DEMOGRAPHIC RESEARCH
*A peer-reviewed, open-access journal of population sciences*

***DEMOGRAPHIC RESEARCH***

**VOLUME 32, ARTICLE 54, PAGES 1469–1486**
**PUBLISHED 10 JUNE 2015**
http://www.demographic-research.org/Volumes/Vol32/54/
DOI: 10.4054/DemRes.2015.32.54

*Descriptive Finding*

# Measuring a neighborhood affluence-deprivation continuum in urban settings: Descriptive findings from four US cities

**Masayoshi Oka**

# Table of Contents

# Measuring a neighborhood affluence-deprivation continuum in urban settings: Descriptive findings from four US cities

**Masayoshi Oka**[1]

## Abstract

**BACKGROUND**
In the United States (US), the area-based measure of neighborhood socioeconomic characteristics used in health research varies considerably from one study to another. However, it is unclear whether different area-based measures capture the same or different dimension of neighborhood context.

**OBJECTIVE**
The purpose of this study is to examine the relationships between single measures (i.e., area-based median household income and median family income) and composite measures (i.e., area-based measures derived from a combination of multiple variables) of neighborhood socioeconomic characteristics.

**METHODS**
Area-based socioeconomic data at the census tract level were obtained from the 2005–09 American Community Survey (ACS) for St. Louis, Missouri; Chicago, Illinois; San Diego, California; and Los Angeles, California. Single measures of neighborhood socioeconomic characteristics were simply taken from the ACS data, and composite measures were derived from the computational methods described in previous studies. Separate correlation statistics were then conducted for four US cities.

**RESULTS**
Despite the differences in how selected area-based measures of neighborhood socioeconomic characteristics were derived from the ACS data, they were highly correlated (either negatively or positively) with one another. In other words, selected area-based measures capture the same dimension of neighborhood context.

**CONCLUSIONS**
A neighborhood affluence-deprivation continuum in US cities may be captured by an area-based median household (or family) income. Nevertheless, to ensure the

[1] Current affiliation: Social and Cardiovascular Epidemiology Research Group, Faculty of Medicine, University of Alcalá, 28871 Alcalá de Henares (Madrid), Spain. E-Mail: moka@post.harvard.edu.
Previous affiliation (at the time the study was conducted): Division of Public Health Sciences, Department of Surgery, School of Medicine, Washington University in St. Louis, St. Louis, MO 63110, U.S.A.

generalizability and transportability of results from four US cities, further comparisons of area-based measures (not limited to those considered in this study) are needed in different US cities.

## 1. Introduction

The social characteristics of a neighborhood may either promote or hinder the health (i.e., health behaviors and health outcomes) of individuals residing in them (Kawachi and Berkman 2003). Review articles have summarized a growing number of studies highlighting the importance of place to the health of its residents in the United States (US) and other countries (e.g., Pickett and Pearl 2001; Riva, Gauvin, and Barnett 2007; Black and Macinko 2008; Kim 2008; Mair, Diez Roux, and Galea 2008; Chaix 2009; Meijer et al. 2012). Various health-related outcomes examined in previous studies include, but are not limited to: cardiovascular disease, depressive symptoms or depression, low birth weight, mortality, obesity, and self-rated health status. Overall, the importance of place of residence has been demonstrated after accounting for individual sociodemographic characteristics (e.g., age, gender, race/ethnicity, educational achievement, income level, and/or employment status) in a multilevel regression analysis.

Among previous studies (Pickett and Pearl 2001; Riva, Gauvin, and Barnett 2007; Black and Macinko 2008; Kim 2008; Mair, Diez Roux, and Galea 2008; Chaix 2009; Meijer et al. 2012), the most common research objective was to examine the association of neighborhood socioeconomic characteristics with health. In US studies, an administrative unit at the census tract or block group level (sometimes zip-code tabulation areas) has been used to denote 'neighborhoods' in which individuals reside. A composite measure (i.e., a measure derived from a combination of multiple variables) of neighborhood socioeconomic characteristics is constructed from the US Census data; area-based socioeconomic data is compiled representing several key domains relevant to health (e.g., educational attainment, income level, poverty status, employment status, housing condition, and occupation). However, the construction of composite measures varies considerably from one study to another. For instance, area-based indexes of socioeconomic advantage (SEA) (Diez Roux et al. 2001), socioeconomic status (SES) (Winkleby and Cubbin 2003), socioeconomic position (SEP) (Krieger et al. 2003), socioeconomic deprivation (SED) (Singh 2003), and deprivation (DEP) (Messer et al. 2006), among others, have been developed by different authors in the US (Table 1). Of equal importance, a single measure of neighborhood socioeconomic characteristics (i.e., an area-based median household or median family income) has also been used in some

studies (e.g., Galea et al. 2007; Stockdale et al. 2007; Tonorezos et al. 2008; Sallis et al. 2009; Black and Macinko 2010; King et al. 2011) instead of a composite measure. Therefore, a natural question arises: do single and composite measures of neighborhood socioeconomic characteristics capture the same dimension of neighborhood context?

In order to better understand the area-based variations in health, the purpose of this study is to examine the relationships between single measures (i.e., area-based median household income and median family income) and composite measures (Diez Roux et al. 2001; Krieger et al. 2003; Singh 2003; Winkleby and Cubbin 2003; Messer et al. 2006) of neighborhood socioeconomic characteristics. Two important remarks in health research are discussed in this paper. St. Louis, Missouri (MO); Chicago, Illinois (IL); San Diego, California (CA); and Los Angeles, CA were used as examples to illustrate these points.

