Journal of Urban Health: Bulletin of the New York Academy of Medicine, Vol. 93, No. 1
doi:10.1007/s11524-015-9959-y

Constructing a Time-Invariant Measure of the Socio-economic Status of U.S. Census Tracts

Jeremy N. Miles, Margaret M. Weden, Diana Lavery,
José J. Escarce, Kathleen A. Cagney, and Regina A. Shih

ABSTRACT *Contextual research on time and place requires a consistent measurement instrument for neighborhood conditions in order to make unbiased inferences about neighborhood change. We develop such a time-invariant measure of neighborhood socio-economic status (NSES) using exploratory and confirmatory factor analyses fit to census data at the tract level from the 1990 and 2000 U.S. Censuses and the 2008–2012 American Community Survey. A single factor model fit the data well at all three time periods, and factor loadings—but not indicator intercepts—could be constrained to equality over time without decrement to fit. After addressing remaining longitudinal measurement bias, we found that NSES increased from 1990 to 2000, and then—consistent with the timing of the "Great Recession"—declined in 2008–2012 to a level approaching that of 1990. Our approach for evaluating and adjusting for time-invariance is not only instructive for studies of NSES but also more generally for longitudinal studies in which the variable of interest is a latent construct.*

KEYWORDS *Neighborhood socio-economic status, Neighborhood disadvantage, Neighborhood change, Confirmatory factor analysis, Measurement bias, Invariance*

A well-established literature documents that the socio-economic characteristics of the places in which we live influence our health and wellbeing.[1–7] For example, neighborhood socio-economic status (NSES), over and above individual socio-economic status,[2] can have lasting effects on outcomes ranging from hypertension,[8] to allostatic load,[9] disability,[10] and depression.[11] Reviews of research on neighborhoods and health have suggested we need to better understand the role of critical periods, sequencing, and the accumulation of (dis)advantages over time.[6,12] Longitudinal studies hoping to address these questions, however, must first address the methodological challenge of appropriately measuring neighborhood characteristics over time: in particular, distinguishing between changes in the *consequences* of a neighborhood construct over time and changes in the *measurement* of the construct over time. The objective of this study is to address these challenges of incorporating time into the study of place by developing a measure of NSES and testing the stability of its measurement (time-invariance) from 1990 through about 2010. By so doing, we intend to not only produce a measure of the NSES of U.S. census tracts that can be used in longitudinal research and surveillance, but also

Miles, Weden, Lavery, Escarce, and Shih are with the RAND Corporation, Santa Monica, CA, USA; Escarce is with the University of California, Los Angeles, CA, USA; Cagney is with the University of Chicago, Chicago, IL, USA.

Correspondence: Margaret M. Weden, RAND Corporation, Santa Monica, CA, USA. (E-mail: mweden@rand.org)

Jeremy N. Miles and Margaret M. Weden contributed equally to this work.

elucidate whether and how indicators of NSES may have changed over the last several decades.

WHAT IS NSES AND HOW HAS IT BEEN MEASURED?

Research on neighborhood socio-economic conditions has its origins in Early Chicago School social theory.[1,6,13–15] Whether labeled "neighborhood disadvantage," "neighborhood affluence and disadvantage," or more broadly "neighborhood socio-economic status" (NSES), these studies have similarly employed factor analyses to describe the social and economic characteristics of U.S. census tracts, typically using decennial census data.[8–10,16–26] Most commonly, a single factor is retained that is measured by all or a subset of the following indicators:[*] level of income, poverty, unemployment, public assistance, female-headed households, educational attainment, and employment in professional or managerial positions. However, these factor analyses have typically been carried out for a single measurement occasion. Thus, it is unknown and may not necessarily be the case that a single factor analytic solution for estimating NSES will be stable over time—i.e., that the factor structure, factor loadings, and item intercepts will be equivalent over time.

At least a decade of reviews highlight the need for longitudinal research on neighborhoods and individuals' health and wellbeing.[6,7,12] However, nearly all known studies have employed static models of the neighborhood in which neighborhood conditions are measured at one point in time. In two known longitudinal studies with dynamic measures of NSES,[27,28] the assumption is made that a measure of neighborhood socio-economic conditions can be similarly estimated using three decades of decennial census data. This assumption of factorial invariance is a standard approach for dealing with potential changes in the measurement of a construct over time or age or across groups.[29] However, as in these two studies, the assumption of invariance is rarely tested. An unfortunate consequence is that comparisons over time may be biased and lead to false conclusions if invariance does not hold.[30]

WHAT IS FACTORIAL INVARIANCE AND WHY DOES IT MATTER?

Factorial invariance refers to the equivalence of factor structures that are used to define latent variables. A latent variable is a "variable for which there is no sample realization"—that is, a variable that is not directly measured but instead is hypothesized as an explanation for relationships between measured variables.[31] Invariance can be tested across groups (i.e., group-invariance),[32–35] or, as in this study, across time (i.e., time-invariance).[3,36–39] Invariance is tested in a sequence of increasingly restrictive models.[40] The first level of invariance is structural and evaluates whether the structure, or number of factors, are the same over time. If structural invariance fails to hold, the dimensionality of the latent variable has changed and comparisons will not be meaningful. However if structural invariance

[*]Although some single factor models also include measures of census tract composition by race/ethnicity and nativity or even English language ability,[19,21] other studies have shown that models including these variables are best fit with a multifactorial structure, with a separate but correlated factor for the presence of racial/ethnic segregation and/or immigrant enclaves.[8,10,21,27]

holds, weak invariance can be assessed by evaluating whether the factor loadings (or relationship between the indicators and the factor) are the same over time. If factor loadings change, then any potential changes in the level of the latent variable will not be reflected appropriately by changes in the measured variables. However, if weak invariance holds, then strong invariance can be tested by evaluating whether the intercepts (or the score) of an indicator variable is the same for the same level of the latent construct at each point in time.[†] Analogous to a regression model, the intercept of the indicator variables represent the expected score on the indicator when the latent variable is equal to zero. If the intercepts of the indicator variables remain constant (along with the factor loadings), as the mean score of the latent construct increases or decreases from a referent time point, then the measurement model meets strong invariance. However, as above, if strong invariance fails to hold and the assumption is made that the measure is time-invariant, then estimates of the latent construct will be biased and not comparable over time. Uncorrected, the resulting implication of a researcher applying a measure that fails the test of strong invariance is that a change in the level of construct will be interpreted from what is in actuality a change in the scaling of one (or more) of the indicators of the construct.

