

RESOURCE SHEET:

RESIDENTIAL SEGREGATION MEASURES – LOCATION, LOCATION, LOCATION!

Residential segregation plays a key role in understanding places, and has significant ties to public health, especially in urban settings like Philadelphia. This brief will present several measures related to residential segregation, including their formula (with key references), interpretations, and a non-exhaustive list of existing empirical evidence that utilizes each measure.

TOPICS COVERED:

- **Residential Segregation Measures**

INTRODUCTION

Social determinants of health (SDOH) refer to the conditions and environments that people are born, live, learn, work, play, worship and age in that not only affect their health and quality of life, but also help to explain the racial, ethnic, and social inequities present in the United States (US) and globally (Healthy People 2030, 2022). These social determinants of health have been documented to be tied to structural racism, which can be defined as the “the totality of ways in which societies foster [racial] discrimination, via mutually reinforcing [inequitable] systems...(e.g., in housing, education, employment, earnings, benefits, credit, media, health care, criminal justice, etc.) that in turn reinforce discriminatory beliefs, values, and distribution of resources” (Krieger, 2014). Recent calls to integrating racism as an important indicator of public health research and surveillance have been documented (White et al., 2023) And, while structural racism can be operationalized several ways, one common approach is to utilize metrics related to residential segregation – which represents a domain of structural racism and has been used to frame the context of health risk and health disparities and inequities (Steil & Arcaya, 2023; White et al., 2023; White & Borrell, 2011; Williams & Collins, 2001).

Historically, in the late 1800s, when considering the importance of neighborhoods, scholars and activists like W. E. B. Du Bois captured the health and social problems present in African American communities in Philadelphia, PA (Du Bois & Eaton, 1899). This fieldwork allowed Du Bois and others to better understand the importance of neighborhood structures, and to highlight inequities in these communities in education, employment, health, and other social outcomes. Notably, in the next century, in the 1930s, neighborhoods were characterized by the federally sponsored Home Owners’ Loan Corporation (HOLC), such that sections of major cities like Philadelphia across the US were deemed unworthy based on their sociodemographic composition (Aaronson et al., 2021; Greer, 2013; Michney, 2022; Rothstein, 2021; Woods et al., 2014).

Specifically, neighborhoods with high percentages of low-income immigrant or Black residents received a lower grade (i.e., “redlined”) compared to other neighborhoods – this legal act perpetuated the structural racism that already existed during this time in US history, and has contributed to contemporary segregation, disinvestment, and inequities in home ownership and wealth (Swope et al., 2022) among other documented unfavorable health and social outcomes.

Today, residential segregation, plays a key role in understanding places, and has significant ties to public health, especially in urban settings like Philadelphia. Health outcomes like cancer (Goel et al., 2022; Krieger, Singh, et al., 2016; Westrick et al., 2020), asthma (Alexander & Currie, 2017; Nardone et al., 2020), hypertension (Feldman et al., 2021), preterm birth (Chambers et al., 2019; Yang et al., 2024), and cardiovascular health (Allgood et al., 2024; Tabb et al., 2022) have all been tied to residential segregation; however, the operationalization of residential segregation as a quantitative measure varies (Tabb et al., 2024) and recent reviews have further examined how these types of measures are operationalized in public health research more broadly (McCulley et al., 2025).

We will present several measures in this resource document, including their formula (with key references), interpretations, and a non-exhaustive list of existing empirical evidence that utilizes each measure. Specifically, this resource document includes information on: the Dissimilarity Index, the Gini Index, the Index of Concentration at the Extremes, and the G_i^* Statistic.

DISSIMILARITY INDEX

Dissimilarity Index, a segregation measure of evenness, represents the percentage of the minority population that would need to move for the two groups (minority and majority) to be evenly distributed in the neighborhood. The index is often based on self-report race/ethnicity from sources like the US Census and the American Community Survey (ACS). The index has been shown to be associated with a number of health and social outcomes like mortality, pregnancy, self-rated health, firearm homicides, air pollution and COVID-19 (see examples of empirical evidence referenced below).