## 2. Methods

### 2.1 Data

Area-based socioeconomic data at the census tract level were obtained from the 2005–09 American Community Survey (ACS) for St. Louis, MO (St. Charles County, St. Louis County, and St. Louis City); Chicago, IL (Cook County); San Diego, CA (San Diego County); and Los Angeles, CA (Los Angeles County). Note that the five-year ACS estimates are based on a larger sample size (therefore, more accurate) than the one- and three-year estimates. In this study, county boundaries were used to denote 'cities' including periurban areas (i.e., suburban areas surrounding the urban area) in each locality, instead of city boundaries that mostly reflect city centers. The use of county boundaries is based on the notion that periurban areas are an important source of annexation for understanding the health of cities (e.g., Vlahov and Galea 2002).

**Table 1: Description of selected area-based indexes developed in the United States (US)**

| | | Area-based Indexes | | | | |
|---|---|---|---|---|---|---|
| | | **SEA** | **SES** | **SEP** | **SED** | **DEP** |
| *Variables* | | | | | | |
| Median Household Income | (US $) | X | | X | | |
| Median Family Income | (US $) | | X | | X | |
| Median Home Value | (US $) | X | X | | X | |
| Median Gross Rent | (US $) | | | | X | |
| Median Monthly Mortgage | (US $) | | | | X | |
| Income Disparity | | | | | X | |
| Household Income <$30,000 | (%) | | | | | X |
| Expensive Homes | (%) | | | X | | |
| With Public Assistance Income | (%) | | | | | X |
| With Interest, Dividend, or Net Rental Income | (%) | X | | | | |
| Below Poverty | (%) | | | X | X | X |
| Below Poverty <150% | (%) | | | | X | |
| Less than 9th Grade Education | (%) | | | | X | |
| Less than High School Degree | (%) | | X | X | | X |
| High School Degree | (%) | X | | | | |
| More than High School Degree | (%) | | | | X | |
| Bachelor's Degree or Higher | (%) | X | | | | |
| White Collar Occupation | (%) | X | | | X | |
| Blue Collar Occupation | (%) | | X | X | | |
| Males in Management Occupation | (%) | | | | | X |
| Unemployed | (%) | | X | X | X | X |
| Owner Occupied Housing Units | (%) | | | | X | |
| Household Crowding | (%) | | | | X | X |
| Single-Parent Households with Dependents | (%) | | | | X | |
| Female Households with Dependents | (%) | | | | | X |
| No Vehicles | (%) | | | | X | |
| No Telephone | (%) | | | | X | |
| No Complete Plumbing | (%) | | | | X | |
| *Methods* | | | | | | |
| Sum of the Z-Scores | | X | | X | | |
| Factor Analysis | | | | | X | X |
| Principal Component Analysis | | | X | | | |

*Notes*: SEA: socioeconomic advantage (Diez Roux et al. 2001), SES: socioeconomic status (Winkleby and Cubbin 2003), SEP: socioeconomic position (Krieger et al. 2003), SED: socioeconomic deprivation (Singh 2003), and DEP: deprivation (Messer et al. 2006). For more details, refer to the original articles.

Four US cities were chosen to compare two large cities of different sizes, where the population size in one city is roughly one-third of another city, in two different regions of the US. Within these cities, census tracts were used to denote 'neighborhoods' for three main reasons: i) an area-based median family income (US $) and various area-based socioeconomic data needed in the construction of composite measures (Diez Roux et al. 2001; Krieger et al. 2003; Singh 2003; Winkleby and Cubbin 2003; Messer et al. 2006) were not available at the block group level, ii) unlike other areal units (i.e., block groups or zip-code tabulation areas), census tracts are designed to be relatively homogeneous in regard to population characteristics, economic status, and living conditions; in urban areas, census tracts generally consist of about 4,000 inhabitants (US Census Bureau 2013), and iii) census tracts are a national creation of democratic governance informed by local inputs, and are also historically in accordance with uniform standards (Krieger 2006).

Based on the 2005–09 ACS estimates, a total population of 1,692,563 was distributed across 340 census tracts in St. Louis, MO; 5,257,001 across 1,327 census tracts in Chicago, IL; 2,987,543 across 605 census tracts in San Diego, CA; and 9,785,295 across 2,047 census tracts in Los Angeles, CA. However, due to missing data, 11 (about 3.24%), 105 (about 7.91%), 13 (about 2.15%), and 82 (about 4.01%) census tracts were removed from St. Louis, MO; Chicago, IL; San Diego, CA; and Los Angeles, CA, respectively.

## 2.2 Measures

Following the computational methods described in previous studies, composite measures of SEA (Diez Roux et al. 2001), SES (Winkleby and Cubbin 2003), SEP (Krieger et al. 2003), SED (Singh 2003), and DEP (Messer et al. 2006) were calculated in R (R Core Team 2013). Area-based median household income (MHI) and median family income (MFI) were used in their raw forms (US $). However, since the distributions of these two single measures were skewed, Box-Cox (i.e., parametric power) transformations (Box and Cox 1964) were applied to follow a normal distribution. In brief, the Box-Cox transformation finds the maximum likelihood estimation of the parameter λ (lambda). The optimal value of λ is then used as a guide for the power transformation. Box-Cox transformations of two single measures were carried out in R (Fox and Weisberg 2011). Box-Cox transformed MHI and MFI are denoted as $MHI^{(\lambda)}$ and $MFI^{(\lambda)}$, respectively.