Assessment of these potential forms of measurement instability and adjustment for any form of time-invariance is critical in producing a measure of NSES—or any other latent measure—for unbiased assessment of longitudinal research questions. In addition, the process of evaluating and constructing a time-invariant measure allows us, and other researchers, to develop new insights about potential stability and change in the indicators and measurement of the construct over time.

METHODS

Data

U.S. Census Bureau data on the social and economic characteristics of U.S. census tracts are obtained from the 1990 and 2000 decennial census long-forms[41,42] and the 2008–2012 American Community Survey (ACS).[43] The ACS, like the decennial census long-form it has replaced, is a national survey of all U.S. housing units and group living quarters. Due to the smaller but more frequent sampling conducted by the ACS, however, data for geographic areas as small as census tracts are released in 5-year estimates, beginning with 2005–2009. We select the 2008–2012 ACS because the centroid is ten years after the 2000 census. Census tracts were employed as the proxy for neighborhood with the study sample restricted to those observed at all three assessments in 1990, 2000, and 2008–2012 (N=65,174). U.S Census Bureau data for the 1990 Census, 2000 Census, and 2008–2012 ACS use, respectively, the 1990, 2000, and 2010 census tract boundary definitions. These data were harmonized to the 2000 census tract boundaries using a transformation matrix we calculated from the Longitudinal Tract Data Base (LTDB[44]).[‡]

†In factor analysis, the measured variables are considered to be outcome variables with their scores predicted by the latent construct; hence the intercept is the expected value of the measured variable when the latent construct takes the value of zero.

‡The LTDB provides transformation coefficients and a tract correspondence matrix for harmonizing 2000 geographic boundaries to 2010 geographic boundaries. The methodology is similar to an earlier harmonization method developed for harmonizing 1990 boundaries to 2000 boundaries.[45] We were able to use these data to produce transformation coefficients for (reverse) harmonizing from 2010 to 2000.

Measures

Nine indicators of NSES were considered:

- Median household income;
- Proportion of households with income below the federal poverty level;
- Level of education;
- Proportion of the population age 16 years or older that is unemployed;
- Proportion of civilian workers age 16 years or older in management, professional, and related occupations;
- Proportion of households that receive public assistance income;
- Proportion of female-headed households (i.e., no husband present) with children under age 18 years;
- Proportion of households with crowded housing (i.e., more than one occupant per room); and
- Median value of owner-occupied housing units.

Median household income and median housing value are reported in dollar values that we adjusted for inflation using a 2000 referent.[46] The measure of the level of education is calculated from the proportion of the population age 25 years and older that reports educational attainment: (a) less than a high school diploma or general education development (GED) equivalent; (b) a high school diploma or GED equivalent, but not a bachelor's degree; and (c) a bachelor's degree or higher.[§] The reference periods for the reporting of household income and any income from public assistance,[||] as well as employment status, differ between the decennial censuses and the ACS.[47,48] The Census Bureau suggests that these variables (and variables, such as poverty status, derived from them) can be compared between the decennial censuses and ACS, albeit with caution.[49] In addition, the proportion of the census tract employed in management, professional and related occupations is obtained from the estimates for detailed occupational categories reported in the decennial censuses and ACS. These detailed occupational categories are based on the U.S. Bureau of Labor Statistics six-digit Standard Occupation Code (SOC) system for 1990, 2000, and 2010. Although reporting categories changed between 1990 and 2000 and again in 2010 on the basis of changes in the SOC, these changes largely entailed those in subcategories below the top level of SOC reporting categories denoting management, professional, and related occupations.[50,51]

Data Preparation and Statistical Analysis

We first examined the distributions of the variables and carried out transformations so that skewed variables more closely approximated a normal distribution. After examining variable distributions, a natural logarithm transformation was applied (adding a constant, where necessary). In addition, because the estimation algorithms

---

[§]The rescaled level of education=$0\times(a)+1\times(b)+2\times(c)$.

[||]The reference period for questions about income and sources of income in 1990 and 2000 Censuses is the last calendar year, while estimates from the ACS 2008–2012 refer to income in the last 12 months and come from respondents surveyed between 2008 and 2012.

The reference period for all surveys was employment in the last week; however, as noted above, ACS estimates come from respondents who may have been interviewed at any of the year-round survey dates between 2008 and 2012.

used in the confirmatory factor analysis models (described below) may not converge when variances differ across indicator variables by more than an order of magnitude, we multiplied indicators by a constant to ensure that variances were broadly similar.

Estimation of a latent construct describing NSES was carried out using exploratory and confirmatory factor analysis with Mplus version 7.11.[52] Models were estimated using the maximum likelihood (robust) algorithm which provides estimates of model fit that are equivalent to maximum likelihood, with corrections to standard errors and model fit to account for non-normality in the data.[53–55]

The data analysis procedure for testing time-invariance typically starts with a measure or measurement instrument that has a known dimensionality. Because the dimensionality of NSES is unknown, our first analysis combines the testing of dimensionality and configural invariance. Based on our review of the previous NSES literature, we fitted a confirmatory factor analysis (CFA) model with a single factor at each time using the nine indicators described above. Unique variances of the equivalent variables were correlated across time. However, the fit of this CFA model (model A1) was poor. Given that our initially specified model had failed we therefore proceeded with exploratory factor analysis to develop an initial model, using maximum likelihood extraction and *geomin* rotation.[56] Using this exploratory factor analysis approach, we identified five NSES indicator variables which appeared to fit a single factor structure at each assessment. We then tested for time-invariance of the NSES measurement instrument defined using these indicators within a confirmatory factor analysis framework. A confirmatory factor analysis approach for modeling NSES longitudinally involves estimating all of the following at each time point: (a) a matrix of factor loadings for the relationship between each indicator variable and the latent construct; (b) a vector of intercepts for each indicator variable; (c) a vector of means of the latent variables, (d) a matrix of variances and covariances of the latent variables; and (e) a matrix of residual variances and covariances for the indicator variables.

