

Figure 3.4 shows the predicted margins of access to transit stops by mode of transportation.

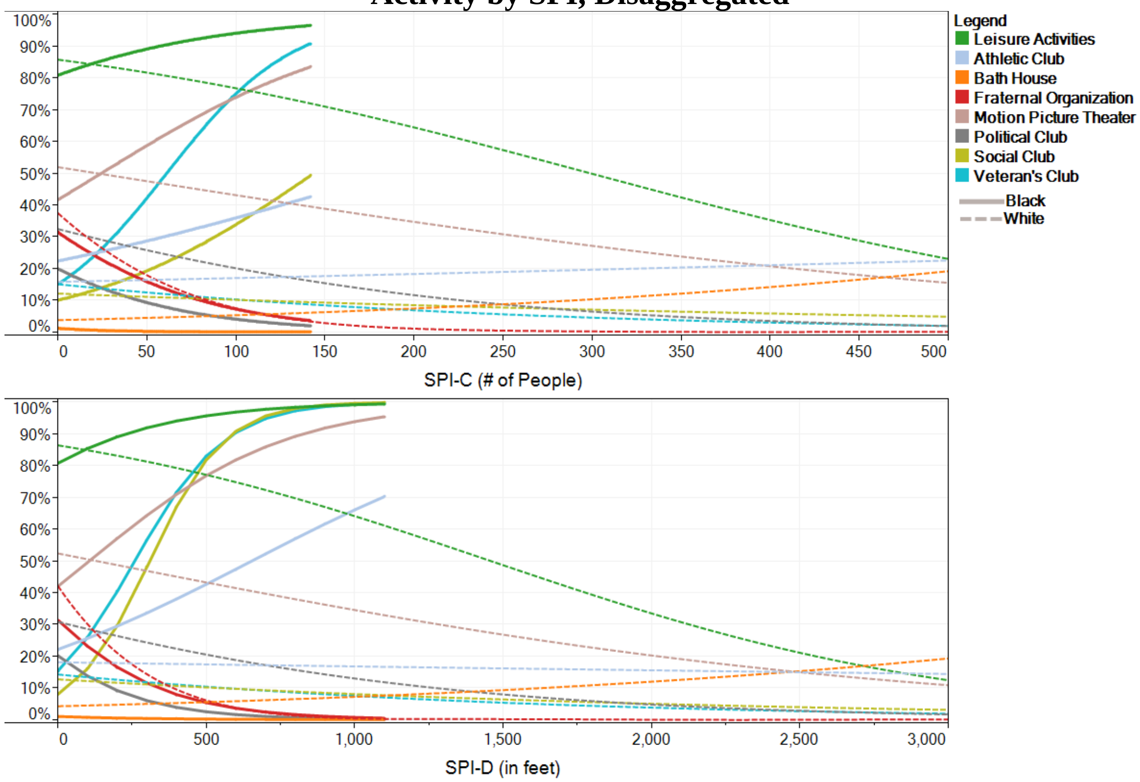
Figure 3.4: Predicted Probabilities of being within a Quarter Mile of a Public Transit Stop by SPI, Disaggregated



What stands out most from this figure is that Black household access to public transit is exactly defined by access to the trolley system; the two lines are exactly on top of each other. Another interesting finding is that which mode of transportation is more likely to be in a White neighborhood changes as isolation increases. White households with low levels of isolation are more likely to live near a trolley station, while more isolated White households are more likely to live near a bus stop. This finding may potentially point to a large racial divide in public transit use, where most of the Black households, regardless of level of isolation, used the trolley while the most racially isolated White households used the bus system. This finding is interesting because it indicates that the most racially isolated White individuals were not only isolated at their residence, but potentially on their commute to work as well. More contemporary segregation literature finds that being racially isolated in multiple activity spaces is more impactful than only in one space (M. Jones & Pebley, 2014; Wong & Shaw, 2011). This difference may be important because of how effective each mode of transportation was at moving the masses for work. While the trolley system was instrumental in helping individuals move out west, it was less than a perfect system (Cheape, 1980; Ericksen & Yancey, 1979; Warner Jr., 1987). The narrow grid-like street system lead to major traffic congestion, making it difficult for the trolleys to maneuver (Cheape, 1980; Hershberg et al., 1981). Additionally, the grid-like structure often meant that commuters would have to transfer trolleys if they wanted to change cardinal directions, and since there were no transfer fare, they would have to pay two separate fares (Cheape, 1980). The Philadelphia bus system, on the other hand, was less restricted than the trolley system (Cutler III, 1980). Not dependent on laying down tracks, the bus system allowed for more direct routes across the city than its other transportation counterparts (Cutler III, 1980). Therefore, this differential usage of transit system by race may indicate the potential differential burden households experienced while commuting to work.

Figure 3.5 shows the predicted probabilities of accessing a leisure activity by type of activity. Since many activities had low, unchanging levels of access for both White and Black households, Figure 3.5 only includes activities whose likelihood of access changed for at least one race as isolation increased.

Figure 3.5: Predicted Probabilities of being within a Quarter Mile of a Leisure Activity by SPI, Disaggregated



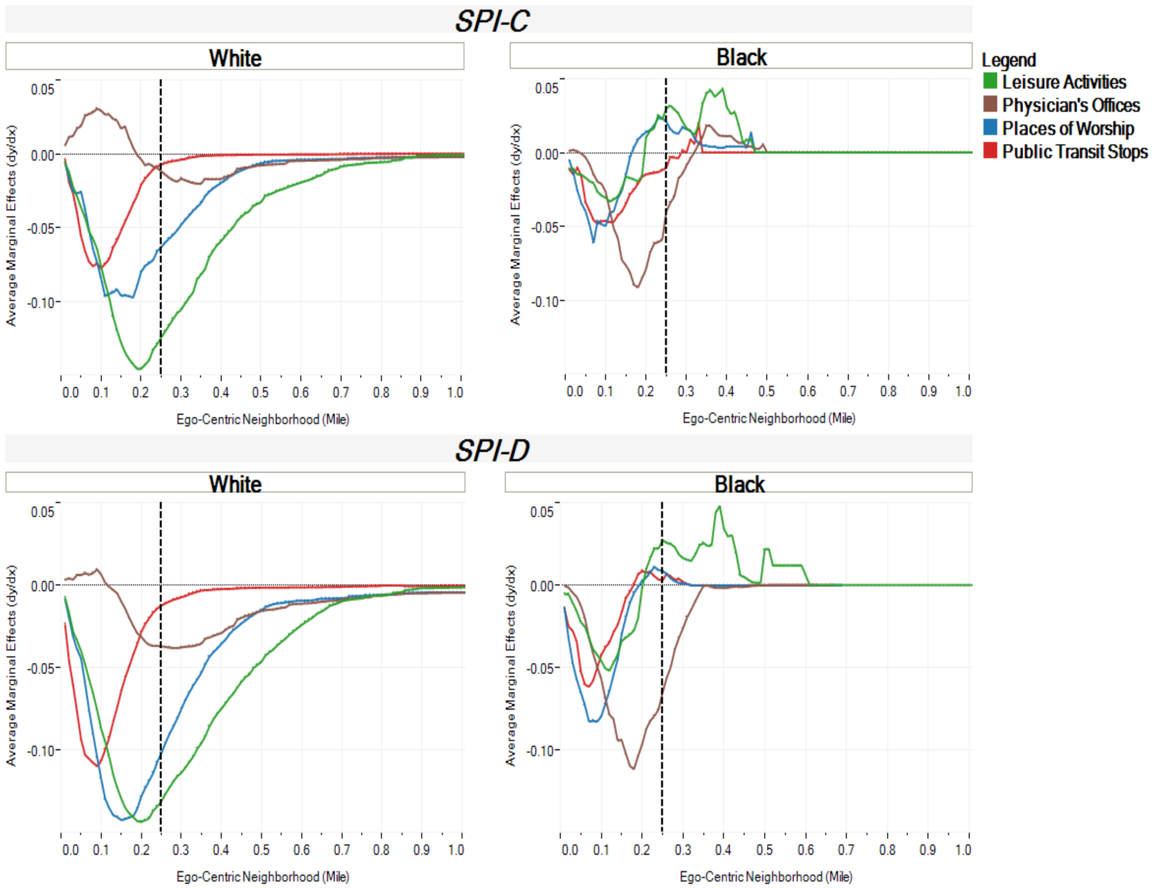
For the most part, it appears that isolated White households and isolated Black households had access to different types of leisure activities. While an increase in SPI for Black population is associated with an increase in access to leisure activities overall, it seems that this is largely due to the increased probability of being near a movie theater or veteran's club, and to a smaller extent athletic and social clubs. This pattern may be because Black households living closer to the local commercial areas compared to White households. An increase in SPI is also associated with a decreased likelihood in having access to a fraternal organization or political club for Black households. Alternatively, while an increase in SPI for White population is associated with an overall decline in accessibility to leisure activities, increased isolation is associated with an increased probability of access to the athletic club and the bath house. While both isolated White and Black households have an increased probability of living near athletic clubs, this did not necessarily mean that White and Black households attended the same athletic clubs. In 1940, West Philadelphia had 18 athletic clubs dispersed throughout the area, meaning it is highly likely that each local neighborhood had their own athletic club and potentially only catered to a single racial group. Therefore, it appears that in most cases isolated White and Black households participated in different leisure activities, apart from athletic clubs.

Together, these findings show a mixed story, where increased levels of isolation are often associated with lower likelihoods of access to resources for households, but not in all cases. However, these findings are largely dependent on the arbitrary choice of a quarter mile accurately representing the correct level of exposure and access within an ego-centric neighborhood. The next part of the analysis will investigate how sensitive the results are to one's choice of distance.

Robustness Analysis:

Figure 3.6 shows the results of what the SPI-C and SPI-D coefficients would be for access to each resource by race if the analysis was repeated using a different sized ego-centric neighborhood, with one mile as the maximum distance. The y-axis shows the average marginal effect increasing the SPI measures by one standard deviation has on the likelihood of having a resource within one's ego-centric neighborhood. A vertical line is placed at a quarter mile for reference. The patterns shown in Figure 3.6 provide insights into how changing neighborhood size can underestimate the magnitude and, in some cases, alter the overall conclusions of the relationship between isolation and access to resources.