In the preliminary analysis, square root and log transformations were also applied to MHI and MFI. However, Box-Cox transformations yielded better normal distributions (data not shown). This reflects the notion that the Box-Cox transformation

provides a potential best practice for optimally normalizing skewed variables than the traditional (e.g., square root, log, and inverse) transformation (Osborne 2010). Therefore, square root and log transformed MHI and MFI were not considered in this study.

### 2.3 Analysis

To examine the relationships between four single measures (i.e., MHI, MFI, $MHI^{(\lambda)}$, and $MFI^{(\lambda)}$) and five composite measures (i.e., SEA, SES, SEP, SED, and DEP) of neighborhood socioeconomic characteristics, four separate correlation statistics (Friendly 2002) were conducted in R (Wright 2012) for St. Louis, MO (Figure 1); Chicago, IL (Figure 2); San Diego, CA (Figure 3); and Los Angeles, CA (Figure 4). Correlations and scatterplot matrices were used to display the linear relationships. The upper off-diagonal panels show the correlation coefficients with associated 95% confidence intervals (in parentheses), and the lower off-diagonal panels show the scatter plots.

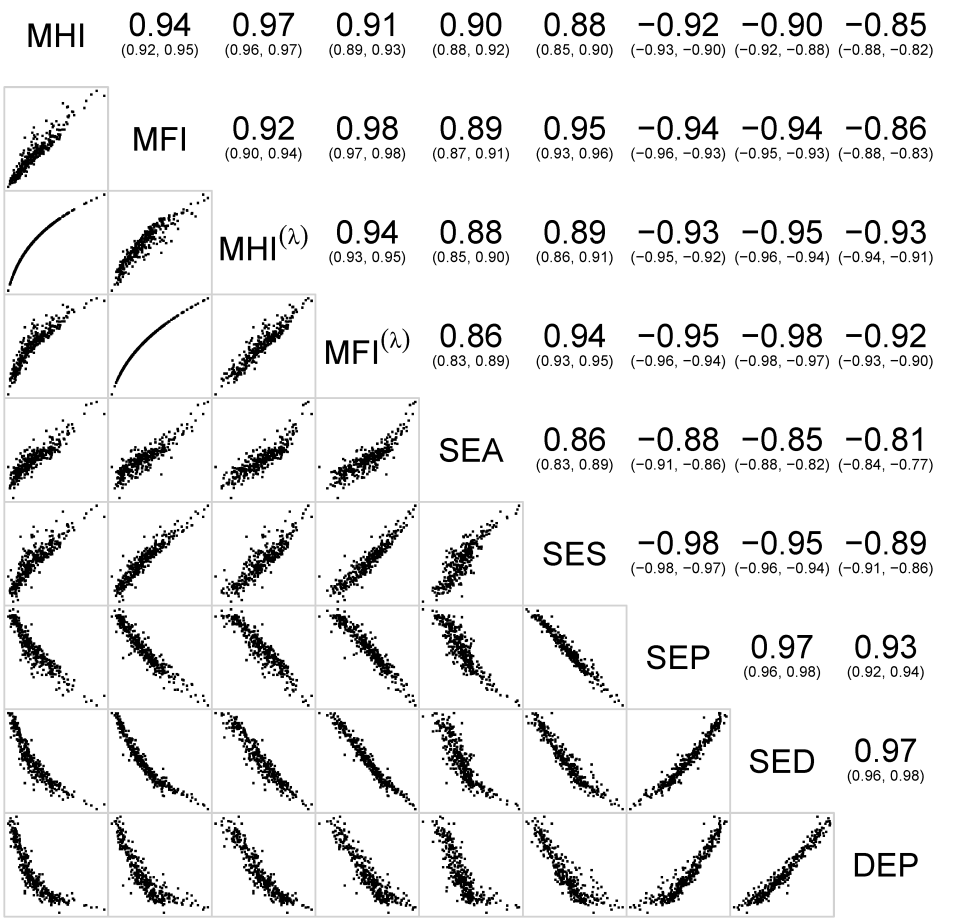
## 3. Results

Four US cities were examined to account for the differences in geographical regions and population sizes. In short, St. Louis, MO and Chicago, IL are located in the Midwestern US, whereas San Diego, CA and Los Angeles, CA are located in the Western US. The total population in St. Louis, MO and San Diego, CA are approximately one-third of that in Chicago, IL and Los Angeles, CA, respectively. Importantly, the racial and ethnic compositions are quite different in these four US cities (data not shown).

Figures 1 through 4 show the relationships between four single and five composite measures of neighborhood socioeconomic characteristics in St. Louis, MO (Figure 1); Chicago, IL (Figure 2); San Diego, CA (Figure 3); and Los Angeles, CA (Figure 4). By and large, correlations and scatterplot matrices display consistent patterns in these four US cities: i) MHI and MFI are highly and positively correlated with SEA and SES ($0.80 \leq r \leq 0.95$), ii) MHI and MFI are highly, but negatively correlated with SEP, SED, and DEP ($-0.82 \leq r \leq -0.94$), except for the relationship between MHI and DEP in San Diego, CA where $r = -0.78$ (Figure 3), and iii) scatterplots of MHI and MFI against the five composite measures indicate rather non-linear and quadratic relationships, whereas $MHI^{(\lambda)}$ and $MFI^{(\lambda)}$ against the five composite measures indicate linear relationships: however, regardless of four single measures following skewed or normal distributions