Several parameter constraints are required to statistically identify the longitudinal model for NSES. To identify the mean and variance of the latent variable at one point in time, it is necessary to identify a reference variable and constrain the loading and intercept of this parameter at each time point. The reference variable can be thought of as 'anchoring' the latent variable and is typically chosen to be the indicator variable believed to be most closely related to the latent construct.[31] We selected median household income as the reference variable and constrained the loadings of this variable to 1.00. Similarly, in order to statistically identify potential change in the latent construct, it was necessary to anchor the mean of the latent variable in time. We selected the first time point (1990) as the referent time point for the latent variable and constrained the mean of the latent variable to be zero at this point. Again, any value could have been employed, but zero was most convenient for model interpretation because it allowed us to observe change in the mean of the hypothesized latent variable (NSES) over time, relative to the first time point once a time-invariant model of NSES was identified. To identify the mean of the latent variables at the later time points, we constrain the intercept of the reference variable to be equal at each time point. By selecting the above parameter constraints to achieve statistical identification we also ensured that the coefficients for the intercepts of the indicator variables could be readily interpreted and evaluated for time-invariance. The intercept of the indicators are interpreted, as in regression, as the expected value of the variable when the mean of the latent variable (which functions as a predictor here) is equal to zero. If the latent variable changes over time, and all of the indicator variables change in a consistent fashion, then the intercepts of the indicator variables will be constant over time.

We employed conventional statistical methods to test for configural, weak, and strong time invariance.[40,57] Weak invariance was tested by comparing a model in which the set of factor loadings for each respective indicator were constrained to equality over time (model B3) versus an otherwise equivalent model (model B2). Strong invariance was tested by comparing a model in which the set of intercept coefficients for each respective indicator were additionally constrained to equality over time (model B4) versus the unconstrained model (model B3).

Given the large sample size, we expected chi-square tests to be overpowered and lead us to reject models based on small discrepancies. Therefore, the RMSEA (root mean square error of approximation[58]) and CFI (comparative fit index[59]) were better choices for evaluating global model fit. We employed a threshold value of 0.06 to indicate adequate fit for RMSEA and 0.95 for CFI;[60] we also evaluated aspects of local model fit (e.g., residuals, modification indices and standardized expected parameter changes). In addition, we use a ΔCFI of greater than 0.010 to indicate a meaningful reduction in model fit when constraints were added.[61] We use modification indices in conjunction with standardized expected parameter changes to guide model modification.[62]

RESULTS

Table 1 shows means and standard deviations for each of the indicator variables prior to and post transformation in 1990, 2000, and 2008–2010.

The process of evaluating the time-invariance involved a series of models whose goodness of fit statistics are detailed in Table 2. The first step was to assess configural invariance and dimensionality of NSES over the three assessments. A single factor model (model A1) failed to adequately account for the data (RMSEA=0.062 and CFI=0.875). We therefore conducted EFA and identified a unidimensional subset of the variables: median household income, educational level, proportion unemployed, proportion below the poverty level, and proportion of female-headed households.

Having identified a unidimensional set of items that appeared to measure the same construct over time, we next proceeded to formally test the time-invariance of these items. We refer to this sequence of CFA models as model series B (See Table 2). The first model in the series (model B1) tested the configural invariance of a single factor model measured with the five indicator variables. It fit the data adequately (RMSEA=0.062 and CFI=0.966). Examination of the modification indices and standardized parameter estimates suggested that a model that incorporated a correlation in the residual variance of the variables median household income and proportion below the poverty level would fit better, and indeed the fit statistics improved considerably for model B2 with this added parameter.

Our next step was to test for weak factorial invariance by constraining factor loadings to equality across all three assessments (model B3). Although the fit worsened when this constraint was added, the decrement to CFI was negligible (ΔCFI=−0.010, Table 2), and thus allowed us to retain the hypothesis of longitudinal invariance of the factor loadings. As further (qualitative) evidence of weak time invariance, consider the stability over time of the factor loadings for model B2 in which these parameters were allowed to be freely estimated and their

TABLE 1 Descriptive statistics for variables describing neighborhood socioeconomic characteristics of U.S. census tracts by assessment year(s), pre- and post-transformation (N=65,174)

	Pre-transformation			Post- transformation	
Variable and year(s)	Mean	SD	Transformation of variable (Y)	Mean	SD
Household income (median in dollars[a])					
1990	39,048.46	18,122.98	$\ln(10+Y)$	10.47	0.46
2000	44,249.33	20,760.14		10.60	0.45
2008–2012	40,128.05	19,805.98		10.49	0.47
Poverty (proportion)					
1990	0.13	0.12	$\ln(0.1+Y)$	−1.55	0.43
2000	0.13	0.11		−1.57	0.41
2008–2012	0.15	0.12		−1.46	0.42
Educational level					
1990	0.94	0.28	$Y\times10$	9.39	2.80
2000	1.02	0.29		10.24	2.86
2008–2012	1.11	0.27		11.15	2.73
Unemployment (proportion)					
1990	0.07	0.05	$\ln(0.01+Y)$	−2.70	0.56
2000	0.06	0.06		−2.79	0.61
2008–2012	0.10	0.06		−2.34	0.53
Professional/ managerial occupations (proportion)					
1990	0.25	0.12	$Y\times10$	2.50	1.22
2000	0.32	0.14		3.18	1.40
2008–2012	0.37	0.16		3.74	1.60
Public assistance (proportion)					
1990	0.08	0.08	$\ln(Y\times10+0.1)$	−0.91	0.83
2000	0.04	0.05		−0.49	0.47
2008–2012	0.03	0.04		−0.41	0.35
Female-headed households (proportion)					
1990	0.11	0.10	$\ln(0.01+Y)$	−2.30	0.64