Figure 3.6: Effect of Changing Ego-Centric Neighborhood Size on Relationship between SPI and Access to Resources



*Values standardized separately by race with mean of 0 and standard deviation of 1

First, Figure 3.6 shows that there is an empirical maximum distance where isolation no longer relates to access to resources, as shown by the coefficients approaching zero and statistical insignificance at longer distances. In general, the marginal effects are relatively small at distances greater than half a mile and are statistically insignificant at distances greater than one mile. However, the specific length at which isolation is no longer associated with access differs by resource. The marginal effect on public transit reaches insignificance around 0.35-0.40 miles while the marginal effect on leisure activities is closer to 0.6 miles for Black households and 0.70-0.80 miles for White households.

Second, Figure 3.6 shows that in certain cases, a choice of 0.25 miles may potentially underestimate the magnitude of the association if one's true neighborhood size is less than that. In most cases, the local minima and maxima of the coefficients are not at 0.25 miles. So how large are these differences? Table 3.4 shows the difference between the coefficient at 0.25 miles and the coefficient at the local maximum or minimum of the same sign.

Table 3.4: Difference between SPI Beta at a Quarter Mile and Extreme Beta

SPI-C

White						Black					
Resource	**β_1**	**Extrema Distance**	**β_2**	**Difference**	**Z-Score**	**Resource**	**β_1**	**Extrema Distance**	**β_2**	**Difference**	**Z-Score**
Leisure Activities	-0.125 (0.001)	0.20	-0.146 (0.002)	0.021	9.34	Leisure Activities	0.03 (0.004)	0.39	0.043 (0.006)	0.013	1.89
Physician's Offices	-0.012 (0.001)	0.35	-0.02 (0.001)	0.009	5.55	Physician's Offices	-0.04 (0.004)	0.18	-0.092 (0.005)	0.051	8.91
Places of Worship	-0.064 (0.001)	0.18	-0.098 (0.002)	0.034	16.75	Places of Worship	0.021 (0.003)	0.23	0.023 (0.003)	0.002	0.52
Public Transit Stops	-0.007 (0.0005)	0.10	-0.078 (0.002)	0.07	40.77	Public Transit Stops	-0.01 (0.001)	0.08	-0.048 (0.006)	0.038	6.68

SPI-D

White						Black					
Resource	**β_1**	**Extrema Distance**	**β_2**	**Difference**	**Z-Score**	**Resource**	**β_1**	**Extrema Distance**	**β_2**	**Difference**	**Z-Score**
Leisure Activities	-0.131 (0.001)	0.20	-0.144 (0.002)	0.012	5.95	Leisure Activities	0.027 (0.004)	0.39	0.048 (0.007)	0.021	2.57
Physician's Offices	-0.037 (0.001)	0.29	-0.038 (0.001)	0.001	0.65	Physician's Offices	-0.065 (0.003)	0.18	-0.112 (0.005)	0.046	8.32
Places of Worship	-0.102 (0.001)	0.15	-0.143 (0.002)	0.041	20.92	Places of Worship	0.009 (0.004)	0.23	0.011 (0.003)	0.002	0.48
Public Transit Stops	-0.012 (0.001)	0.09	-0.11 (0.002)	0.097	52.13	Public Transit Stops	0.003 (0.002)	0.2	0.008 (0.003)	0.006	1.67

Note: Each model controls for: % homeowner, % married, % US born, % in labor force, and per capita income
Robust standard error in parentheses
* p <0.05 ** p < 0.01 *** p < 0.001

For White households, the differences in magnitude for public transit stops and places of worship are quite stark, with z-scores above 15; the difference in magnitude for leisure activity is also large. For Black households, the largest difference in magnitude is access to physician's offices. Therefore, it is likely that many of the findings in this study may be a more conservative estimate of the relationship between isolation and access.

Third, Figure 3.6 shows that changing neighborhood size may affect the overall conclusions of the relationship between isolation and access to resources differs by race. In general, the overall conclusion about isolation for White households remains the same regardless of ego-centric neighborhood size. While the magnitude of the coefficients change as distance increases, the overall sign of the coefficients stay the same for to transit stops, places of worship, and leisure activities. While the coefficient related to access to physician's offices changes signs at around 0.12-0.20, depending on the measure, it is more likely that individuals are willing to travel longer than a quarter mile for health care, not less (Hine & Kamruzzaman, 2012). Therefore, it is safe to conclude that for White households, isolation is negatively associated with access to public transit, places of worship, and leisure activities, and most likely physician's offices as well. On the other hand, the conclusions about the relationship between isolation and access to resources is heavily dependent on defined neighborhood size for Black households. While isolation is negatively associated with access to resources at shorter distance, all coefficients change sign at some distances, except for access to physician's offices for SPI-D. While it is difficult to conclude the exact reason for this pattern, it most likely relates to the fact that as the ego-centric neighborhood area increases, there is a higher likelihood that those households would include the resources situated in White areas just outside the Black concentrated area. Therefore, without having a general theory about the size of a household's ego-centric neighborhoods, there is the potential risk of coming to the incorrect conclusions.

Discussion:

In their book, "American Apartheid," Massey & Denton (1993) conclude that residents living in highly segregated areas "necessarily live within a very circumscribed and limited social world," rarely travelling outside of their perceived neighborhood bounds (Massey & Denton, 1993, p. 161). While this level of isolation may produce a lack of personal connection with individuals outside of their neighborhood, the findings of this paper show that isolation may also lead to a lack of connection to social institutions and other neighborhood resources as well. Isolated White households were less likely to have a physician's office, place of worship, leisure activity, or public transit stop in their defined neighborhood than non-isolated White households. Isolated Black households were also less likely to have a physician's office in their neighborhood; however, they were more likely to have access to a place of worship or leisure activity. From these findings, we learn that racially isolated households potentially experience a very different world compared to their non-isolated counterparts.

These findings present a somewhat different view of segregation than prior literature might expect. Prior literature often displays the White population as always benefitting from segregation and the Black population, more than any other minority group, suffering from segregation (Bauman, 1974; Charles, 2003; Cloutier, 1984; Duncan & Duncan, 1957; Taeuber & Taeuber, 1965). This study finds a much less clear cut story. Extremely isolated White households have a much lower likelihood of access to these neighborhood resources than the average Black household. Additionally, isolated Black households have a higher likelihood of access to places of worship and leisure activities than non-isolated Black households. While it is not the intent of this article to compare the relative importance of a physician's office to a place of worship, it is nonetheless interesting to note that an increase in isolation does not always correspond to a reduced likelihood in access. Therefore, the investigation of how isolation relates to access to neighborhood resources does not necessarily fall along the same racial divide as access to other resources, such as jobs (Hershberg et al., 1979).

While Marcuse (1997) describe segregated areas as either Citadels, Ghettos, or Enclaves, these findings potentially indicate that segregated areas may not fit perfectly into these rigid definitions. While an area may be an economic Ghetto, lacking job accessibility or businesses, it may be a social Enclave, with many institutions providing social support for residents. By looking at an area as either completely beneficial or harmful overlooks the complexity of all the different resources that impact residents' wellbeing. Therefore, segregation literature may need to refine the ways it details segregated areas to allow for these nuances.

By utilizing individual-level data, this study delves deeper into how an individual household's level of racial isolation relates to the likelihood of access to certain neighborhood resources. By using a people-based, rather than place-based, representation, this study provides new insights into how individuals and households experienced their neighborhood. This level of detail would not be possible if not for the ability to measure individual-level isolation. Therefore, the use of the Shortest Path Isolation (SPI) index in this study is advantageous for many reasons. First, it provides a more realistic measurement of a household's level of racial isolation that might have been otherwise obscured using area-level data. By utilizing the street-networks, instead of Euclidean distance, the SPI index could provide a more detailed picture of what household's may have experienced in their daily life, which also allows for a more detailed investigation of the robustness of our assumptions on the size of individual's perceived neighborhood. Second, the SPI index allows for a more micro-level investigation of the relationship between isolation and access to resources, rather than investigate how the average level of isolation relates to the average degree of access for households residing in a general arbitrarily bounded area. The SPI index calculates the level of isolation for each individual household, meaning that this study could investigate how even the smallest differences in household racial isolation relates to each household's likelihood access to resources. Third, because each household has its own SPI value, this study performs micro-level analysis, controlling for household characteristics, without introducing the ecological fallacy and reducing the risk of multicollinearity. The ability to perform individual-level regression analysis also allows for more statistical power to detect an effect size. Given that West Philadelphia is comprised of either 41 census tracts or 5 wards, having enough statistical power to detect an effect is a legitimate concern for area-level analysis. While there are 342 enumeration districts in which non-institutionalized population live, this information is not readily available without micro-level data which begs the question why not simply study individuals anyways. Therefore, the Shortest Path Isolation index's ability to measure individual-level differences in isolation provided a lot of strength to this analysis.

Limitations:

There are a few limitations in this study that are worth noting. In every study, researchers must make certain analytical decisions and assumptions that are important to understand their potential impact on the findings.