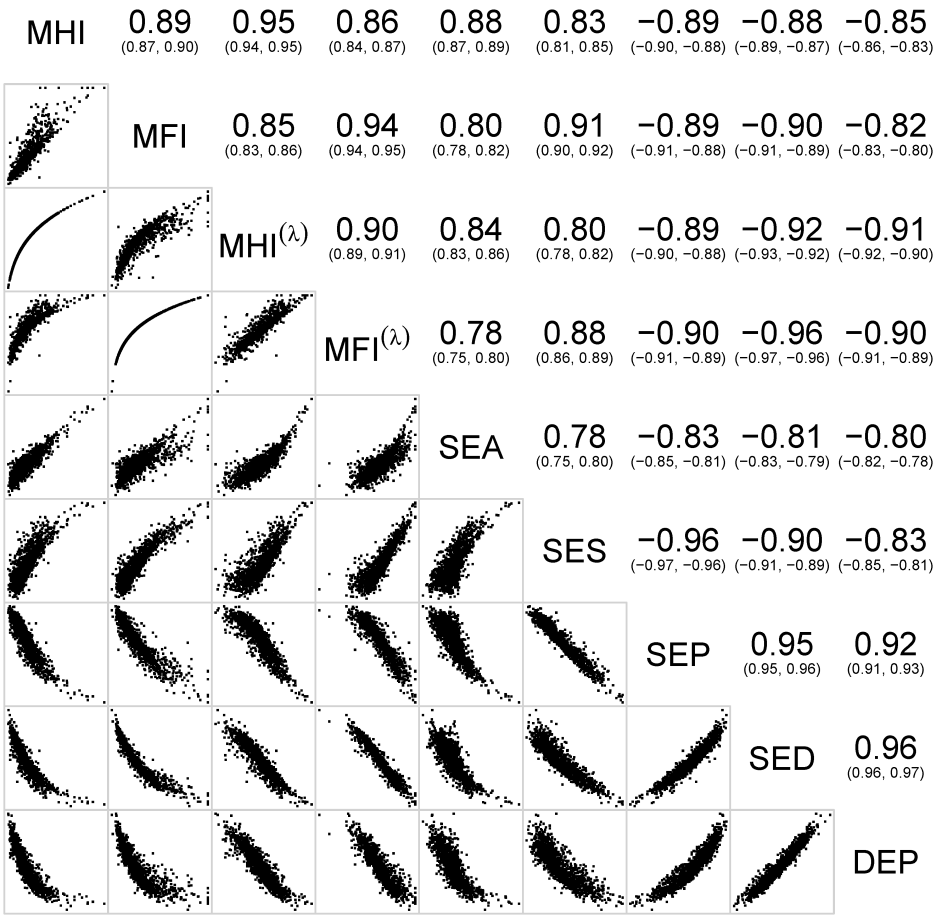
(in respect to their original or transformed forms, respectively), the high degree of correlations between composite measures remained unchanged.

**Figure 1: Correlations and scatterplot matrices of selected area-based measures in St. Louis, MO (329 census tracts)**



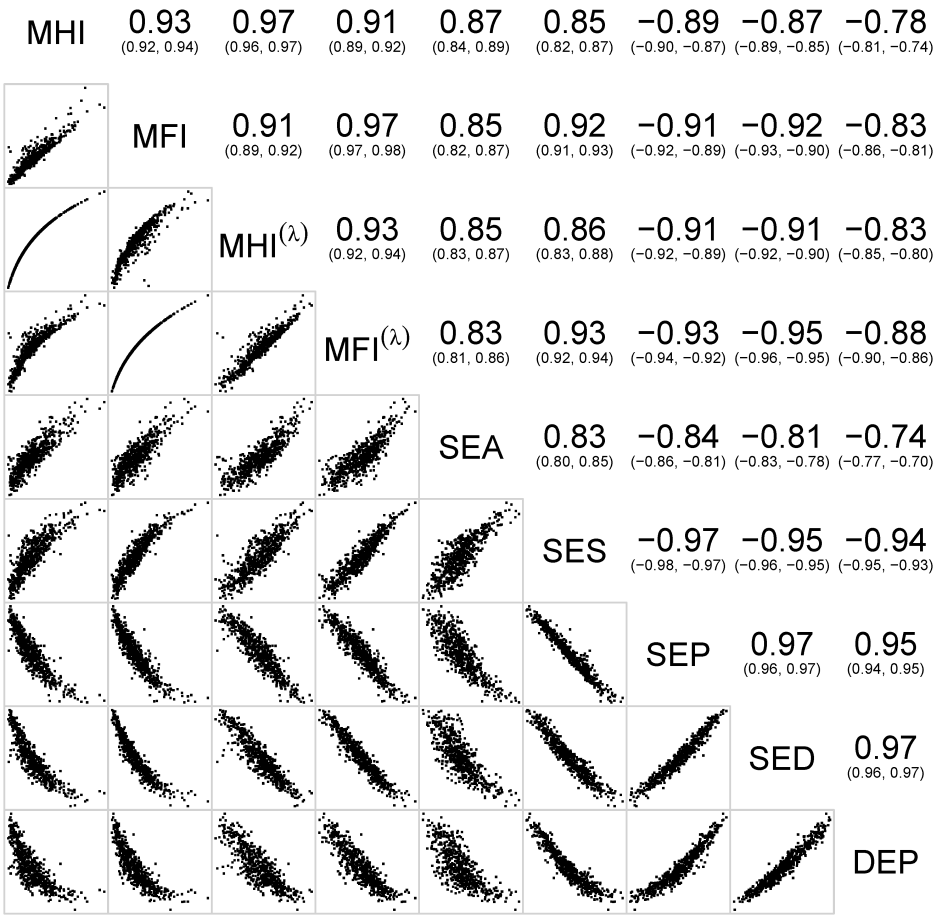

*Notes*: MHI: median household income (US $), MFI: median family income (US $), $MHI^{(\lambda)}$: Box-Cox transformed median households income, $MFI^{(\lambda)}$: Box-Cox transformed median family income, SEA: socioeconomic advantage (Diez Roux et al. 2001), SES: socioeconomic status (Winkleby and Cubbin 2003), SEP: socioeconomic position (Krieger et al. 2003), SED: socioeconomic deprivation (Singh 2003), and DEP: deprivation (Messer et al. 2006).