TABLE 1 ***Continued***

Variable and year(s)	Pre-transformation Mean	SD	Transformation of variable (*Y*)	Post- transformation Mean	SD
2000	0.11	0.09		−2.34	0.70
2008–2012	0.12	0.10		−2.28	0.77
Crowded housing (proportion)					
1990	0.05	0.07	$\ln(Y \times 10 + 0.1)$	−0.60	0.77
2000	0.06	0.09		−0.71	0.90
2008–2012	0.04	0.06		−0.45	0.55
Housing unit value (median in dollars[a])					
1990	124,881.55	103,603.53	$\ln(0.01 + Y)$	11.22	2.21
2000	134,904.00	110,838.00		11.34	2.20
2008–2012	124,731.27	100,068.51		11.38	1.48

Note: *SD* standard deviation

[a]Reported dollar values are adjusted for inflation using the CPI-U-RS with a 2000 referent

TABLE 2 Fit statistics for confirmatory factor analysis models of neighborhood socio-economic status, testing levels of time-invariance

Model	Chi-square (*df*)	RMSEA	CFI	ΔCFI
A1: Configural invariance, 9 indicator variables[a]	73,527 (294)	0.062	0.875	NA
B1: Configural invariance, 5 indicator variables[b]	17,118 (72)	0.060	0.966	NA
B2: Model B1 with correlated error[c]	15,122 (69)	0.058	0.970	+0.004
B3: Model B2 with weak invariance[d]	20,068 (77)	0.063	0.960	−0.010
B4: Model B3 with strong invariance[e]	86,023 (85)	0.125	0.829	−0.131
C1: Model B4 with intercepts corrected[f]	21,090 (85)	0.063	0.958	NA

Notes: *df* degrees of freedom, *RMSEA* root mean square error of approximation, *CFI* comparative fit index, *ΔCFI* change in comparative fit index, *CFA* confirmatory factor analysis

[a]Model A1 is a CFA model estimated using 9 indicator variables for the socio-economic status of U.S. census tracts (i.e., household income, educational level, housing unit value, and proportions of poverty, unemployment, professional/managerial occupations, public assistance, female-headed households, and crowded housing)

[b]Model B1 is a CFA model estimated using five indicator variables (i.e., household income, educational level, and proportions of poverty, unemployment, and female-headed households)

[c]Model B2 adds to model B1 a correlation of the residual error terms for poverty and household income

[d]Model B3 adds to model B2 the constraint that factor loadings for respective indicators are equal over time

[e]Model B4 adds to model B3 the constraint that the intercepts for respective indicators are equal over time

[f]Model C1 estimates model B4 using data on the five indicator variables that corrects for strong invariance

close equivalence to the constrained factor loadings for model B3 (Table 3). Recall also that because median household income was selected as the reference variable, the factor loadings for this indicator are constrained to 1.00.

The final step was to consider the stability of the indicator intercepts over time (model B4), or strong factorial invariance. To test strong invariance, we constrained the intercepts of all equivalent variables to be equal across the three assessments. In implementation, this means that the constrained intercepts will be equal to a weighted average of the unconstrained intercepts (see Table 4). As reported in Table 2, model B4 had considerably worse fit than model B3 (ΔCFI=−0.131) which indicated that we cannot conclude that the measure is strongly invariant over time. As further (qualitative) evidence of the lack of time-invariance of the indicator intercepts, consider how the freely estimated values of the intercepts for model B3 reported in Table 4 change over time. Recall that in order to statistically identify the time trend in NSES in all of the longitudinal models, we employ 1990 as the referent time point and median household income as the referent indicator variable. Thus, in model B3 the indicator intercepts in 1990 are equal to the means of the respective variables at 1990 in Table 1, and the intercept for median household income is constrained to be equal at all time points. Inspection of the values of the indicator intercepts in model B3 show that the indicators with the largest changes are educational level and unemployment. For both indicators, the intercepts are increasing and thus indicate that the rate of change in education and unemployment is out of pace with the rate of change in the other indicators. In order to achieve the objective of employing a consistent measurement instrument for NSES at each time point, this change in the scaling of the indicator variables relative to one another will need to be addressed.

In summary, our CFA analyses provided evidence supporting two out of the three levels of longitudinal invariance: i.e., configural invariance and weak factorial invariance, but not strong factorial invariance. The poor fit of the model testing strong factorial invariance implied that, for the same overall level of NSES, the expected value of the measured variables was not equal. Bias that arises from a lack of strong invariance can be corrected by adding a constant to measured variables to correct for change; in the item response theory literature, this lack of strong invariance is referred to as differential item functioning.[63] The constant that is added is the difference between the intercepts shown in Table 3. For example, for education in 2000, one would add the constant (9.39–9.61=) −0.22 to the values of education in the 2000 wave. When we corrected the data on our indicator variables for strong invariance, and then tested the fit of a model in which we constrained both the factor loadings and indicator intercepts to be equal (model C1), we observed about as good of fit (RMSEA=0.063 and CFI=0.958) as the reference model, model B2, in which the loadings and intercepts were unconstrained and allowed to vary freely. Our final, best time-invariant model of NSES was thus model C1.

In Table 4, we compare the factor loadings from the unconstrained reference model (model B2) and our final time-invariant model (model C1). As expected, on the basis of the presence of factorial invariance in model B2, all loadings are similar in magnitude over the three assessments and about equal to those of model C1. The bottom panel of Table 4 displays the standardized loadings which are useful for comparing the magnitude of loadings across indicator variables. Although all variables are good indicators of NSES with loadings above 0.60, the strongest indicators are median household income and proportion below the poverty level, both of which have standardized loadings greater than 0.90 in model C1 and at each assessment in model B2.