First, proximity, access, and utilization are not the same. Just because a household lives near a resource does not mean that that household has access to this resource. Black households were often discriminated against and barred from using certain nearby resources (Schafer Jr., 2014; Warner Jr., 1987). Additionally, even if someone had access to a resource, this does not mean that that individual utilized this resource. Individual households have different tastes and behaviors, meaning they may not want to go to every church or leisure activity in their neighborhood. Additionally, the automobile was introduced in Philadelphia in 1910, meaning that it is likely that some of the upper-income households commuted to work via their own transportation (Gin & Sonstelie, 1992). Therefore, the findings of this study should be thought of as the potential options afforded to a household if they were allowed access everywhere and not necessarily the true level of usage. However, even though the analysis in this project cannot identify if specific households went to specific locations within their ego-centric neighborhood, it does not mean these findings are not relevant for understanding residential segregation and resource allocation. Households are impacted by living near resources, even if they do not use them. Local residents would have benefited from the paved roads completed by the transit company to prevent trains from breaking down (Cheape, 1980). Households would have benefitted from the social services and programs provided by church patrons, even if they were not members themselves (Cnaan et al., 2006). The overall health of the area would have benefited from physicians providing treatment and education to neighborhoods, reducing the chances of spreading infection (Acevedo-Garcia, 2000; Diez-Roux & Aiello, 2005). Additionally, households may be negatively affected by their proximity to these resources. Public transportation may have contributed to the overall pollution in the area, impacting residents' physical and mental health (Downey & Van Willigen, 2005). Certain leisure activities may have attracted criminal activity into the area (Kinney, Brantingham, Wuschke, Kirk, & Brantingham, 2008). Therefore, even if a household did not directly use the resources within their neighborhood, they were undoubtable impacted indirectly by their presence. While the question of how proximity relates to utilization is an important one, it is beyond the scope of this paper.

Second, ego-centric neighborhoods are not always isotropic, equidistant from all directions (Chaix et al., 2009). Contemporary empirical evidence finds that most individual's perceived neighborhood is not simply a circle with them in the center, but are an array of different patterns depending on a series of factors (Basta, Richmond, & Wiebe, 2010; Chaix et al., 2009; Coulton, 2012; Coulton et al., 2001). While knowledge of the local boundaries, both physical and perceived, could have potentially created more oriented neighborhoods, without specific knowledge of travel patterns, this process would be guesswork at best (Chaix et al., 2009). Therefore, while isotropic neighborhoods may lead to some measurement error, it is uncertain if oriented neighborhoods would not have as well. Additionally, even if it could be done, the marginal decrease in error may be small compared to the benefit of transitioning from Euclidean distance to network-distance (Gutiérrez & García-Palomares, 2008).

Implications for Future Research:

There are a few ways in which to advance upon this study for future research. First, just because isolated Black households had access to neighborhood resources does not mean that they had access to the same type of resources as White households. White and Black households often had access to different places of worship, leisure activities, and modes of public transit. Therefore, future research might benefit from using city directories, as well as church and social club membership lists, to link individuals to their specific neighborhood resource. Second, future research may benefit from teasing out how access to these different types of resources may differentially impact household's wellbeing. Third, while this project treated the White and Black population as homogeneous, these racial groups are in fact comprised of many important subgroups. West Philadelphia was home to many White ethnic groups including Russians (Jews), Irish, Italians, and Central/Eastern Europeans (Marsh, 1980; F. Miller, 1984). Therefore, it would be interesting to investigate how ethnic isolation relates to access to resources as well. The Black population in West Philadelphia was not homogeneous either; while some families were originally from Philadelphia, many migrated from the South decades earlier (F. Miller, 1984). These Southern Black families were different than the Northern Black families, often having larger families (Lane, 1992). Therefore, in addition to investigating White ethnic isolation, investigating Southern/Northern Black isolation may also provide new insights. Lastly, given that the neighborhood context can impact individual's wellbeing, future research might also benefit from investigating how isolation and access directly relate to individual outcomes such as employment, health, and social wellbeing. Fortunately, the complete count 1940 Census data would allow for these types analyses.

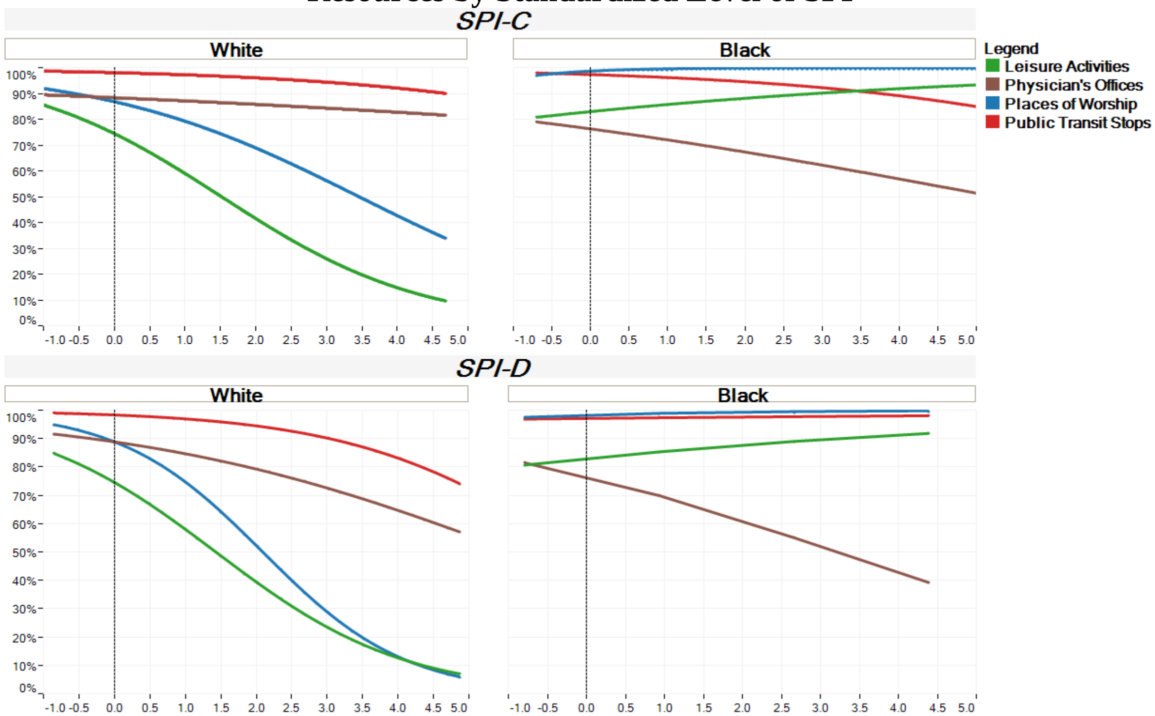
Conclusion:

By utilizing individual-level data, this study was able to show the relationship between household racial isolation and access to neighborhood resources in greater detail that might not have been as evident using an aggregate-level place-based representation. This study shows that while racial isolation is associated with differential access to neighborhood resources, in which way it relates depends on the specific neighborhood resource. Additionally, this study shows that racial isolation is not always negatively associated with access to resources for Black households. So, while many researchers are inclined to refer to any concentrated Black area as a Ghetto, in some circumstances, these areas may have also acted as an Enclave (Hershberg et al., 1979; J. R. Logan, Zhang, & Chunyu, 2015; Massey & Denton, 1993; Morgan, 1983b). While segregation literature should continue to study how residential segregation impacts job accessibility, it should also expand its focus to include other neighborhood resources that are also relevant to residents' wellbeing.

Appendix:

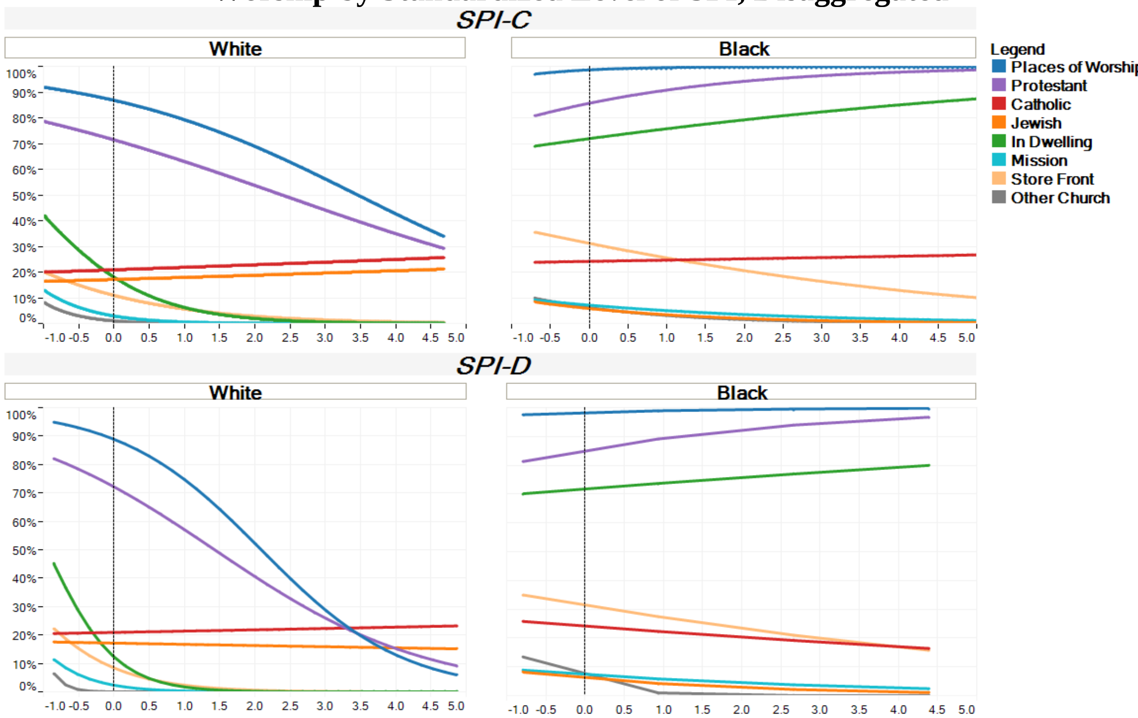
Figure 3.7 through Figure 3.10 display the same information as Figure 3.2 through Figure 3.5, but with the SPI measures standardized separately by race with a mean of 0 and a standard deviation of 1.