**Figure 2: Correlations and scatterplot matrices of selected area-based measures in Chicago, IL (1,222 census tracts)**



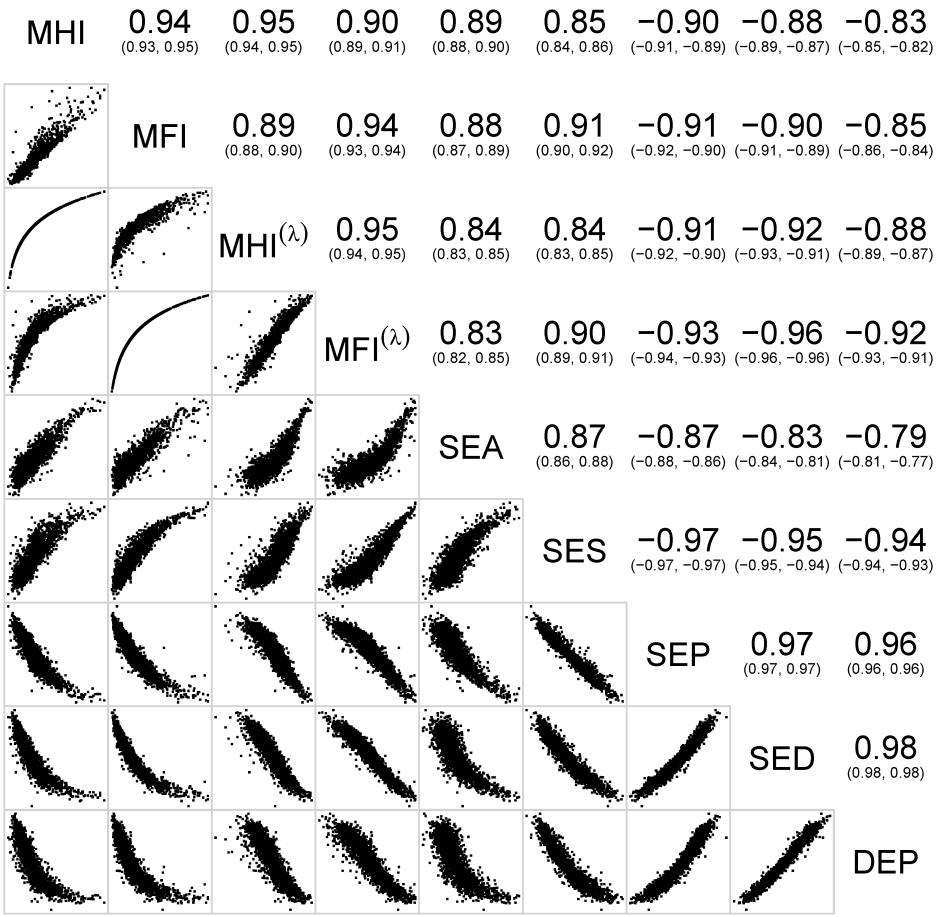

*Notes*: MHI: median household income (US $), MFI: median family income (US $), $MHI^{(\lambda)}$: Box-Cox transformed median households income, $MFI^{(\lambda)}$: Box-Cox transformed median family income, SEA: socioeconomic advantage (Diez Roux et al. 2001), SES: socioeconomic status (Winkleby and Cubbin 2003), SEP: socioeconomic position (Krieger et al. 2003), SED: socioeconomic deprivation (Singh 2003), and DEP: deprivation (Messer et al. 2006).

**Figure 3: Correlations and scatterplot matrices of selected area-based measures in San Diego, CA (592 census tracts)**




*Notes*: MHI: median household income (US $), MFI: median family income (US $), $MHI^{(\lambda)}$: Box-Cox transformed median households income, $MFI^{(\lambda)}$: Box-Cox transformed median family income, SEA: socioeconomic advantage (Diez Roux et al. 2001), SES: socioeconomic status (Winkleby and Cubbin 2003), SEP: socioeconomic position (Krieger et al. 2003), SED: socioeconomic deprivation (Singh 2003), and DEP: deprivation (Messer et al. 2006).

**Figure 4: Correlations and scatterplot matrices of selected area-based measures in Los Angeles, CA (1,965 census tracts)**




*Notes*: MHI: median household income (US $), MFI: median family income (US $), $MHI^{(\lambda)}$: Box-Cox transformed median households income, $MFI^{(\lambda)}$: Box-Cox transformed median family income, SEA: socioeconomic advantage (Diez Roux et al. 2001), SES: socioeconomic status (Winkleby and Cubbin 2003), SEP: socioeconomic position (Krieger et al. 2003), SED: socioeconomic deprivation (Singh 2003), and DEP: deprivation (Messer et al. 2006).

## 4. Discussion

Despite the differences in how selected area-based measures of neighborhood socioeconomic characteristics were derived from the 2005–09 ACS data, Figures 1 through 4 suggest that area-based median household and median family income (i.e., MHI and MFI) capture the same dimension of neighborhood context as do composite measures of neighborhood socioeconomic characteristics (i.e., SEA, SES, SEP, SED, and DEP). In other words, they all capture the neighborhood affluence-deprivation (or, its counterpart, deprivation-affluence) continuum. This, in turn, highlights two important remarks in health research.