Formula (comparing Black and White residents, as an example)

$$D_w^b = 100 \times \frac{1}{2} \sum_i |w_i/W - b_i/B|$$

Where i indexes a smaller region (i.e., census tract), compared to a larger region (i.e., a metropolitan area defined as a core-based statistical area), and b_i , w_i , B , W represent the census tract population of Black and White residents, and the core-based statistical area population of Black and White residents, respectively. The index is multiplied by 100 to represent a percentage – rather than a proportion.

The support of the Dissimilarity Index ranges from 0 (indicative of no residential segregation or complete integration) to 100 (indicative of complete residential segregation), where a value over 60 typically signifies high segregation (Massey & Fischer, 1999). While this representation compares Black and White residents, the comparison can be made between any two racial/ethnic groups.

INTERPRETATION

A Dissimilarity Index of 75 comparing Black and White residents indicates that 75% of the Black residents would need to move to achieve a uniform (i.e. even) distribution of the population by race – based on these two groups.

KEY REFERENCE

Duncan, O. D., & Duncan, B. (1955). A methodological analysis of segregation indexes. *American sociological review*, 20(2), 210–217.

EXAMPLES OF EMPIRICAL EVIDENCE (ORDERED ALPHABETICALLY)

Cooper, R. S., Kennelly, J. F., Durazo-Arvizu, R., Oh, H. J., Kaplan, G., & Lynch, J. (2001). Relationship between premature mortality and socioeconomic factors in black and white populations of US metropolitan areas. *Public health reports*.

Ellen, I. G., Cutler, D. M., & Dickens, W. (2000). Is segregation bad for your health? The case of low birth weight [with comments]. *Brookings-Wharton papers on urban affairs*, 203-238.

Frankenfeld, C. L., Hakes, J. K., & Leslie, T. F. (2022). All-cause mortality and residential racial and ethnic segregation and composition as experienced differently by individual-level race, ethnicity, and gender: mortality disparities in American communities data. *Annals of Epidemiology*, 65, 38-45.

Hart, K. D., Kunitz, S. J., Sell, R. R., & Mukamel, D. B. (1998). Metropolitan governance, residential segregation, and mortality among African Americans. *American journal of public health*, 88(3), 434-438.

Knopov, A., Rothman, E. F., Cronin, S. W., Franklin, L., Cansever, A., Potter, F., ... & Hemenway, D. (2019). The role of racial residential segregation in black-white disparities in firearm homicide at the state level in the United States, 1991-2015. *Journal of the National Medical Association*, 111(1), 62-75.

Kramer, M. R., & Hogue, C. R. (2008). Place matters: variation in the black/white very preterm birth rate across US metropolitan areas, 2002–2004. *Public Health Reports*, 123(5), 576-585.

Niemesh, G. T., & Shester, K. L. (2020). Racial residential segregation and black low birth weight, 1970–2010. *Regional Science and Urban Economics*, 83, 103542.

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.

Woo, B., Kravitz-Wirtz, N., Sass, V., Crowder, K., Teixeira, S., & Takeuchi, D. T. (2019). Residential segregation and racial/ethnic disparities in ambient air pollution. *Race and social problems*, 11, 60-67.

Yang, T. C., Emily Choi, S. W., & Sun, F. (2021). COVID-19 cases in US counties: roles of racial/ethnic density and residential segregation. *Ethnicity & Health*, 26(1), 11-21.

Yu, Q., Salvador, C. E., Melani, I., Berg, M. K., Neblett, E. W., & Kitayama, S. (2021). Racial residential segregation and economic disparity jointly exacerbate COVID-19 fatality in large American cities. *Annals of the New York Academy of Sciences*, 1494(1), 18-30.

GINI INDEX

The Gini Index is a segregation measure that was initially developed as a measure of statistical dispersion to operationalize income inequality. The index is sometimes referred to as the Gini coefficient or the Gini ratio. The index has been shown to be associated with a number of health and social outcomes like mortality, suicide, cancer and COVID-19 (see examples of empirical evidence referenced below).