Figure 3.7: Predicted Probabilities of being within a Quarter Mile of Select Resources by Standardized Level of SPI



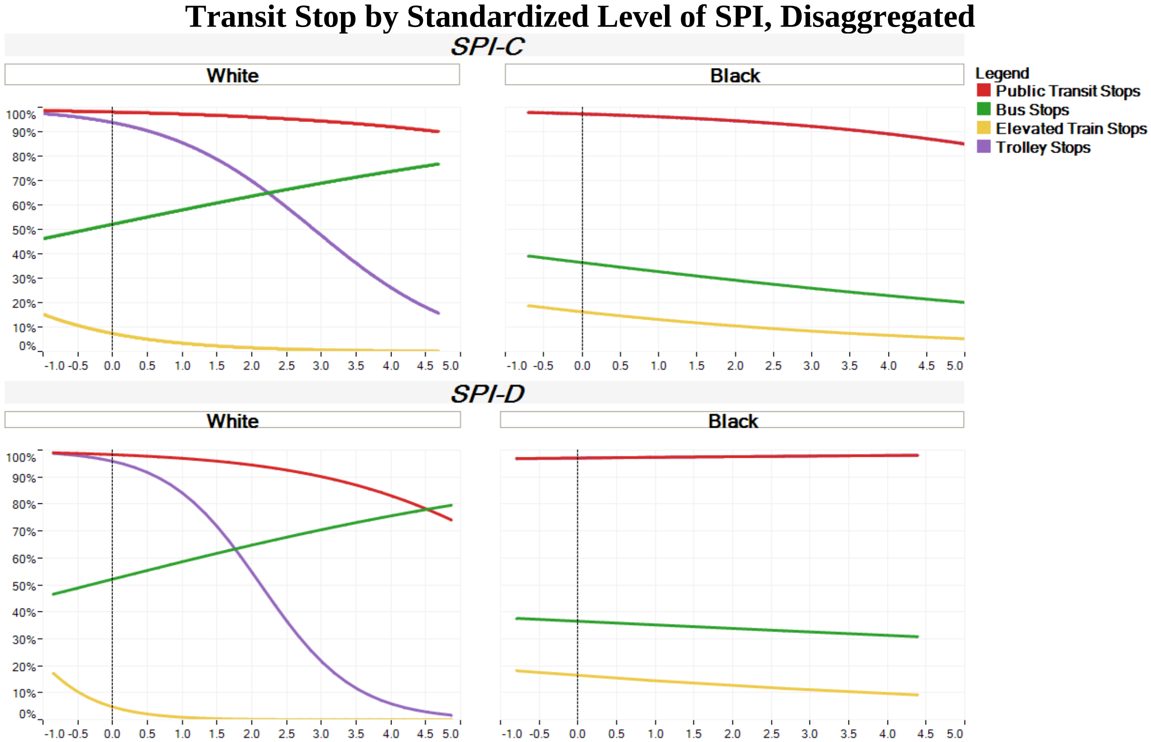
*Values standardized separately by race with mean of 0 and standard deviation of 1

Figure 3.8: Predicted Probabilities of being within a Quarter Mile of a Place of Worship by Standardized Level of SPI, Disaggregated



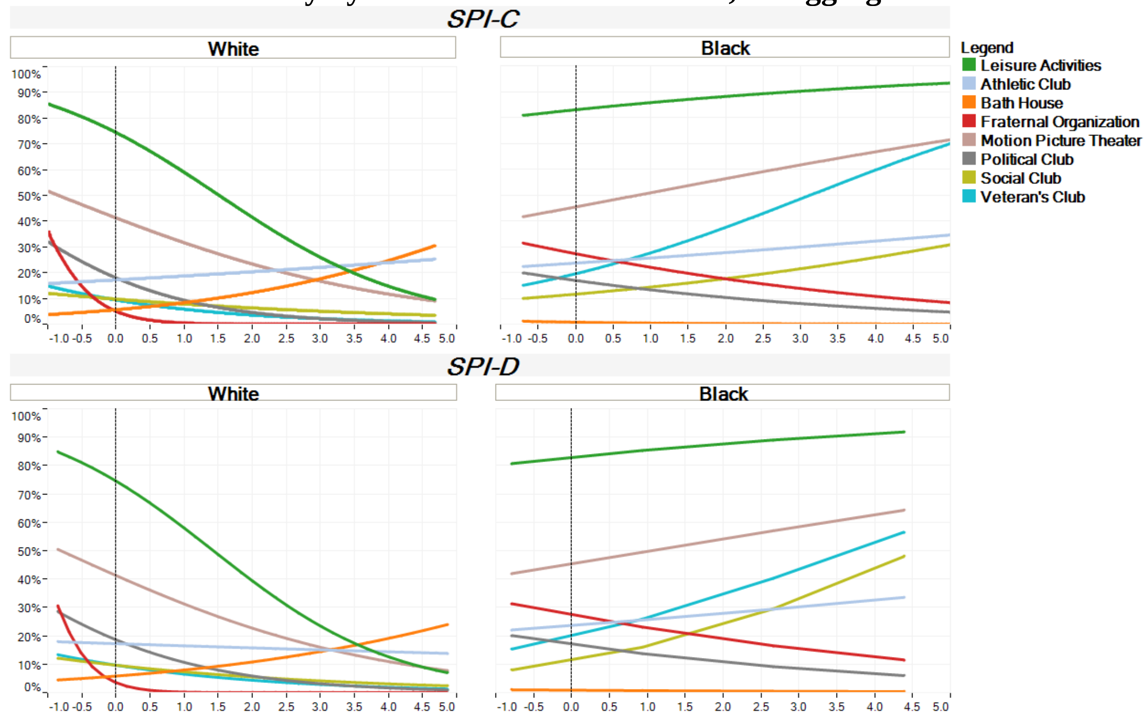
*Values standardized separately by race with mean of 0 and standard deviation of 1

Figure 3.9: Predicted Probabilities of being within a Quarter Mile of a Public Transit Stop by Standardized Level of SPI, Disaggregated



*Values standardized separately by race with mean of 0 and standard deviation of 1

Figure 3.10: Predicted Probabilities of being within a Quarter Mile of a Leisure Activity by Standardized Level of SPI, Disaggregated



*Values standardized separately by race with mean of 0 and standard deviation of 1

CHAPTER 4: CONCLUDING REMARKS

Accurate measurement is necessary for rigorous analysis. Without it, the study of segregation cannot push forward to truly understanding the proximate and distal causes and consequences of segregation. Unfortunately, the literature has relied too heavily on arbitrary areal units for analysis, forcing researchers to study segregation in a contrived way. Instead of modifying the data to fit the theory, researchers have been forced to modify the theory to match the data (D. R. James & Taeuber, 1985; Sin, 2002). For example, while researchers want to investigate how different segregation mechanisms operate at different scales, they are often forced to study segregation at the level of the census tract, assuming that it represents a "neighborhood" (Farrell, 2008; Fischer et al., 2004; Kramer & Hogue, 2009; Kwan, 2012; Massey & Hajnal, 1995; Massey et al., 2009). This process misaligns the data from the contextually relevant area of study, thereby weakening the link between the theoretical models and empirical findings (Cloutier, 1984; Kramer et al., 2010; Kwan, 2012).

This reliance on areal units has had long-lasting effects. In their seminal work "The Dimensions of Residential Segregation," Massey & Denton (1988) categorized segregation into the five dimensions commonly studied in the literature, each with its own spatial patterning and assumptions as to how segregation operates (Sundstrom, 2004). This article substantially influenced the field, initiating a new set of research and theory about racial inequality in America (Massey, 2012). Based on these criteria, Massey & Denton (1989) coined the term "hypersegregation," or areas that have high levels of segregation across multiple dimensions, which has been regarded as a "fundamental mechanism of socioeconomic stratification" (Massey, 2012, p. 42). However, as Reardon & O'Sullivan

(2004) note, this classification system arose due to arbitrary areal units; the distinction between many of the dimensions do not hold when one knows the exact location of individuals. Therefore, regrettably, a potentially sizeable amount of research has been dedicated to detailing certain spatial processes that may, in fact, be data artifacts from an erroneous classification system.

For these reasons, we as researchers need to make a greater push for individual-level geographic information while still protecting privacy and confidentiality concerns.

Unlike areal units, individual-level point data provides more freedom to the researcher to mold data to the areal structures that accurately match theory, allowing for more rigorous testing of theory. For example, if researchers want to study how local policing practices relate to segregation, then they could aggregate the information into police precincts. Alternatively, if researchers wanted to study housing discrimination, researchers could aggregate the information into real estate districts. Alternatively, they could assign each housing unit to its specific realtor to see if certain agencies had discriminatory practices. As data becomes more precise, the options will become limitless.

Additionally, as shown in this Dissertation, individual-level data can also allow for the inspection of people-based exposures. Like the process performed by the Shortest Path Isolation (SPI) Index, researchers can simulate individuals' behaviors and actions in an environment and investigate what they might experience in their daily life. While most data currently available are static, relating only to one's place of residence, this is certain to change as more geocoded and travel data becomes more readily available. With the level of detailed information, researchers can study a variety of exposures individuals may experience in their daily lives, including, but not limited to: crime events, jobs, neighborhood institutions, schools, transportation, banks, grocery stores, liquor stores, and parks.

While this Dissertation details one point-level people-based measure, it is the belief of this author that more measures need to be developed to bring in the new wave of segregation research. Not only does this type of data avoid the problems plaguing measures that rely on arbitrary areal units, but it allows for new analytical perspectives. Individual-level data allows for new forms of both aggregate-level and individual-level analysis which may lead to new insights about what is relevant to an individual's experience of segregation. In the end, by better matching data to theory, researchers will be better suited to test our theories on segregation.

Bibliography

95th Congress. An act to amend chapter 21 of title 44, United States Code, to include new provisions relating to the acceptance and use of records transferred to the custody of the Administrator of General Services (1978). Washington, D.C.: 95th Congress.

Acevedo-Garcia, D. (2000). Residential Segregation and the Epidemiology of Infectious Diseases. *Social Science and Medicine*, *51*(8), 1143–1161.

Adams, C., Bartelt, D., Elesh, D., Goldstein, I., Kleniewski, N., & Yancey, W. (1993). *Philadelphia: Neighborhoods, Division, and Conflict in a Postindustrial City*. Philadelphia: Temple University Press.

Agresti, B. F. (1980). Measuring Residential Segregation in Nineteenth-Century American Cities. *Sociological Methods & Research*, *8*(4), 389–399.

Alba, R. D., & Logan, J. R. (1992). Analyzing Locational Attainments: Constructing Individual-Level Regression Models Using Aggregate Data. *Sociological Methods & Research*, *20*(3), 367–397.