First, an area-based median household (or family) income may be used in future studies instead of a composite measure of neighborhood socioeconomic characteristics. Since selected area-based measures were highly correlated (either positively or negatively) with one another (Figures 1 through 4), they are indicative of having the same geographic distribution (or the inverse thereof). From an analytical point of view, incorporating either one of them into a regression analysis would convey the same information (e.g., similar regression estimates and standard errors). In fact, Morris and Carstairs (1991) demonstrated this more than two decades ago by comparing highly correlated area-based composite measures of deprivation (not included in this study) developed in the United Kingdom (UK). However, neither the area-based median household income nor the area-based median family income was included in their analysis. This is mainly because such socioeconomic data have not been routinely collected in the UK Census; a composite measure, often derived from the Townsend index (Townsend 1987) or the Carstairs index (Carstairs and Morris 1989) has been widely used in UK studies as a proxy for area-based income or wealth data (Morgan and Baker 2006). For this reason, it may be important to recognize that the availability of area-based socioeconomic data varies from country to country. Indeed, both area-based median household income and median family income are readily available at the census tract level in the US Census (recently from the ACS). Therefore, using either one would be an appropriate approach in US studies as long as the high degree of correlations between three to five composite measures (not limited to the ones considered in this study) can be validated in the preliminary analysis.

Second, an area-based median household (or family) income may provide a better representation of neighborhood context than a composite measure of neighborhood socioeconomic characteristics does. From a theoretical point of view, the use of a composite measure has been favored in health research due to its conceptually comprehensive portrayal of neighborhood socioeconomic characteristics (Diez Roux et al. 2001; Krieger et al. 2003; Singh 2003; Winkleby and Cubbin 2003; Messer et al. 2006). However, a driving objective behind the development and construction of

composite measures is the representation of a concept that is too complex to be represented by a single measure; multivariate techniques (e.g., factor analysis and principal component analysis) have been commonly used to gain a well-rounded perspective on a concept in various fields of study (Hair Jr. et al. 2009), not limited to health research. As shown in Figures 1 through 4, such a conceptual expectation does not hold, and thus composite measures may not be superior to single measures. From a methodological point of view, socioeconomic characteristics at both the individual and neighborhood levels need to be measured with as much relevant and specific information (not overall) as possible; importantly, such measures need to distinguish the key domains (e.g., income, wealth, poverty, educational achievement, and occupation) at different levels (Krieger, Williams, and Moss 1997; Braveman et al. 2005). By definition, composite measures (e.g., SEA, SES, SEP, SED, and DEP) are an overall characterization of a neighborhood. Therefore, an area-based median household (or family) income may capture the area-based variations in health along a neighborhood affluence-deprivation continuum in a coherent and consistent manner.

As a technical note, area-based median household income and median family income (i.e., MHI and MFI) may be reversed and rescaled before being incorporated into a regression analysis as a neighborhood covariate. Since the increases in MHI and MFI conform to the increase in neighborhood affluence, reversing them (i.e., multiplying by -1) correspondingly conforms to the increase in neighborhood deprivation. Then, the following approaches each provide a simple but common trick in applied statistics to ease the interpretation of regression estimates: i) reversed MHI and MFI may be divided by its range so that the values are bounded between 0 (zero) and 1 (one), or ii) reversed MHI and MFI may be divided by its interquartile range (IQR). The former approach offers an interpretable comparison of the changes from low or minimum (i.e., zero) to high or maximum (i.e., one) values, whereas the latter approach offers the difference between the upper and lower quartiles (i.e., the third quartile minus the first quartile). Both approaches do not alter or influence the regression analysis but rather ease the interpretation of regression estimates in a standardized manner.

In order to ensure the usefulness of using area-based median household (or family) income as the measure of a neighborhood affluence-deprivation continuum, however, further studies are needed for cities of different population sizes in different geographic locations (e.g., suburban and rural settings) and regions (e.g., Eastern and Southern states) across the US. Moreover, another important area of study (which is beyond the scope of this study) is to compare various area-based measures of county- and metropolitan-level socioeconomic characteristics used in US studies. Once a reliable single or composite measure is identified at those levels, area-based measures of socioeconomic characteristics at the neighborhood- and county-levels, neighborhood- and metropolitan-levels, or neighborhood-, county- and metropolitan-levels can be

incorporated into a multilevel regression analysis. Because different socioeconomic factors may shape health at different levels (e.g., individuals, neighborhoods, and counties or metropolitan areas) and through different causal pathways (Braveman et al. 2005), such efforts contribute to a better understanding of the hierarchical structure of contextual effects on health in different types of human settlement.

The practical arguments discussed so far may be useful in promoting informative research and in interpreting research findings. However, two limitations warrant mention. First, many area-based measures of neighborhood socioeconomic characteristics have been developed (or modified) in health research. However, not all area-based measures are alike. For example, the composite measure of concentrated disadvantage (Sampson, Raudenbush, and Earls 1997) as well as unemployment and below poverty rates were not correlated with MHI and MFI or did not show consistent patterns in four US cities (data not shown). Therefore, further comparisons of area-based measures are needed in future studies. Second, the US Census no longer gathers detailed information on households through the so-called long form; after 2000 the ACS replaced the long form. Due to the design of ACS (Herman 2008), its estimates for census tracts (and block groups) may not be as reliable (with a relatively large margin of errors) as the past decennial censuses. However, how the quality of ACS estimates may affect the computations of composite measures of neighborhood socioeconomic characteristics has not been investigated.