TABLE 3 Factor loadings from longitudinal confirmatory factor analysis models of neighborhood socioeconomic status in which factor loadings are unconstrained (model B2) versus constrained (model B3 and model C1)

	Factor loadings unconstrained			Factor loadings constrained	
	Model B2[a]			Model B3[b]	Model C1[c]
Variable	1990	2000	2008–2012	All time points	All time points
Median household income	1.00	1.00	1.00	1.00	1.00
Educational level	4.85	5.13	4.44	4.88	4.88
Unemployment	−1.01	−1.09	−0.72	−0.96	−0.96
Female-headed households	−1.11	−0.96	−1.01	−1.09	−1.09
Poverty	−0.96	−0.93	−0.88	−0.93	−0.93
Standardized loadings[d]					
Median household income	0.92	0.92	0.93	0.92	0.92
Educational level	0.74	0.74	0.72	0.74	0.74
Unemployment	0.77	0.74	0.60	0.75	0.75
Female-headed households	0.74	0.65	0.59	0.73	0.73
Poverty	0.94	0.93	0.92	0.93	0.93

All factor loadings are highly statistically significant ($p<0.0001$)

[a]Model B2 is estimated using five indicator variables for the socio-economic status of U.S. census tracts, with correlated errors of two of the indicator variables. Median household income is selected as the reference variable for statistical identification of all models, and thus the factor loadings are constrained to 1.00 at each time point

[b]Model B3 adds to model B2 constraints on the factor loadings, but like model B2 allows the indicator intercepts to be freely estimated

[c]Model C1 adds to model B3 constraints on intercepts for each respective indicator to be equal over time, and it adjusts the data for strong invariance over time. All factor loadings are highly statistically significant ($p<0.0001$)

[d]Model C1, the standardized loadings will change slightly over time due to differences in the variance of the indicator variables at each time point. Thus, we display the standardized loadings for the 1990 assessment

TABLE 4 Indicator variable intercepts from longitudinal confirmatory factor analysis models of neighborhood socio-economic status in which indicator variable intercepts are unconstrained (model B3) versus constrained (models B4 and C1)

	Indicator intercepts unconstrained			Indicator intercepts constrained	
	Model B3[a]			Model B4[b]	Model C1[c]
Variable	1990	2000	2008–2012	All time points	All time points
Median household income	10.47	10.47	10.47	10.47	10.47
Educational level	9.39	9.61	11.06	10.35	9.39
Unemployment	−2.70	−2.67	−2.32	−2.59	−2.70
Female-headed households	−2.30	−2.20	−2.26	−2.28	−2.30
Poverty	−1.55	−1.45	−1.44	−1.50	−1.55

All indicator intercepts are highly statistically significant ($p<0.0001$)

[a]Model B3 is estimated using five indicator variables for the socio-economic status of U.S. census tracts, with correlated errors of two of the indicator variables and constraints on the factor loadings (not shown) for each respective indicator to be equal over time. Median household income is selected as the reference variable and 1990 is selected as the reference time point for statistical identification of all models; thus the indicator intercepts for median household income are constrained to their 1990 value at each time point

[b]Model B4 adds to model B3 constraints on the intercepts for each respective indicator to be equal over time

[c]Model C1 is equivalent to model B4, and also adjusts the model for strong invariance over time

In Table 5 we summarize the characteristics of the final latent NSES measure (using model C1), including the mean score of NSES indexed to 1990 (i.e., with a mean of zero in 1990), its variance and composite reliability. We find that NSES increased from 1990 to 2000 by $\left(\frac{0.13-0.00}{\sqrt{0.41}}=\right)0.20$ standard deviation units, but then decreased by nearly the same amount, (i.e., $\frac{(0.02-0.13)}{\sqrt{0.40}}=-0.17$ standard deviation units) by the 2008–2012 assessment. The composite reliability of the latent variables is high, with values above 0.95 at each assessment. The correlation of NSES over assessments is also very high with the correlation of 0.97 between 1990 and 2000 assessments and 0.92 between 2000 and 2008–2012 assessments, indicating stability of the NSES measure over time.

DISCUSSION

Our objective was to develop a time-invariant measure of NSES for 1990 through about 2010. We achieved this goal using five indicators and a unidimensional model that met conditions for configural and weak factorial time-invariance (pertaining to factor structure and factor loadings) but not strong factorial time-invariance (pertaining to indicator intercepts). Change in indicator intercepts is also described as differential item functioning and can lead to bias in the longitudinal application of a measure.[64] In our final model, we corrected for differential item functioning and found that, although NSES increased between 1990 and 2000 (by about 0.2 standard deviations units), this gain was almost entirely lost about 10 years later. The pattern of expansion and collapse we observe for NSES parallels changes in housing, financial and labor markets over these decades, whereby the Great Recession reversed the "economic boom" of the 1990s. Other studies have described the negative consequences of the Great Recession for levels of employment, income and poverty,[65] as well as family arrangements and their potential for leveraging resources.[66,67] To our knowledge, ours is the first to describe the apparent consequences for trends in socio-economic conditions of U.S. census tracts.

Through the process of developing our final time-invariant NSES measure, we determined that four of the hypothesized indicators failed to consistently load with the other indicators at all three assessments. We speculate that macro social and economic changes occurring over the 1990s and 2000s may account for the changing relationship between these indicators and NSES. For example, fundamental restructuring of policies to assist low-income families initiated with the 1996 Welfare Reform, including conditions on assistance such as time-limits,[68] may have made the proportion of individuals receiving assistance at any given time in a community a poorer indicator of that community's underlying NSES. Similarly, the poor performance of the proportion employed in managerial and professional occupations may reflect the recent critique of "big class" stratification models,[69–71] including that they have become insufficiently nuanced to capture current class cleavages. Finally, we suspect that the housing market bubble, collapse, and

Although the SOC classifies occupations employ four levels of hierarchical detail, an indicator for professional and managerial occupations collapses over ten of the major occupation groups at the highest level of this detail, providing a less nuanced indicator of occupations than even the highest level of the SOC. By so doing, this indicator for professional and managerial occupations is understood to respectively capture either two of the ten "big class" categories of the Featherman-Hauser classification system, or one of the five "big class" categories in the Erikson-Goldthorpe classification system, after excluding self-employed occupations.[72]