Alba, R. D., & Logan, J. R. (1993). Minority Proximity to Whites in Suburbs: An Individual-Level Analysis of Segregation. *American Journal of Sociology*, *98*(6), 1388.

Amrhein, C. G., & Flowerdew, R. (1992). The effect of data aggregation on a Poisson regression model of Canadian migration. *Environment and Planning A*, *24*(10), 1381–1391.

Apparicio, P., Martori, J. C., Pearson, A. L., Fournier, E., Apparicio, D., Fournier, É., & Apparicio, D. (2014). An Open-Source Software for Calculating Indices of Urban Residential Segregation. *Social Science Computer Review*, *32*(1), 117–128.

Atash, F. (1994). Redesigning Suburbia for Walking and Transit: Emerging Concepts. *Journal of Urban Planning and Development*, *120*(1), 48–57.

Aultman-Hall, L., Roorda, M., & Baetz, B. W. (1997). Using GIS for Evaluation of Neighborhood Pedestrian Accessibility. *Journal of Urban Planning and Development*, *123*(1), 10–17.

Austin, W. L., & Hopkins, H. L. (1940). *Instructions to Enumerators: Population and Agriculture 1940* (No. Form PA-1). Washington, D.C.

Basta, L. A., Richmond, T. S., & Wiebe, D. J. (2010). Neighborhoods, daily activities, and measuring health risks experienced in urban environments. *Social Science and Medicine*, *71*(11), 1943–1950.

Bauman, J. F. (1974). Black Slums/Black Projects: The New Deal and Negro Housing in Philadelphia. *Pennsylvania History*, *41*(3), 311–338.

Bauman, J. F. (1981). Downtown Versus Neighborhood: Focusing on Philadelphia in the Metropolitan Era, 1920-1980. *Pennsylvania History*, *48*(1), 3–20.

Bayer, P., McMillan, R., & Rueben, K. S. (2004). What Drives Racial Segregation? New Evidence Using Census Microdata. *Journal of Urban Economics*, *56*, 514–535.

Bell, W. (1954). A Probability Model for the Measurement of Ecological Segregation. *Social Forces*, *32*(4), 357–364.

Blalock, H. M. (1971). Aggregation and Measurement Error. *Social Forces*, *50*(2), 151–165.

Blanchard, T. C. (2007). Conservative Protestant Congregations and Racial Residential Segregation: Evaluating the Closed Community Thesis in Metropolitan and Nonmetropolitan Counties. *American Sociological Review*, *72*(3), 416–433.

Boyce, R. R., & Clark, W. A. V. (1964). The Concept of Shape in Geography. *Geographical Review*, *54*(4), 561–572.

Boyd, R. L. (2001). Black Enterprise in the Retail Trade during the Early Twentieth Century. *Sociological Focus*, *34*(3), 241–250.

Boyer, H. (1896). *North of Market Street: Being the Adventures of a New York Woman in Philadelphia*. Philadelphia: Avil Printing Company.

Brassel, K. E., & Reif, D. (1979). A Procedure to Generate Thiessen Polygons. *Geographical Analysis*, *11*(3), 289–303.

Briant, A., Combes, P.-P., & Lafourcade, M. (2010). Dots to boxes: Do the size and shape of spatial units jeopardize economic geography estimations? *Journal of Urban Economics*, *67*(3), 287–302.

Brown, L. A., & Chung, S. Y. (2006). Spatial segregation, segregation indices and the geographical perspective. *Population, Space and Place*, *12*(2), 125–143.

Brown, Q., Milam, A. J., Bowie, J. V., Ialongo, N. S., Gaskin, D. J., & Furr-Holden, D. (2016). The Moderating Role of Gender in the Relationship Between Tobacco Outlet Exposure and Tobacco Use Among African American Young Adults. *Prevention Science*, *17*(3), 338–346.

Burgess, E. W. (1928). Residential Segregation in American Cities. *The Annals of the American Academy*, *140*(1), 105–115.

Butkiewicz, T., Meentemeyer, R. K., Shoemaker, D. A., Chang, R., Wartell, Z., & Ribarsky, W. (2010). Alleviating the Modifiable Areal Unit Problem within Probe-Based Geospatial Analyses. *Computer Graphics Forum*, *29*(3), 923–932.

Carrington, W. J., & Troske, K. R. (1997). On Measuring Segregation in Samples with Small Units. *Journal of Business & Economic Statistics*, *15*(4), 402–409.

Chaix, B., Merlo, J., Evans, D., Leal, C., & Havard, S. (2009). Neighbourhoods in eco-epidemiologic research: delimiting personal exposure areas. A response to Riva, Gauvin, Apparicio and Brodeur. *Social Science & Medicine*, *69*(9), 1306–10.

Charles, C. Z. (2003). The Dynamics of Racial Residential Segregation. *Annual Review of Sociology*, *29*(1), 167–207.

Charlton, M. (2008). Modifiable Areal Unit Problem (MAUP). In *Encyclopedia of*

Geographic Information Science (pp. 289–290). Sage Publications, Inc.

Cheape, C. W. (1980). *Moving the Masses: Urban Public Transit in New York, Boston, and Philadelphia, 1880-1912* (1st ed.). Cambridge, MA: Harvard University Press.

Clark, W. A. V., & Avery, K. L. (1976). The Effects of Data Aggregation in Statistical Analysis. *Geographical Analysis*, *8*(October), 428–438.

Cloutier, N. R. (1984). The Measurement and Modeling of Segregation: A Survey of Recent Empirical Research. *Regional Science Perspectives*, *14*(2), 15–32.

Cnaan, R. A., Boddie, S. C., McGrew, C. C., & Kang, J. (2006). *The Other Philadelphia Story: How Local Congregations Support Quality of Life in Urban America*. Philadelphia, PA: University of Pennsylvania Press.

Cockings, S., & Martin, D. (2005). Zone design for environment and health studies using pre-aggregated data. *Social Science & Medicine*, *60*(12), 2729–2742.

Cohn, M. J., & Jackman, S. P. (2011). A Comparison of Aspatial and Spatial Measures of Segregation. *Transactions in GIS, 15*(S1), 47–66.

Collins, H. (1970). Store Front Churches. *Negro American Literature Forum*, *4*(2), 64–68.

Cortese, C. F., Falk, R. F., & Cohen, J. K. (1976). Further Considerations on the Methodological Analysis of Segregation Indices. *American Sociological Review*, *41*(4), 630–637.

Coulton, C. J. (2012). Defining Neighborhoods for Research and Policy. *Cityscape*, *14*(2), 231–236.

Coulton, C. J., Korbin, J., Chan, T., & Su, M. (2001). Mapping Residents' Perceptions of Neighborhood Boundaries: A Methodological Note. *American Journal of Community Psychology*, *29*(2), 371–383.

Culhane, J. F., & Elo, I. T. (2005). Neighborhood context and reproductive health. *American Journal of Obstetrics and Gynecology*, *192*(5 Suppl), S22-9.

Cutler, W. W., & Gillette Jr., H. (Eds.). (1980). *The Divided Metropolis: Social and Spatial Dimensions of Philadelphia, 1800-1975*. Westport, Connecticut: Greenwood Press.

Cutler III, W. W. (1980). The Persistent Dualism: Centralization and Decentralization in Philadelphia, 1854-1975. In W. W. Cutler & H. Gillette Jr. (Eds.), *The Divided Metropolis: Social and Spatial Dimensions of Philadelphia, 1800-1975* (pp. 249–284). Westport, Connecticut: Greenwood Press.

Dark, S. J., & Bram, D. (2007). The modifiable areal unit problem (MAUP) in physical geography. *Progress in Physical Geography*, *31*(5), 471–479.

Diez-Roux, A. V. (2001). Investigating Neighborhood and Area Effects on Health. *American Journal of Public Health*, *91*(11), 1783–9.

Diez-Roux, A. V. (2004). Estimating Neighborhood Health Effects: the Challenges of Causal Inference in a Complex World. *Social Science & Medicine, 58*(10), 1953–60.

Diez-Roux, A. V., & Aiello, A. E. (2005). Multilevel Analysis of Infectious Diseases. *The*

Journal of Infectious Diseases, *191*(S1), S25–S33.

Diez-Roux, A. V., & Mair, C. (2010). Neighborhoods and health. *Annals of the New York Academy of Sciences*, *1186*(1), 125–45.

Downey, L., & Van Willigen, M. (2005). Environmental Stressors: The Mental Health Impacts of Living near Industrial Activity. *Journal of Health and Social Behavior*, *46*(3), 289–305.

Duke, J. (2012). ProQuest Backs 1940 U.S. Census Community Project. *Advanced Technology Libraries*, *41*(6), 2.

Dumedah, G., Schuurman, N., & Yang, W. (2008). Minimizing effects of scale distortion for spatially grouped census data using rough sets. *Journal of Geographical Systems*, *10*(1), 47–69.

Duncan, O. D., & Duncan, B. (1955). A Methodological Analysis of Segregation Indexes. *American Sociological Review*, *20*(2), 210–217.

Duncan, O. D., & Duncan, B. (1957). *The Negro Population of Chicago: A Study of Residential Succession*. Chicago: University of Chicago Press.

Egan, K. L., Anderton, D. L., & Weber, E. (1998). Relative Spatial Concentration among Minorities: Addressing Errors in Measurement. *Social Forces*, *76*(3), 1115–1121.

Ellis, M., Holloway, S. R., Wright, R., & Fowler, C. S. (2012). Agents of Change: Mixed-Race Households and the Dynamics of Neighborhood Segregation in the United States. *Annals of the Association of American Geographers*, *102*(3), 549–570.

Ericksen, E. P., & Yancey, W. L. (1979). Work and Residence in Industrial Philadelphia. *Journal of Urban History*, *5*(2), 147–182.

Farley, J. E. (1984). P* Segregation Indices: What Can They Tell Us about Housing Segregation in 1980? *Urban Studies*, *21*(3), 331–336.