FORMULA

A simplified version of the Gini Index assumes an exponential function ($y(x) = x^P$) along with an assumed Pareto distribution ($y(x) = 1 - (1 - x)^P$). When combined, the following is assumed:

$$y(x) = (1 - k)x^P + k(1 - (1 - x)^{1/P})$$

The parameters' support is $0 \leq k \leq 1$ and $P \geq 1$, where the parameter k is the weight that controls the curvature of the Lorenz curve, and the parameter P represents the degree of inequality in income distribution. Because of this relationship, the area under the Lorenz curve and the mathematical formulation of the Gini Index is as follows:

$$\int_0^1 y(x)dx = \frac{1}{P+1}$$

$$\text{Gini} = 1 - 2 \int_0^1 y(x)dx = \frac{P-1}{P+1}$$

$$0 \leq \text{Gini} \leq 1$$

A graphical depiction of the Lorenze curve is presented in Figure 1, and, as the degree of income inequality increases, the Lorenz curve deepens and the area between the perfect equality line and the curve widens.

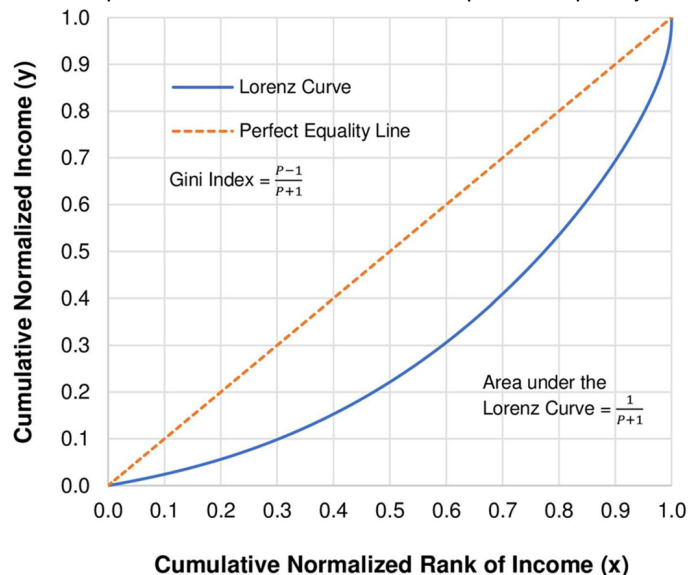


Fig. 1. Lorenz curve with Gini Index calculation. Given the specified functional form that is constructed based on the weighted average of the exponential function and the functional form implied by Pareto distribution, the closed form expression for the Gini index can be conveniently computed as $Gini = 1 - 2 \int_0^1 y(x)dx = \frac{P-1}{P+1}$.

Image reference: Fig. 1 presented from Sitthiyot, T., & Holasut, K. (2020). A simple method for measuring inequality. Palgrave Communications, 6(1), 1–9. <https://www.nature.com/articles/s41599-020-0484-6#Fig1>.

The support of the Gini Index ranges from 0 (indicative of perfect income equality) to 1 (indicative of perfect income inequality), where less than 0.2 typically corresponds to perfect income equality, 0.3-0.4 indicates reasonable income gaps, 0.4-0.5 indicates high income disparity and larger than 0.5 indicates severe income disparity (Haddad et al., 2024).

INTERPRETATION

A Gini Index of 0 indicates that all the residents living in the region have an equal share of the income that's possible in the region; whereas a Gini Index of 1 indicates that one group of individuals receives all the income.

KEY REFERENCE

Sitthiyot, T., & Holasut, K. (2020). A simple method for measuring inequality. Palgrave Communications, 6(1), 1-9.

EXAMPLES OF EMPIRICAL EVIDENCE (ORDERED ALPHABETICALLY)

Herb, J., Holmes, M., & Stitzenberg, K. (2022). Trends in rural-urban disparities among surgical specialties treating cancer, 2004-2017. The Journal of Rural Health, 38(4), 838-844.

Kawachi, I., & Kennedy, B. P. (1997). The relationship of income inequality to mortality: does the choice of indicator matter?. *Social science & medicine*, 45(7), 1121-1127.