## 5. Conclusion

A tension between theory and application may arise in applied statistics (Cox 1995; Chatfield 2002). As discussed above, the same analogy may be used to infer a potential conflict of interest in the use of composite measure (e.g., Diez Roux et al. 2001; Krieger et al. 2003; Singh 2003; Winkleby and Cubbin 2003; Messer et al. 2006) versus single measure (e.g., Galea et al. 2007; Stockdale et al. 2007; Tonorezos et al. 2008; Sallis et al. 2009; Black and Macinko 2010; King et al. 2011) of neighborhood socioeconomic characteristics in previous studies. Figures 1 through 4 suggest that an area-based median household (or family) income in a regression analysis would convey the same information as that of selected composite measures considered in this study, despite the conceptual and theoretical differences.

Place of residence may play a role in the persistence and/or widening of health disparities (Robinson 2005; Bernard et al. 2007; Diez Roux and Mair 2010). To this end, an important component in health research is to identify the risk factors that vary across affluent and deprived neighborhoods. Hence, an area-based median household

(or family) income may be used as a measure of a neighborhood affluence-deprivation continuum within US cities in such future research endeavors.

## 6. Acknowledgements

This work was supported by the National Cancer Institute (NCI) Centers for Transdisciplinary Research on Energetics and Cancer (U54CA155496). The NCI had no further role in the study design; in the collection, analysis and interpretation of data; in the writing of the report; or in the decision to submit the paper for publication. The opinions or assertions contained herein are the private one of the author, and are not considered as official or reflecting the views of the NCI.

## References

Bernard, P., Charafeddine, R., Frohlich, K.L., Daniel, M., Kestens, Y., and Potvin, L. (2007). Health inequalities and place: A theoretical conception of neighbourhood. *Social Science & Medicine* 65(9): 1839–1852. doi:10.1016/j.socscimed.2007.05.037.

Black, J.L. and Macinko, J. (2008). Neighborhoods and obesity. *Nutrition Reviews* 66(1): 2–20. doi:10.1111/j.1753-4887.2007.00001.x.

Black, J.L. and Macinko, J. (2010). The changing distribution and determinants of obesity in the neighborhoods of New York City, 2003–2007. *American Journal of Epidemiology* 171(7): 765–775. doi:10.1093/aje/kwp458.

Box, G.E. and Cox, D.R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society. Series B* 26(2): 211–252.

Braveman, P.A., Cubbin, C., Egerter, S., Chideya, S., Marchi, K.S., Metzler, M., and Posner, S. (2005). Socioeconomic status in health research: One size does not fit all. *Journal of the American Medical Association* 294(22): 2879–2888. doi:10.1001/jama.294.22.2879.

Carstairs, V. and Morris, R. (1989). Deprivation, mortality and resource allocation. *Community Medicine* 11(4): 364–372.

Chaix, B. (2009). Geographic life environments and coronary heart disease: A literature review, theoretical contributions, methodological updates, and a research agenda. *Annual Review of Public Health* 30: 81–105. doi:10.1146/annurev.publhealth.031308.100158.

Chatfield, C. (2002). Confessions of a pragmatic statistician. *Journal of the Royal Statistical Society. Series D* 51(1): 1–20. doi:10.1111/1467-9884.00294.

Cox, D.R. (1995). The relation between theory and application in statistics. *Test* 4(2): 207–261. doi:10.1007/BF02562627.

Diez Roux, A.V., Kiefe, C.I., Jacobs, D.R.J., Haan, M., Jackson, S.A., Nieto, F.J., Paton, C.C., and Schulz, R. (2001). Area characteristics and individual-level socioeconomic position indicators in three population-based epidemiologic studies. *Annals of Epidemiology* 11(6): 395–405. doi:10.1016/S1047-2797(01)00221-6.

Diez Roux, A.V. and Mair, C. (2010). Neighborhoods and health. *Annals of the New York Academy of Sciences* 1186: 125–145. doi:10.1111/j.1749-6632.2009.05333.x.

Fox, J. and Weisberg, S. (2011). *An R companion to applied regression*. Thousand Oaks, CA: SAGE Publications.

Friendly, M. (2002). Corrgrams: Exploratory display for correlation matrices. *The American Statistician* 56(4): 316–324. doi:10.1198/000313002533.

Galea, S., Ahern, J., Tracy, M., and Vlahov, D. (2007). Neighborhood income and income distribution and the use of cigarettes, alcohol, and marijuana. *American Journal of Preventive Medicine* 32(6S): S195–S202. doi:10.1016/j.amepre.2007.04.003.

Hair Jr., J.F., Black, W.C., Babin, B.J., and Anderson, R.E. (2009). *Multivariate data analysis*. Upper Saddle River, NJ: Prentice Hall.

Herman, E. (2008). The American community survey: An introduction to the basics. *Government Information Quarterly* 25(3): 504–519. doi:10.1016/j.giq.2007.08.006.

Kawachi, I. and Berkman, L.F. (2003). *Neighborhoods and health*. New York, NY: Oxford University Press. doi:10.1093/acprof:oso/9780195138382.001.0001.

Kim, D. (2008). Blues from the neighborhood? Neighborhood characteristics and depression. *Epidemiologic Reviews* 30(1): 101–117. doi:10.1093/epirev/mxn009.

King, A.C., Sallis, J.F., Frank, L.D., Saelens, B.E., Cain, K., Conway, T.L., Chapman, J.E., Ahn, D.K., and Kerr, J. (2011). Aging in neighborhoods differing in walkability and income: Associations with physical activity and obesity in older adults. *Social Science & Medicine* 73(10): 1525–1533. doi:10.1016/j.socscimed.2011.08.032.