Farrell, C. R. (2008). Bifurcation, Fragmentation or Integration? The Racial and Geographical Structure of US Metropolitan Segregation, 1990-2000. *Urban Studies*, *45*(3), 467–499.

Firebaugh, G. (1978). A Rule for Inferring Individual-Level Relationships from Aggregate Data. *American Sociological Review*, *43*(4), 557–572.

Fischer, C. S., Stockmayer, G., Stiles, J., & Hout, M. (2004). Distinguishing the Geographic Levels and Social Dimensions of U.S. Metropolitan Segregation, 1960-2000. *Demography*, *41*(1), 37–59.

Flowerdew, R. (2011). How serious is the Modifiable Areal Unit Problem for analysis of English census data? *Population Trends*, *145*(1), 106–118.

Flowerdew, R., Manley, D. J., & Sabel, C. E. (2008). Neighbourhood effects on health: Does it matter where you draw the boundaries? *Social Science & Medicine*, *66*(6), 1241–1255.

Fogarty, E. A. (2010). Modifiable Areal Unit Problem. In *Encyclopedia of Geography* (pp. 1935–1937). Sage Publications, Inc.

Fotheringham, A. S. (1989). Scale-independent spatial analysis. In M. F. Goodchild & S. Gopal (Eds.), *Accuracy of Spatial Databases* (pp. 221–228). London: Taylor & Francis.

Fotheringham, A. S. (1998). Trends in quantitative methods II: stressing the computational. *Progress in Human Geography*, *22*(2), 283–292.

Fotheringham, A. S., & Wong, D. W. S. (1991). The modifiable areal unit problem in multivariate statistical analysis. *Environment and Planning A*, *23*(7), 1025–1044.

Gaskin, D. J., Dinwiddie, G. Y., Chan, K. S., & McCleary, R. (2012a). Residential Segregation and Disparities in Health Care Services Utilization. *Medical Care Research and Review*, *69*(2), 158–175.

Gaskin, D. J., Dinwiddie, G. Y., Chan, K. S., & McCleary, R. R. (2012b). Residential Segregation and the Availability of Primary Care Physicians. *Health Services Research*, *47*(6), 2353–2376.

Gehlke, C. E., & Biehl, K. (1934). Certain Effects of Grouping Upon the Size of the Correlation Coefficient in Census Tract Material. *Journal of the American Statistical Association*, *29*(185A), 169–170.

Gin, A., & Sonstelie, J. (1992). The Streetcar and Residential Location in Nineteenth Century Philadelphia. *Journal of Urban Economics*, *32*, 92–107.

Glazer, N., & Moynihan, D. P. (1963). *Beyond the Melting Pot: The Negroes, Puerto Ricans, Jews, Italians, and Irish of New York City*. Cambridge, MA: The Massachusetts Institute of Technology Press.

Goodchild, M. F. (2011). Scale in GIS: An overview. *Geomorphology*, *130*(1), 5–9.

Grannis, R. (1998). The Importance of Trivial Streets: Residential Streets and Residential Segregation. *American Journal of Sociology*, *103*(6), 1530–1564.

Grannis, R. (2005). T-Communities: Pedestrian Street Networks and Residential Segregation in Chicago, Los Angeles, and New York. *City & Community*, *4*(3), 295–321.

Greenberg, S. W. (1980). The Relationship between Work and Residence in an Industrializing City: Philadelphia, 1880. In W. W. Cutler & H. Gillette Jr. (Eds.), *The Divided Metropolis: Social and Spatial Dimensions of Philadelphia, 1800-1975* (pp. 141–168). Westport, Connecticut: Greenwood Press.

Grigoryeva, A., & Ruef, M. (2015). The Historical Demography of Racial Segregation. *American Sociological Review*, *80*(4), 1–29.

Guo, J. Y., & Bhat, C. R. (2004). Modifiable Areal Units: Problem or Perception in Modeling of Residential Location Choice? *Transportation Research Record: Journal of the Transportation Research Board*, *1898*(1), 138–147.

Gutiérrez, J., & García-Palomares, J. C. (2008). Distance-measure impacts on the calculation of transport service areas using GIS. *Environment and Planning B: Planning and Design*, *35*(3), 480–503.

Harrison, I. E. (1966). The Storefront Church as a Revitalization Movement. *Review of Religious Research*, *7*(3), 160–163.

Harrison, R. (2001). Segregation Indices. In *International Encyclopedia of Social and Behavioral Sciences* (1st ed., pp. 13791–13795). Elsevier Science Ltd.

Haynes, R., Daras, K., Reading, R., & Jones, A. (2007). Modifiable neighbourhood units, zone design and residents' perceptions. *Health and Place*, *13*(4), 812–825.

Hershberg, T., Burstein, A. N., Ericksen, E. P., Greenberg, S. W., & Yancey, W. L. (1979). A Tale of Three Cities: Blacks and Immigrants in Philadelphia: 1850-1880, 1930 and 1970. *The Annals of the American Academy of Political and Social Science*, *441*(1), 55–81.

Hershberg, T., Cox, H. E., Light Jr., D., & Greenfield, R. R. (1981). The "Journey-to-Work": An Empirical Investigation of Work, Residence and Transportation, Philadelphia, 1850 and 1880. In T. Hershberg (Ed.), *Philadelphia: Work, Space, Family, and Group Experience in the Nineteenth Century: Essays Toward an Interdisiplinary History of the City* (pp. 128–173). New York: Oxford University Press.

Hine, J., & Kamruzzaman, M. (2012). Health & Place Journeys to health services in Great Britain : An analysis of changing travel. *Health & Place*, *18*(2), 274–285.

Hipp, J. R. (2012). Segregation Through the Lens of Housing Unit Transition: What Roles Do the Prior Residents, the Local Micro-Neighborhood, and the Broader Neighborhood Play. *Demography*, *49*(4), 1285–1306.

Hipp, J. R., & Perrin, A. J. (2009). The Simultaneous Effect of Social Distance and Physical Distance on the Formation of Neighborhood Ties. *City and Community*, *8*(1), 5–25.

Hoehner, C. M., Brennan Ramirez, L. K., Elliott, M. B., Handy, S. L., & Brownson, R. C. (2005). Perceived and Objective Environmental Measures and Physical Activity Among Urban Adults. *American Journal of Preventive Medicine*, *28*(2S2), 105–116.

Holt, D. T., Steel, D. G., & Tranmer, M. (1996). Area homogeneity and the Modifiable Areal Unit Problem. *Geographical Systems*, *3*(2–3), 181–200.

Holt, D. T., Steel, D. G., Tranmer, M., & Wrigley, N. (1996). Aggregation and Ecological Effects in Geographically Based Data. *Geographical Analysis*, *28*(3), 244–261.

Hong, S.-Y., O'Sullivan, D., & Sadahiro, Y. (2014). Implementing Spatial Segregation Measures in R. *PLoS ONE*, *9*(11), e113767.

Horner, M. W., & Murray, A. T. (2002). Excess Commuting and the Modifiable Areal Unit Problem. *Urban Studies*, *39*(1), 131–139.

Iceland, J. (2004). Beyond Black and White: Metropolitan Residential Segregation in Multi-Ethnic America. *Social Science Research*, *33*(2), 248–271.

IPUMS-USA. (2014). IPUMS Complete Count Data. Retrieved January 1, 2015, from https://usa.ipums.org/usa/complete_count.shtml

Jahn, J., Schmid, C. F., & Schrag, C. (1947). The Measurement of Ecological Segregation.

American Sociological Review, *12*(3), 293–303.

Jakubs, J. F. (1981). A Distance-Based Segregation Index. *Socio-Economic Planning Sciences*, *15*(3), 129–136.

James, D. R., & Taeuber, K. E. (1985). Measures of Segregation. *Sociological Methodology*, *15*(1), 1–32.

James, F. J. (1986). A New Generalized "Exposure-Based" Segregation Index: Demonstration in Denver and Houston. *Sociological Methods & Research*, *14*(3), 301–316.

Johnston, R., Poulsen, M., & Forrest, J. (2010). Moving On from Indices, Refocusing on Mix: On Measuring and Understanding Ethnic Patterns of Residential Segregation. *Journal of Ethnic and Migration Studies*, *36*(4), 697–706.

Jones, K., Johnston, R., Manley, D. J., Owen, D., & Charlton, C. (2015). Ethnic Residential Segregation: A Multilevel, Multigroup, Multiscale Approach Exemplified by London in 2011. *Demography*, *52*(6), 1995–2019.

Jones, M., & Pebley, A. R. (2014). Redefining Neighborhoods Using Common Destinations: Social Characteristics of Activity Spaces and Home Census Tracts Compared. *Demography*, *51*(3), 727–52.

Kaplan, D. H., & Woodhouse, K. (2005). Research in Ethnic Segregation II: Measurements, Categories and Meanings. *Urban Geography*, *26*(8), 737–745.

Kawachi, I., & Berkman, L. F. (Eds.). (2003). *Neighborhoods and Health*. New York: Oxford University Press.

Kennedy, H. (Ed.). (2000). *Dictionary of GIS Terminology*. Redlands, CA: ESRI Press.

Kinney, J. B., Brantingham, P. L., Wuschke, K., Kirk, M. G., & Brantingham, P. J. (2008). Crime Attractors, Generators and Detractors: Land Use and Urban Crime Opportunities. *Built Environment*, *34*(1), 62–74.

Klein, R. (1988). Abstract Voronoi Diagrams and their Applications (Extended Abstract). *Computational Geometry and Its Applications*, *333*(1), 148–157.

Kostarelos, F. (1995). The Historical and Social Context: Charisma and Black Liberation. In *Feeling the Spirit: Faith and Hope in an Evangelical Black Storefront Church* (pp. 7–12). Columbia, SC: University of South Carolina Press.

Kramer, M. R., Cooper, H. L., Drews-Botsch, C. D., Waller, L. A., & Hogue, C. R. (2010). Do measures matter? Comparing surface-density-derived and census-tract-derived measures of racial residential segregation. *International Journal of Health Geographics*, *9*(29).

Kramer, M. R., & Hogue, C. R. (2009). Is Segregation Bad for Your Health? *Epidemiologic Reviews*, *31*(1), 178–194.