Oronce, C. I. A., Scannell, C. A., Kawachi, I., & Tsugawa, Y. (2020). Association between state-level income inequality and COVID-19 cases and mortality in the USA. *Journal of general internal medicine*, 35, 2791-2793.

Rodgers, G. B. (1979). Income and inequality as determinants of mortality: an international cross-section analysis. *Population studies*, 33(2), 343-351.

Swan, B. Q. (2021). Economic disparities and suicides: The dynamic panel data analyses of 50 states in the United States. *J. Forensic Sci. Res*, 5, 20-29.

Truman, B. I., Smith, K. C., Roy, K., Chen, Z., Moonesinghe, R., Zhu, J., ... & Centers for Disease Control and Prevention (CDC). (2011). Rationale for regular reporting on health disparities and inequalities—United States. *MMWR Surveill Summ*, 60(Suppl 01), 3-10.

Wilkinson, R. G. (1990). Income distribution and mortality: a 'natural' experiment. *Sociology of health & illness*, 12(4), 391-412.

INDEX OF CONCENTRATION OF THE EXTREMES (ICE)

The Index of Concentration of the Extremes (ICE) was initially developed as a measure of concentrated affluence and poverty in a neighborhood. The index originally captured one social construct for extreme comparisons but has since been extended to capture multiple social constructs like race/ethnicity and income combined – when measuring segregation in a neighborhood. The index has been shown to be associated with a number of health and social outcomes like police-related deaths, fatal and non-fatal assaults, and COVID-19 (see examples of empirical evidence referenced below).

FORMULA

The original ICE formulation (Massey, 1996) is:

$$ICE_i = \frac{(A_i - P_i)}{T_i}$$

where it is a function of three distinct populations: A_i represents the total number of affluent/privileged residents in neighborhood i , P_i represents the total number of poor/disadvantaged residents, and T_i represents the total number of residents in a neighborhood where income is measured and available – where “poor” and “affluent” are subjective but noted income (often based on self-report) usage has assumed “poor” earning under \$25,000 and “affluent” earning over \$100,000 annually in the US (Feldman et al., 2015).

The support of the ICE measure ranges from -1 (disadvantaged) to 1 (privileged), where 0 indicates none of the residents in the neighborhood are in the best-off or worst-off categories (an equal number of persons are in the best-off and worst-off categories).

INTERPRETATION

An ICE measure of -0.25 indicates that the neighborhood has a larger proportion of disadvantaged residents than privileged residents; whereas an ICE measure of 0.85 indicates that a noticeably larger proportion of residents are privileged. If the ICE metric is based on education, for example, the two comparison groups or extremes could be defined as less educated (i.e., disadvantaged) and more educated (i.e., privileged).

KEY REFERENCE(S) – CHRONOLOGICALLY ORDERED

Massey, D. S. (1996). The age of extremes: Concentrated affluence and poverty in the twenty-first century. *Demography*, 33, 395-412.

Krieger, N., Waterman, P. D., Gryparis, A., & Coull, B. A. (2015). Black carbon exposure, socioeconomic and racial/ethnic spatial polarization, and the Index of Concentration at the Extremes (ICE). *Health & place*, 34, 215-228.

Krieger, N., Waterman, P. D., Spasojevic, J., Li, W., Maduro, G., & Van Wye, G. (2016). Public health monitoring of privilege and deprivation with the index of concentration at the extremes. *American journal of public health*, 106(2), 256-263.

EXAMPLES OF EMPIRICAL EVIDENCE (ORDERED ALPHABETICALLY)

Chen, J. T., & Krieger, N. (2021). Revealing the unequal burden of COVID-19 by income, race/ethnicity, and household crowding: US county versus zip code analyses. *Journal of Public Health Management and Practice*, 27(Supplement 1), S43-S56.

Feldman, J. M., Waterman, P. D., Coull, B. A., & Krieger, N. (2015). Spatial social polarisation: using the Index of Concentration at the Extremes jointly for income and race/ethnicity to analyse risk of hypertension. *J Epidemiol Community Health*, 69(12), 1199-1207.