Krieger, N. (2006). A century of census tracts: Health & the body politic (1906–2006). *Journal of Urban Health* 83(3): 355–361. doi:10.1007/s11524-006-9040-y.

Krieger, N., Chen, J.T., Waterman, P.D., Soobader, M.-J., Subramanian, S.V., and Carson, R. (2003). Choosing area based socioeconomic measures to monitor social inequalities in low birth weight and childhood lead poisoning: The public health disparities geocoding project (US). *Journal of Epidemiology and Community Health* 57(3): 186–199. doi:10.1136/jech.57.3.186.

Krieger, N., Williams, D.R., and Moss, N.E. (1997). Measuring social class in US public health research: Concepts, methodologies, and guidelines. *Annual Review of Public Health* 18: 341–378. doi:10.1146/annurev.publhealth.18.1.341.

Mair, C., Diez Roux, A.V., and Galea, S. (2008). Are neighbourhood characteristics associated with depressive symptoms? A review of evidence. *Journal of Epidemiology and Community Health* 62(11): 940–946. doi:10.1136/jech.2007.066605.

Meijer, M., Röhl, J., Bloomfield, K., and Grittner, U. (2012). Do neighborhoods affect individual mortality? A systematic review and meta-analysis of multilevel studies. *Social Science & Medicine* 74(8): 1204–1212. doi:10.1016/j.socscimed.2011.11.034.

Messer, L.C., Laraia, B.A., Kaufman, J.S., Eyster, J., Holzman, C., Culhane, J., Elo, I., Burke, J.G., and O'Campo, P. (2006). The development of a standardized neighborhood deprivation index. *Journal of Urban Health* 83(6): 1041–1062. doi:10.1007/s11524-006-9094-x

Morgan, O. and Baker, A. (2006). Measuring deprivation in England and Wales using 2001 Carstairs scores. *Health Statistics Quarterly / Office for National Statistics* Autumn(31): 28–33.

Morris, R. and Carstairs, V. (1991). Which deprivation? A comparison of selected deprivation indexes. *Journal of Public Health Medicine* 13(4): 318–326.

Osborne, J.W. (2010). Improving your data transformations: Applying the Box-Cox transformation. *Practical Assessment, Research & Evaluation* 15(12).

Pickett, K.E. and Pearl, M. (2001). Multilevel analyses of neighborhood socioeconomic context and health outcomes: A critical review. *Journal of Epidemiology and Community Health* 55(2): 111–122. doi:10.1136/jech.55.2.111.

R Core Team (2013). R: A language and environment for statistical computing [computer program]. Vienna, Austria: R Foundation for Statistical Computing. [http://www.R-project.org]

Riva, M., Gauvin, L., and Barnett, T.A. (2007). Toward the next generation of research into small area effects on health: A synthesis of multilevel investigations published since July 1998. *Journal of Epidemiology and Community Health* 61(10): 853–861. doi:10.1136/jech.2006.050740.

Robinson, R.G. (2005). Community development model for public health applications: Overview of a model to eliminate population disparities. *Health Promotion Practice* 6(3): 338–346. doi:10.1177/1524839905276036.

Sallis, J.F., Saelens, B.E., Frank, L.D., Conway, T.L., Slymen, D.J., Cain, K.L., Chapman, J.E., and Kerr, J. (2009). Neighborhood built environment and income: Examining multiple health outcomes. *Social Science & Medicine* 68(7): 1285–1293. doi:10.1016/j.socscimed.2009.01.017.

Sampson, R.J., Raudenbush, S.W., and Earls, F. (1997). Neighborhoods and violent crime: A multilevel study of collective efficacy. *Science* 277(5328): 918–924. doi:10.1126/science.277.5328.918.

Singh, G.K. (2003). Area deprivation and widening inequalities in US mortality, 1969–1998. *American Journal of Public Health* 93(7): 1137–1143. doi:10.2105/AJPH.93.7.1137.

Stockdale, S.E., Wells, K.B., Tang, L., Belin, T.R., Zhang, L., and Sherbourne, C.D. (2007). The importance of social context: Neighborhood stressors, stress-buffering mechanisms, and alcohol, drug, and mental health disorders. *Social Science & Medicine* 65(9): 1867–1881. doi:10.1016/j.socscimed.2007.05.045.

Tonorezos, E.S., Karpati, A.M., Wang, Y., and Barr, R.G. (2008). Does the relationship between asthma and obesity differ by neighborhood? *Respiratory Medicine* 102(12): 1797–1804. doi:10.1016/j.rmed.2008.06.018.

Townsend, P. (1987). Deprivation. *Journal of Social Policy* 16(2): 125–146. doi:10.1017/S0047279400020341.

US Census Bureau. (2013). Geographic areas – definitions [electronic resource]. Washington, DC: American Community Survey Office. [http://www.census.gov/acs/www/data_documentation/custom_tabulation_request_form/geo_def.php]

Vlahov, D. and Galea, S. (2002). Urbanization, urbanicity, and health. *Journal of Urban Health* 79(4): S1–S12. doi:10.1017/S0047279400020341.

Winkleby, M.A. and Cubbin, C. (2003). Influence of individual and neighbourhood socioeconomic status on mortality among black, Mexican-American, and white women and men in the United States. *Journal of Epidemiology and Community Health* 57(6): 444–452. doi:10.1136/jech.57.6.444.

Wright, K. (2012). Corrgram: Plot a correlogram [computer program]. R package version 1.4. [http://CRAN.R-project.org/package=corrgram]