Krupka, D. J. (2007). Are Big Cities More Segregated? Neighbourhood Scale and the Measurement of Segregation. *Urban Studies*, *44*(1), 187–197.

Kwan, M.-P. (2009). From place-based to people-based exposure measures. *Social Science*

& Medicine, *69*(9), 1311–1313.

Kwan, M.-P. (2012). The Uncertain Geographic Context Problem. *Annals of the Association of American Geographers*, *102*(5), 958–968.

Lane, R. (1992). Black Philadelphia, then and now. *The Public Interest*, *108*(1), 35–52.

Larsen, J., El-Geneidy, A., & Yasmin, F. (2010). Beyond the Quarter Mile: Examining Travel Distances by Walking and Cycling, Montréal, Canada. *Canadian Journal of Urban Research*, *19*(1 Supp), 70–88.

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.

Lee, M. R., & Ousey, G. C. (2005). Institutional Access, Residential Segregation, and Urban Black Homicide. *Sociological Inquiry*, *75*(1), 31–54.

Lichter, D. T., Parisi, D., Grice, S. M., & Taquino, M. C. (2007). National Estimates of Racial Segregation in Rural and Small-Town America. *Demography*, *44*(3), 563–581.

Lieberson, S. (1963). *Ethnic Patterns in American Cities: A Comparative Study Using Data from Ten Urban Centers*. New York: Free Press.

Lieberson, S. (1980). *A Piece of the Pie: Blacks and White Immigrants Since 1880*. Berkeley, CA: University of California Press.

Lieberson, S., & Carter, D. K. (1982). Temporal Changes and Urban Differences in Residential Segregation: A Reconsideration. *American Journal of Sociology*, *88*(2), 296–310.

Logan, J. R., & Bellman, B. (2016). Before The Philadelphia Negro: Residential Segregation in a Nineteenth-Century Northern City. *Social Science History*, *40*(4), 683–706.

Logan, J. R., Zhang, W., & Chunyu, M. D. (2015). Emergent Ghettos: Black Neighborhoods in New York and Chicago, 1880-1940. *American Journal of Sociology*, *120*(4), 1055–1094.

Logan, T. D., & Parman, J. M. (2017). The National Rise in Residential Segregation. *The Journal of Economic History*, *77*(1), 127–170.

Macintyre, S., Ellaway, A., & Cummins, S. (2002). Place effects on health: how can we conceptualise, operationalise and measure them? *Social Science & Medicine*, *55*(1), 125–139.

Maly, M. T. (2000). The Neighborhood Diversity Index: A Complementary Measure of Racial Residential Settlement. *Journal of Urban Affairs*, *22*(1), 37–47.

Manley, D. J., Flowerdew, R., & Steel, D. G. (2006). Scales, levels and processes: Studying spatial patterns of British census variables. *Computers, Environment and Urban Systems*, *30*(2), 143–160.

Marcuse, P. (1997). The Enclave, the Citadel, and the Ghetto: What Has Changed in the Post-Fordist U.S. City. *Urban Affairs Review*, *33*(2), 228–264.

Marsh, M. S. (1973). Suburbanization and the Search for Community: Residential Decentralization in Philadelphia, 1880-1900. *Pennsylvania History*, *44*(2), 29–44.

Marsh, M. S. (1980). The Impact of the Market Street "El" on Northern West Philadelphia: Environmental Change and Social Transformation, 1900-1930. In W. W. Cutler & H. Gillette Jr. (Eds.), *The Divided Metropolis: Social and Spatial Dimensions of Philadelphia, 1800-1975* (pp. 169–192). Westport, Connecticut: Greenwood Press.

Martin, D. (1998). Optimizing census geography: the separation of collection and output geographies. *International Journal of Geographical Information Science*, *12*(7), 673–685.

Massey, D. S. (2012). Reflections on the Dimensions of Segregation. *Social Forces*, *91*(1), 39–43.

Massey, D. S., & Denton, N. A. (1988). The Dimensions of Residential Segregation. *Social Forces*, *67*(2), 281–315.

Massey, D. S., & Denton, N. A. (1989). Hypersegregation in U.S. Metropolitan Areas: Black and Hispanic Segregation along Five Dimensions. *Demography*, *26*(3), 373–391.

Massey, D. S., & Denton, N. A. (1993). *American Apartheid: Segregation and the Making of the Underclass*. Cambridge, MA: Harvard University Press.

Massey, D. S., & Denton, N. A. (1998). The Elusive Quest for the Perfect Index of Concentration: Reply to Egan, Anderton, and Weber. *Social Forces*, *76*(3), 1123–1132.

Massey, D. S., & Hajnal, Z. L. (1995). The Changing Geographic Structure of Black-White Segregation in the United States. *Social Science Quarterly*, *76*(3), 527–542.

Massey, D. S., Rothwell, J., & Domina, T. (2009). The Changing Bases of Segregation in the United States. *The Annals of the American Academy of Political and Social Science*, *626*(1), 74–90.

Massey, D. S., White, M. J., & Phua, V.-C. (1996). The Dimensions of Segregation Revisited. *Sociological Methods & Research*, *25*(2), 172–206.

McCormack, G. R., Giles-Corti, B., & Bulsara, M. (2008). The relationship between destination proximity, destination mix and physical activity behaviors. *Preventive Medicine*, *46*(1), 33–40.

Merino, S. M. (2011). Neighbors Like Me? Religious Affiliation and Neighborhood Racial Preferences among Non-Hispanic Whites. *Religions*, *2*, 165–183.

Miller, F. (1984). The Black Migration to Philadelphia: A 1924 Profile. *The Pennsylvania Magazine of History and Biography*, *108*(3), 315–350.

Miller, R., & Siry, J. (1980). The Emerging Suburb: West Philadelphia, 1850-1880. *Pennsylvania History*, *47*(2), 99–145.

Millward, H., Spinney, J., & Scott, D. (2013). Active-transport walking behavior: destinations, durations, distances. *Journal of Transport Geography*, *28*, 101–110.

Mitra, S. (1984). On "The Measurement of Spatial Segregation." *American Journal of Sociology*, *90*(1), 187–189.

Morgan, B. S. (1982). The Properties of a Distance-Based Segregation Index. *Socio-Economic Planning Sciences*, *16*(4), 167–171.

Morgan, B. S. (1983a). A Distance-Decay Interaction Index to Measure Residential Segregation. *Area*, *15*(3), 211–217.

Morgan, B. S. (1983b). An Alternate Approach to the Development of a Distance-Based Measure of Racial Segregation. *American Journal of Sociology*, *88*(6), 1237–1249.

Morrill, R. L. (1991). On the Measure of Geographic Segregation. *Geography Research Forum*, *11*(1), 25–36.

Murdock, S. H., Hwang, S.-S., & Hoque, N. (1994). Nonmetropolitan Residential Segregation Revisited. *Rural Sociology*, *59*(2), 236–254.

Nagel, C. L., Carlson, N. E., Bosworth, M., & Michael, Y. L. (2008). The Relation between Neighborhood Built Environment and Walking Activity among Older Adults. *American Journal of Epidemiology*, *168*(4), 461–468.

O'Neill, W. A., Ramsey, R. D., & Chou, J. (1992). Analysis of Transit Service Areas Using Geographic Information Systems. *Transportation Research Record: Journal of the Transportation Research Board*, *1364*, 131–138.

Oka, M., & Wong, D. W. S. (2015). Spatializing Segregation Measures: An Approach to Better Depict Social Relationships. *Cityscape: A Journal of Policy Development and Research*, *17*(1), 97–113.

Openshaw, S. (1984). Ecological fallacies and the analysis of areal census data. *Environment & Planning A*, *16*(1), 17–31.

Openshaw, S. (1989). *Modifiable Areal Unit Problem*. Geo Books.

Openshaw, S. (1996). Developing GIS relevant zone based spatial analysis methods. In P. A. Longley & M. Batty (Eds.), *Spatial Analysis: Modelling in a GIS Environment* (pp. 55–78). Cambridge: GeoInformation International.

Openshaw, S., & Rao, L. (1995). Algorithms for reengineering 1991 Census geography. *Environment and Planning A, 27*(3), 425–446.

Openshaw, S., & Taylor, P. J. (1979). A million or so correlation coefficients: three experiments on the modifiable areal unit problem. In N. Wrigley (Ed.), *Statistical Applications in the Spatial Sciences* (pp. 127–144). London: Pion Ltd.

Östh, J., Clark, W. A. V., & Malmberg, B. (2015). Measuring the Scale of Segregation Using k-Nearest Neighbor Aggregates. *Geographical Analysis*, *47*(1), 34–49.

Páez, A., & Scott, D. M. (2004). Spatial statistics for urban analysis: A review of techniques with examples. *GeoJournal*, *61*, 53–67.

Park, R. E., Burgess, E. W., Mckenzie, R. D., & Wirth, L. (1925). *The City: Suggestions for Investigation of Human Behavior in the Urban Environment*. Chicago: University of Chicago Press.

Peach, C. (2009). Slippery Segregation: Discovering or Manufacturing Ghettos? *Journal of Ethnic and Migration Studies*, *35*(9), 1381–1395.

Philadelphia City Planning Commission. (1956). *Land Use in Philadelphia, 1944-1954*. Philadelphia, PA.

Philadelphia Rapid Transit Company. (1932). Surface High Speed and Bus Routes. Philadelphia, PA: Philadelphia Rapid Transit Company.

Philadelphia Transportation Company. (1944). Street Car, Bus and Subway-Elevated Lines. Philadelphia, PA: Philadelphia Transportation Company.

Pickett, K. E., & Pearl, M. (2001). Multilevel analyses of neighbourhood socioeconomic context and health outcomes: a critical review. *Journal of Epidemiology and Community Health*, *55*(2), 111–22.

Pierson, D. (1994). *An Ethnic History of West Philadelphia, 1870-1980: A Research Tool for Demographic Studies*. Philadelphia: Balch Institute for Ethnic Studies.