Feldman, J. M., Gruskin, S., Coull, B. A., & Krieger, N. (2019). Police-related deaths and neighborhood economic and racial/ethnic polarization, United States, 2015–2016. *American journal of public health*, 109(3), 458-464.

Krieger, N., Feldman, J. M., Waterman, P. D., Chen, J. T., Coull, B. A., & Hemenway, D. (2017). Local residential segregation matters: stronger association of census tract compared to conventional city-level measures with fatal and non-fatal assaults (total and firearm related), using the index of concentration at the extremes (ICE) for racial, economic, and racialized economic segregation, Massachusetts (US), 1995–2010. *Journal of urban health*, 94, 244-258.

Krieger, N., Kim, R., Feldman, J., & Waterman, P. D. (2018). Using the Index of Concentration at the Extremes at multiple geographical levels to monitor health inequities in an era of growing spatial social polarization: Massachusetts, USA (2010–14). *International journal of epidemiology*, 47(3), 788-819.

Williams, D. R., & Collins, C. (2001). Racial residential segregation: a fundamental cause of racial disparities in health. *Public health reports*.

G_i^* STATISTIC

The G_i^* Statistic (also referred to as Getis-Ord Gi-star) is a segregation measure that shows how a neighborhood's racial/ethnic composition deviates from the mean racial/ethnic composition of the larger region. The index has been shown to be associated with a number of health and social outcomes like cardiovascular health, BMI, spatial epidemiology broadly, and COVID-19 (see examples of empirical evidence referenced below).

FORMULA

$$G_i^* = \frac{\sum_j w_{i,j} x_j - \bar{X} \sum_j w_{i,j}}{S \sqrt{(\sum_j w_{i,j}^2) - (\sum_j w_{i,j})^2 / n}}$$

Where x_j is the value at location j , $w_{i,j}$ is the spatial weight between locations i and j , $\bar{X}$ and S are the mean and standard deviation of all x_j 's, respectively, and n represents the total number of observations (i.e., geographical areas).

The support for the G_i^* is $-\infty$ to $+\infty$. Larger positive G_i^* values indicate spatial clustering of higher values and often indicate hot spots in the region of consideration; whereas, low negative G_i^* values indicate spatial clustering of low values (i.e. cold spots).

INTERPRETATION

A G_i^* value of 2.05 indicates that the neighborhood, and its surrounding neighboring areas, have significantly higher, say, proportions of a particular racial/ethnic group than expected. The value of 2.05 would suggest evidence of a hot spot.

KEY REFERENCE

Getis, A., & Ord, J. K. (1992). The analysis of spatial association by use of distance statistics. *Geographical analysis*, 24(3), 189-206.

INTERPRETATION

A G_i^* value of 2.05 indicates that the neighborhood, and its surrounding neighboring areas, have significantly higher, say, proportions of a particular racial/ethnic group than expected. The value of 2.05 would suggest evidence of a hot spot.

EXAMPLES OF EMPIRICAL EVIDENCE (ORDERED ALPHABETICALLY)

Barber, S., Roux, A. V. D., Cardoso, L., Santos, S., Toste, V., James, S., ... & Chor, D. (2018). At the intersection of place, race, and health in Brazil: Residential segregation and cardio-metabolic risk factors in the Brazilian Longitudinal Study of Adult Health (ELSA-Brasil). *Social science & medicine*, 199, 67-76.

Lopes, M. S., Caiaffa, W. T., de Souza Andrade, A. C., Malta, D. C., Barber, S., & de Lima Friche, A. A. (2020). Disparities in food consumption between economically segregated urban neighbourhoods. *Public health nutrition*, 23(3), 525-537.

Kershaw, K. N., Barber, S., & Hicken, M. T. (2024). Current approaches to measuring local racial and ethnic residential segregation in population health studies. *Current Epidemiology Reports*, 11(1), 32-43. Kershaw, K. N., Osypuk, T. L., Do, D. P., De Chavez, P. J., & Diez Roux, A. V. (2015). Neighborhood-level racial/ethnic residential segregation and incident cardiovascular disease: the multi-ethnic study of atherosclerosis. *Circulation*, 131(2), 141-148.