TABLE 5 Characteristics of the latent measure for neighborhood socio-economic status (NSES) estimated in the final confirmatory factor analysis model (model C1)

	Assessment		
	1990	2000	2008–2012
Mean of NSES[a]	0.00	0.13	0.02
Standard deviation of NSES	0.41	0.40	0.40
Composite reliability of NSES	.953	.951	.951
Correlation of NSES over assessments			
Assessment 1990	1.00		
Assessment 2000	0.97	1.00	
Assessment 2008–2012	0.92	0.97	1.00

Model C1 is estimated using five indicator variables for the socio-economic status of U.S. census tracts that are corrected for strong invariance over time, with correlated errors of two of the indicator variables and with constraints on the factor loadings and intercepts for each respective indicator to be equal over time

[a]The mean of the latent measure for NSES is constrained to zero in 1990 for identification purposes

subprime mortgage crisis[73] may have changed the relationship between socioeconomic position and our housing-related variables making these indicators too volatile or inadequately discriminating of NSES.

With respect to the time-invariance of the indicator intercepts, we observed that educational composition changed the most with a higher level of education required to attain the same level of NSES in each decade. These results are consistent with findings from our descriptive statistics and elsewhere,[74,75] that the upward trend in U.S. educational attainment was not equally matched by changes in other socioeconomic indicators. Our adjustments for strong time-invariance ensures that, in the application of the NSES model, these changes in the measurement of NSES will not be misinterpreted as changes in the consequences of NSES.

While there are many possible alternative indicators and models of NSES, and though we drew upon existing literature to inform our selection, our indicators were necessarily restricted to those from publicly accessible data on census tracts. The limitations of this approach while still the most feasible for nationally-generalizable longitudinal research, are well established.[2,6,7] Furthermore, as we describe in our methods, a reference variable is required in CFA for purposes of identification. Standard practice is to select the variable hypothesized to be most closely related to the latent variable, and although both poverty and income met this criterion,[76] income offered a more intuitive interpretation. It is important to note that our findings on the lack of intercept stability over time are predicated on our choice of income as the reference variable. Finally, it was not possible to assess time-invariance for intercensal years when no publicly available nationally inclusive census tract data exists, and it was beyond the scope of this paper to assess time-invariance using censuses prior to 1990 or earlier ACS 5-year estimates (beginning in 2005–2009). Thus, our findings should not be extrapolated to points prior to 1990 and should be applied cautiously to intercensal years. That said, we have conducted sensitivity analyses using the 2005–2009 ACS and found similar results on the five-factor unidimensional structure, presence of configural and weak time-invariance, and absence of strong time-invariance. We have also tried to be careful about language throughout the text to recognize that, although the centroid of the ACS 2008–2012 assessment is 2010 about 10 years after the 2000 Census, multiyear estimates capture a "window" of time rather than time-point.[77,78]

All latent variable models suffer from the disadvantage that they cannot be objectively defined—this is the indeterminacy problem.[31,79] Hence, there are potentially other models of NSES which may also evidence the properties of fitting the data well and meeting statistical qualifications for time-invariant measurement. As with any latent construct, we therefore caution the reader against over-interpretation and/or reification of the construct[80] The latent variable is a hypothesized variable which explains the relations among the variables. It cannot attain the status of 'truth' but may be strengthened through additional independent validation studies. We nonetheless subscribe to the view that an invariant latent variable based approach is useful for two reasons. First, the use of a latent variable rather than a (potentially large) number of separate indicator variables leads to more parsimonious statistical models. Second, latent variable models are in use by NSES researchers, regardless. Yet, on the basis of our findings, it is likely that these models are not time-invariant and thereby are introducing bias into longitudinal inferences about NSES change or the consequences of NSES change. By ensuring that our NSES measure is consistent over time, we introduce additional rigor into the research on time-varying assessment of latent contextual measures. Finally, there is no reason to believe that our measurement approach, or any other, will continue to function coherently into the future (or further into the past). The utility of different indicators may change—for example, if an overarching majority of individuals obtain a college degree in the future, education will cease to be a useful indicator of NSES. It is possible that models developed using methods for accelerated longitudinal designs[81] or integrative data analysis[82] might be useful in such a case. In these designs, indicator variables can be added or removed over time as they become appropriate, or cease to be appropriate, retaining the comparability of the latent variable. The disadvantage of such an approach is that they would require a stronger theoretical understanding of the nature of the NSES construct than we believe currently exists and a much longer series of measurement occasions then was feasible within the scope of the existing study.

Our final NSES measure provides, to our knowledge, the only such measure to have been evaluated for, and adjusted to ensure, time-invariance. We, thus, offer not only a research tool for the longitudinal study and surveillance of NSES, but a more general methodology for measuring changing neighborhood conditions. It is unclear how much the failure to address potential time-invariance may bias inferences about the accumulation and change in neighborhood conditions, but continuing debate about neighborhood change as a lever for social policy underscores the importance of examining these questions using the methodologies and tools from this study.