Pietrzak, M. B. (2014a). Redefining the Modifiable Areal Unit Problem Within Spatial Econometrics, the Case of the Aggregation Problem. *Equilibrium. Quarterly Journal of Economics and Economic Policy*, *9*(3), 131–151.

Pietrzak, M. B. (2014b). Redefining The Modifiable Areal Unit Problem Within Spatial Econometrics, The Case of the Scale Problem. *Equilibrium. Quarterly Journal of Economics and Economic Policy*, *9*(2), 111–132.

Poulsen, M., Johnston, R., & Forrest, J. (2001). Intraurban ethnic enclaves: introducing a knowledge-based classification method. *Environment and Planning A*, *33*(1978), 2071–2082.

Poulsen, M., Johnston, R., & Forrest, J. (2002). Plural Cities and Ethnic Enclaves: Introducing a Measurement Procedure for Comparative Study. *International Journal of Urban and Regional Research*, *26*(2), 229–243.

Proudfoot, M. J. (1937). *Intra-City Business Census Statistics for Philadelphia, Pennsylvania*. Washington, D.C.

Reardon, S. F., Farrell, C. R., Matthews, S. A., O'Sullivan, 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., O'Sullivan, D., Lee, B. A., Firebaugh, G., Farrell, C. R., & Bischoff, K. (2008). The Geographic Scale of Metropolitan Racial Segregation. *Demography*, *45*(3), 489–514.

Reardon, S. F., & O'Sullivan, D. (2004). Measures of Spatial Segregation. *Sociological Methodology*, *34*(1), 121–162.

Roberts, J. P. (1980). Railroads and the Downtown: Philadelphia 1830-1900. In W. W. Cutler & H. Gillette Jr. (Eds.), *The Divided Metropolis: Social and Spatial*

Dimensions of Philadelphia, 1800-1975 (pp. 27–55). Westport, Connecticut: Greenwood Press.

Robinson, W. S. (1950). Ecological Correlations and the Behavior of Individuals. *American Sociological Review*, *15*(3), 351–357.

Ruggles, S., Genadek, K., Goeken, R., Grover, J., & Sobek, M. (2015). Integrated Public Use Microdata Series: Version 6.0 [Machine-readable database]. Minneapolis, MN: University of Minnesota. http://doi.org/10.18128/D010.V6.0

Savery, M. (1980). Instability and Uniformity: Residential Patterns in Two Philadelphia Neighborhoods, 1880-1970. In W. W. Cutler & H. Gillette Jr. (Eds.), *The Divided Metropolis: Social and Spatial Dimensions of Philadelphia, 1800-1975* (pp. 193–226). Westport, Connecticut: Greenwood Press.

Schafer Jr., J. A. (2014). *Critical Issues in Health and Medicine: The Business of Private Medical Practice: Doctors, Specialization, and Urban Change in Philadelphia, 1900-1940*. New Brunswick, New Jersey: Rutgers University Press.

Sin, C. H. (2002). The Interpretation of Segregation Indices in Context: The Case of P* in Singapore. *The Professional Geographer*, *54*(3), 422–437.

Spielman, S. E., & Logan, J. R. (2013). Using High-Resolution Population Data to Identify Neighborhoods and Establish Their Boundaries. *Annals of the Association of American Geographers*, *103*(1), 67–84.

Steel, D. G., Holt, D. T., & Tranmer, M. (1994). Modelling and Adjusting Aggregation Effects. In *Conference and CASIC Technologies Interchange* (pp. 382–408). Arlington (Rosslyn), Virginia: Bureau of the Census.

Steel, D. G., Tranmer, M., & Holt, D. T. (2003). Analysis Combining Survey and Geographically Aggregated Data. In R. L. Chambers & C. J. Skinner (Eds.), *Analysis of Survey Data* (pp. 323–343). West Sussex, England: John Wiley & Sons, Ltd.

Subramanian, S. V., Acevedo-Garcia, D., & Osypuk, T. L. (2005). Racial Residential Segregation and Geographic Heterogeneity in Black/White Disparity in Poor Self-Rated Health in the US: A Multilevel Statistical Analysis. *Social Science & Medicine*, *60*(8), 1667–1679.

Sundstrom, R. R. (2004). Racial politics in residential segregation studies. *Philosophy & Geography*, *7*(1), 61–78.

Taeuber, K. E., & Taeuber, A. F. (1965). *Negroes in Cities: Residential Segregation and Neighborhood Change*. New York: Aldine.

Taeuber, K. E., & Taeuber, A. F. (1976). A Practitioner's Perspective on the Index of Dissimilarity. *American Sociological Review*, *41*(5), 884–889.

Taylor, P. J. (1971). Distances within Shapes: An Introduction to a Family of Finite Frequency Distributions. *Geografiska Annaler. Series B, Human Geography*, *53*(1), 40–53.

Thomas, G. E. (1980). Architectural Patronage and Social Stratification in Philadelphia between 1840 and 1920. In W. W. Cutler & H. Gillette Jr. (Eds.), *The Divided*

Metropolis: Social and Spatial Dimensions of Philadelphia, 1800-1975 (pp. 85–124). Westport, Connecticut: Greenwood Press.

Tobler, W. R. (1990). Frame independent spatial analysis. In M. F. Goodchild (Ed.), *The Accuracy of Spatial Databases* (pp. 115–122). London: Taylor & Francis.

Tomlin, C. D. (2017). The Bird's-Eye View from a Worm's-Eye Perspective. In D. A. Griffith, Y. Chun, & D. J. Dean (Eds.), *Advances in Geocomputation: Geocomputation 2015-The 13th International Conference* (pp. 21–31). Cham, Switzerland: Springer.

Tranmer, M., & Steel, D. G. (1998). Using census data to investigate the causes of the ecological fallacy. *Environment and Planning A*, *30*(5), 817–831.

Truesdell, L. E. (1942). *16th Census of the United States, 1940 : Population and Housing : Statistics for Census Tracts, Philadelphia, Pa.* Washington, D.C.: United State Government Printing Office.

Wade, T., & Sommer, S. (Eds.). (2006). *A to Z GIS: An Illustrated Dictionary of Geographic Information Systems* (2nd ed.). Redlands, CA: ESRI Press.

Warner Jr., S. B. (1987). *The Private City: Philadelphia in Three Periods of its Growth* (2nd ed.). Philadelphia: University of Pennsylvania Press.

Watson, K. B., Carlson, S. A., Humbert-Rico, T., Carroll, D. D., & Fulton, J. E. (2015). Walking for Transportation: What do U.S. Adults Think is a Reasonable Distance and Time? *Journal of Physical Activity and Health*, *12*(S1), S53–S61.

Weaver, W. W. (1930). *West Philadelphia: A Study of Natural Social Areas*. University of Pennsylvania.

Weintraub, J. (2008). Why the 72 Year Rule for U.S. Census Privacy? *Journal of the Jewish Genealogical Society of Los Angeles*, *28*(2), 16–17.

White, M. J. (1983). The Measurements of Spatial Segregation. *The American Journal of Sociology*, *88*(5), 1008–1018.

White, M. J. (2015). Segregation, Demographic Effects of. In *International Encyclopedia of Social & Behavioral Sciences* (2nd ed., Vol. 21, pp. 407–410). Elsevier.

Williams, J. J. (1948). Another Commentary on So-Called Segregation Indices. *American Sociological Review*, *13*(3), 298–303.

Wineman, J. D., Marans, R. W., Schulz, A. J., Westhuizen, D. L. Van Der, Mentz, G. B., & Max, P. (2014). Designing Healthy Neighborhoods: Contributions of the Built Environment to Physical Activity in Detroit. *Journal of Planning Education and Research*, *34*(2), 180–189.

Wong, D. W. S. (1993). Spatial Indices of Segregation. *Urban Studies*, *30*(3), 559–572.

Wong, D. W. S. (1997). Spatial Dependency of Segregation Indices. *The Canadian Geographer*, *41*(2), 128–136.

Wong, D. W. S. (2002a). Modeling Local Segregation: A Spatial Interaction Approach. *Geographical and Environmental Modelling*, *6*(1), 81–97.

Wong, D. W. S. (2002b). Spatial Measures of Segregation and GIS. *Urban Geography*, *23*(1), 85–92.

Wong, D. W. S. (2003a). Implementing spatial segregation measures in GIS. *Computers, Environment and Urban Systems*, *27*(1), 53–70.

Wong, D. W. S. (2003b). Spatial Decomposition of Segregation Indices: A Framework Toward Measuring Segregation at Multiple Levels. *Geographical Analysis*, *35*(3), 179–194.

Wong, D. W. S. (2004). Comparing Traditional and Spatial Segregation Measures: A Spatial Scale Perspective. *Urban Geography*, *25*(1), 66–82.

Wong, D. W. S. (2009). The Modifiable Areal Unit Problem (MAUP). In *The SAGE Handbook of Spatial Analysis* (pp. 105–124). Sage Publications Ltd.

Wong, D. W. S., & Chong, W. K. (1998). Using Spatial Segregation Measures in GIS and Statistical Modeling Packages. *Urban Geography*, *19*(5), 477–485.

Wong, D. W. S., Lasus, H. A., & Falk, R. F. (1999). Exploring the variability of segregation index D with scale and zonal systems: an analysis of thirty US cities. *Environment and Planning A*, *31*(3), 507–522.

Wong, D. W. S., & Shaw, S.-L. (2011). Measuring Segregation: An Activity Space Approach. *Journal of Geographical Systems*, *13*(2), 127–145.

Works Progress Administration (WPA). (1939). WPA Project 20879: Social Base Map of Philadelphia. Philadelphia, PA.

Yang, T.-C., & Matthews, S. A. (2015). Death by Segregation: Does the Dimension of Racial Segregation Matter? *PLoS ONE*, *10*(9), 1–26.

Yang, Y., & Diez-Roux, A. V. (2012). Walking Distance by Trip Purpose and Population Subgroups. *American Journal of Preventive Medicine*, *43*(1), 11–19.

Zhang, M., & Kukadia, N. (2005). Metrics of Urban Form and the Modifiable Areal Unit Problem. *Transportation Research Record: Journal of the Transportation Research Board*, *1902*, 71–79.