Kershaw, K. N., Barber, S., & Hicken, M. T. (2024). Current approaches to measuring local racial and ethnic residential segregation in population health studies. *Current Epidemiology Reports*, 11(1), 32-43.

Maroko, A. R., Nash, D., & Pavilonis, B. T. (2020). COVID-19 and inequity: a comparative spatial analysis of New York City and Chicago hot spots. *Journal of Urban Health*, 97, 461-470.

Nayak, P. P., Pai, J. B., Singla, N., Somayaji, K. S., & Kalra, D. (2021). Geographic information systems in spatial epidemiology: Unveiling new horizons in dental public health. *Journal of International Society of Preventive and Community Dentistry*, 11(2), 125-131.

ADDITIONAL MEASURES

While this resource document focused on the description, formula, interpretation, key references and examples of empirical evidence for the Dissimilarity Index, the Gini Index, the Index of Concentration at the Extremes, and G_i^* Statistic, this does not represent an exhaustive list of residential segregation measures. A few additional measures, as well as the ones presented above, can be found in the table below.

Name	What It Measures	Type	Range	Spatial?	Key References
Dissimilarity Index	Uneven distribution of two groups	Unevenness	0 – 1	No	(Massey & Denton, 1988)
Gini Index	Inequality in distribution of a resource or group	Inequality	0 – 1	No	(Reardon & Firebaugh, 2002)

ICE	Relative concentration of privilege vs. deprivation	Polarization	-1 to +1	No	(Krieger et al., 2015) (Krieger, Waterman, et al., 2016)
G_i^* Statistic	Spatial clustering of high/low values	Hotspot analysis (z-score)	$-\infty$ to $+\infty$	Yes	(Getis & Ord, 1992)
ADI (Area Deprivation Index)	Composite socioeconomic disadvantage	Deprivation index	Varies	Some versions	(Kind & Buckingham, 2018)
Entropy Index (Theil's H)	Multi-group segregation/diversity	Information theory	0 – 1	No	(Reardon & Firebaugh, 2002)
Isolation Index	Exposure of group to itself	Exposure	0 – 1	No	(Massey & Denton, 1988)
Interaction Index	Exposure of group to other group(s)	Exposure	0 – 1	No	(Massey & Denton, 1988)
Clustering Index	Adjacency of group members across areas	Spatial contiguity	0 – 1	Yes	(White, 1983)
Centralization Index	Proximity to city center	Spatial form	0 – 1	Yes	(Massey & Denton, 1988)
Local Moran's I	Local spatial autocorrelation	Spatial clustering	-1 to +1	Yes	(Anselin, 1995)
Neighborhood Sorting Index	Sorting of households by traits into neighborhoods	Sorting/economics	0 – 1	No	(Bayer et al., 2007)
Multigroup Entropy / Divergence Index	Multi-ethnic/racial segregation	Information theory	0 – 1	Can be	(Reardon & Firebaugh, 2002)

FINAL THOUGHTS

This resource document presents several commonly utilized residential segregation measures within the urban health research and practice space, including their formula (with key references), interpretations, and a non-exhaustive list of existing empirical evidence that utilizes each measure. All measures have been shown

to be associated with a myriad of health and social outcomes, especially within urban settings. While each measure has distinct ways of capturing the separation of people in neighborhoods based on constructs like race/ethnicity, income, education, and housing, for example, it's imperative to ensure researchers understand the operationalization and functionality to determine what's most appropriate for their study design and setting.

RESIDENTIAL SEGREGATION DATA ON THE UHC DATA PORTAL

The UHC Research and Data Core has calculated a number of residential segregation measures for US core-based statistical areas (CBSA), counties, and census tracts using 2010 census boundaries including dissimilarity indices, isolation indices, G_i^* Statistic, and Index of Concentration (ICE). These measures were derived from US Census and American Community Survey (ACS) data for a number of years. Details about measures available can be found on the UHC Data Portal (data.uhcdata.org).

WHO DO I CONTACT IF I HAVE QUESTIONS?

If you have any questions about the available data or would like to discuss your needs, please contact the Urban Heath Collaborative data team at uhcdata@drexel.edu.