REFERENCES

1. Yen IH, Syme SL. The social environment and health: a discussion of the epidemiologic literature. *Annu Rev Public Health*. 1999; 20: 287–308.
2. Robert SA. Socioeconomic position and health: The independent contribution of community socioeconomic context. *Annu Rev Sociol*. 1999; 25: 489–516.
3. Obradović J, Pardini DA, Long JD, Loeber R. Measuring interpersonal callousness in boys from childhood to adolescence: An examination of longitudinal invariance and temporal stability. *J Clin Child Adolesc Psychol*. 2007; 36(3): 276–292.
4. Pickett KE, Pearl M. Multilevel analyses of neighbourhood socioeconomic context and health outcomes: a critical review. *J Epidemiol Community Health*. 2001; 55(2): 111–122.
5. Galea S, Vlahov D, Tracy M, Hoover DR, Resnick H, Kilpatrick D. Hispanic ethnicity and post-traumatic stress disorder after a disaster: Evidence from a general population survey after September 11, 2001. *Ann Epidemiol*. 2004; 14(8): 520–531.
6. Sampson RJ, Morenoff JD, Gannon-Rowley T. Assessing neighborhood effects: social processes and new directions in research. *Annu Rev Sociol*. 2002; 28(1): 443–478.
7. Diez Roux AV. Investigating neighborhood and area effects on health. *Am J Public Health*. 2001; 91(11): 1783–1789.
8. Morenoff JD, House JS, Hansen BB, Williams DR, Kaplan GA, Hunte HE. Understanding social disparities in hypertension prevalence, awareness, treatment, and control: the role of neighborhood context. *Soc Sci Med*. 2007; 65(9): 1853–1866.
9. Bird CE, Seeman T, Escarce JJ, et al. Neighbourhood socioeconomic status and biological 'wear and tear' in a nationally representative sample of US adults. *J Epidemiol Community Health*. 2010; 64(10): 860–865.
10. Freedman VA, Grafova IB, Schoeni RF, Rogowski J. Neighborhoods and disability in later life. *Soc Sci Med*. 2008; 66(11): 2253–2267.
11. Ross CE. Neighborhood disadvantage and adult depression. *J Health Soc Behav*. 2000; 41(2): 177–187.

12. Robert SA, Cagney KA, Weden MM. A life-course approach to the study of neighborhoods and health. In: Chloe E, Bird PC, Fremont AM, Timmermans S, eds. *Handbook of Medical Sociology.* 6th ed. Nashville: Vanderbilt University Press; 2010.
13. Janson CG. Factorial social ecology - an attempt at summary and evaluation. *Annu Rev Sociol*. 1980; 6: 433–456.
14. Schwirian KP. Models of neighborhood change. *Annu Rev Sociol*. 1983; 9: 83–102.
15. Browning CR, Cagney KA, Morris K. Early Chicago School theory. In: G. Bruinsma DW, ed. *Encyclopedia of Criminology and Criminal Justice*. New York: Springer Science and Business Media; 2014: 1233–1242.
16. Hogan DP, Kitagawa EM. The Impact of social status, family structure, and neighborhood on the fertility of Black adolescents. *Am J Sociol*. 1985; 90(4): 825–855.
17. Diez Roux AV, Nieto FJ, Muntaner C, et al. Neighborhood environments and coronary heart disease: A multilevel analysis. *Am J Epidemiol*. 1997; 146(1): 48–63.
18. Duncan GJ, Aber L. *Neighborhood models and measures. Neighborhood Poverty: Context and Consequences for Children*. New York: Russell Sage Foundation; 1997: 62–78.
19. Ross CE, Mirowsky J. Neighborhood disadvantage, disorder, and health. *J Health Soc Behav*. 2001; 42(3): 258–276.
20. Boardman JD, Finch BK, Ellison CG, Williams DR, Jackson JS. Neighborhood disadvantage, stress, and drug use among adults. *J Health Soc Behav*. 2001; 42(2): 151–165.
21. Browning CR, Cagney KA. Neighborhood structural disadvantage, collective efficacy, and self-rated physical health in an urban setting. *J Health Soc Behav*. 2002; 43(4): 383–399.
22. Weden MA, Carpiano RA, Robert SA. Subjective and objective neighborhood characteristics and adult health. *Soc Sci Med*. 2008; 66(6): 1256–1270.
23. Messer LC, Laraia BA, Kaufman JS, et al. The development of a standardized neighborhood deprivation index. *J Urban Health-Bull New York Acad Med*. 2006; 83(6): 1041–1062.
24. Wen M, Hawkley LC, Cacioppo JT. Objective and perceived neighborhood environment, individual SES and psychosocial factors, and self-rated health: an analysis of older adults in Cook County, Illinois. *Soc Sci Med*. 2006; 63(10): 2575–2590.
25. Winkleby M, Cubbin C, Ahn D. Effect of cross-level interaction between individual and neighborhood socioeconomic status on adult mortality rates. *Am J Public Health*. 2006; 96(12): 2145–2153.
26. Dubowitz T, Heron M, Bird CE, et al. Neighborhood socioeconomic status and fruit and vegetable intake among whites, blacks, and Mexican Americans in the United States. *Am J Clin Nutr*. 2008; 87(6): 1883–1891.
27. Clarke P, Morenoff J, Debbink M, Golberstein E, Elliott MR, Lantz PM. Cumulative exposure to neighborhood context: consequences for health transitions over the adult life course. *Research on Aging*. 2014; 36(1): 115–142.
28. Wodtke GT, Harding DJ, Elwert F. Neighborhood effects in temporal perspective: The impact of long-term exposure to concentrated disadvantage on high school graduation. *Am Sociol Rev*. 2011; 76(5): 713–736.
29. Little TD. *Longitudinal structural equation modeling*. New York: Guilford Press; 2013.
30. Borsboom D. When does measurement invariance matter? *Med Care*. 2006; 44(11): S176–S181.
31. Bollen KA. Latent variables in psychology and the social sciences. *Annu Rev Psychol*. 2002; 53: 605–634.
32. Marshall GN. Posttraumatic Stress Disorder Symptom Checklist: Factor structure and English-Spanish measurement invariance. *J Trauma Stress*. 2004; 17(3): 223–230.
33. McDonald SD, Beckham JC, Morey R, Marx C, Tupler LA, Calhoun PS. Factorial invariance of posttraumatic stress disorder symptoms across three veteran samples. *J Trauma Stress*. 2008; 21(3): 309–317.

34. Miles JNV, Shih RA, Tucker JS, Zhou A, D'Amico EJ. Assessing measurement invariance of familism and parental respect across race/ethnicity in adolescents. *BMC Med Res Methodol*. 2012; 12(1): 61.
35. Shevlin M, Brunsden V, Miles JNV. Satisfaction With Life Scale: analysis of factorial invariance, mean structures and reliability. *Personal Individ Differ*. 1998; 25: 569–574.
36. Herrero J, Gracia E. Measuring perceived community support: Factorial structure, longitudinal invariance, and predictive validity of the PCSQ (Perceived Community Support Questionnaire). *J Comm Psychol*. 2007; 35(2): 197–217.
37. Motl RW, DiStefano C. Longitudinal invariance of self-esteem and method effects associated with negatively worded items. *Struct Equ Model*. 2002; 9(4): 562–578.
38. Pentz M, Chou C. Measurement invariance in longitudinal clinical research assuming change from development and intervention. *J Consult Clin Psychol*. 1994; 62(3): 450–462.
39. Schaie KW, Maitland SB, Willis SL, Intrieri RC. Longitudinal invariance of adult psychometric ability factor structures across 7 years. *Psychol Aging*. 1998; 13(1): 8.
40. Millsap RE. *Statistical approaches to measurement invariance*. New York: Routledge; 2011.
41. US Census Bureau. 1990 Census. http://www.census.gov/main/www/cen1990.html. Accessed 9/09/2014.
42. US Census Bureau. 2000 Census. www.census.gov/main/www/cen2000.html. Accessed 9/09/2014.
43. US Census Bureau. 2008–2012 American Community Survey. www2.census.gov/acs2012_5yr/summaryfile Accessed 9/092014.
44. Logan JR, Xu Z, Stults B. Interpolating US decennial census tract data from as early as 1970 to 2010: a longitudinal tract database. *Prof Geogr*. 2014; 66(3): 412–420.
45. Tatian PA. *Neighborhood Change Database (NCDB) 1970-2000 Tract Data: Data Users Guide*. Washington: Urban Institute; 2003.
46. US Bureau of Labor Statistics. Updated Consumer Price Index Research Series Using Current Methods (CPI-U-RS), All items. 1978–2013. http://www.bls.gov/cpi/cpiurs.htm. Accessed 9/10/2014.
47. Posey K, Welniak E, Nelson C. Income in the American Community Survey: comparisons to Census 2000. Paper presented at: Annual Meeting of the American Statistical Association, San Francisco, California, 2003; http://www2.census.gov/programs-surveys/acs/methodology/ASA_nelson.pdf.
48. US Census Bureau. A Compass for Understanding and Using the American Community Survey: What Researchers Need to Know. Washington, DC: U.S. Government Printing Office; 2009.
49. US Census Bureau. Comparing 2012 American Community Survey Data. 2014; https://www.census.gov/acs/www/guidance_for_data_users/comparing_2012/. Accessed 9/1/2014, 2014.
50. Scopp T. The relationship between the 1990 Census and Census 2000 industry and occupation classification systems. US Census Bureau Technical Paper #65. 2003; http://beta.census.gov/people/io/files/techpaper2000.pdf. Accessed 9/23/2014.
51. US Bureau of Labor Statistics. What's New in the 2010 SOC. 2010; http://www.bls.gov/soc/soc_2010_whats_new.pdf. Accessed 9/23/2014.
52. Muthén BO, Muthén L. Mplus Version 7.11. Los Angeles: Muthen and Muthen; 2013.
53. Yuan K, Bentler PM. Three likelihood-based methods for mean and covariance structure analysis with nonnormal missing data. *Sociol Methodol*. 2000; 30: 167–202.
54. Satorra A, Bentler PM. Corrections to test statistics and standard errors in covariance structure analysis. In: von Eye A, Clogg CC, eds. *Latent Variables Analysis: Applications for Developmental Research*. Thousand Oaks: Sage; 1994.
55. Curran PJ, West SG, Finch JF. The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. *Psychol Methods*. 1996; 1(1): 16–29.

56. Browne MW. An overview of analytic rotation in exploratory factor analysis. *Multivariate Behav Res*. 2001; 36(1): 111–150.
57. Millsap RE, Meredith. Factorial invariance: historical perspectives and new problems. In: Cudeck R, MacCallum R, eds. *Factor Analysis at 100: Historical Developments and Future Directions*. Hillsdale: Erlbaum; 2007.
58. Steiger JH, Lind JC. *Statistically based tests for the number of factors*. Iowa City: Psychometric Society; 1980.
59. Bentler PM, Bonett DG. Significance tests and goodness of fit in the analysis of covariance structures. *Psychol Bull*. 1980; 88: 588–606.
60. Hu L, Bentler P. Cutoff criteria for fit indices in covariance structure analysis: Conventional criteria versus new alternatives. *Struct Equ Model*. 1999; 6: 1–55.
61. Cheung GW, Rensvold RB. Evaluation goodness-of-fit indexes for testing measurement invariance. *Struct Equ Model*. 2002; 9: 235–255.
62. Kaplan D. On the modification and predictive validity of covariance structure models. *Qual Quant*. 1991; 25: 307–314.
63. Osterlind SJ, Everson HT. *Differential Item Functioning*. Thousand Oaks: Sage; 2009.
64. Embretson SE. The new rules of measurement. *Psychol Assess*. 1996; 8(4): 341–349.
65. Hurd MD, Rohwedder S. *Effects of the financial crisis and Great Recession on American households*. Washington: National Bureau of Economic Research; 2010.
66. Harknett K, Kuperberg A. Education, labor markets and the retreat from marriage. *Social Forces*. 2011; 90(1): 41–63.
67. Livingston G, Parker K. Since the start of the Great Recession, more children raised by grandparents. Pew Research Center. 2010; http://www.pewsocialtrends.org/2010/09/09/since-the-start-of-the-greatrecession-more-children-raised-by-grandparents/.
68. Grogger J, Karoly LA. *Welfare reform: effects of a decade of change*. Cambridge: Harvard University Press; 2009.
69. Weeden KA, Grusky DB. Are there any big classes at all? *Res Soc Stratif Mobil*. 2004; 22: 3–56.
70. Weeden KA, Grusky DB. The case for a new class map. *Am J Sociol*. 2005; 111(1): 141–212.
71. Hauser RM, Warren JR. Socioeconomic indexes for occupations: A review, update, and critique. *Sociol Methodol*. 1997; 27(1): 177–298.
72. Weeden KA, Kim Y-M, Di Carlo M, Grusky DB. Social class and earnings inequality. *Am Behav Sci*. 2007; 50(5): 702–736.
73. Moench E, Ng S. A hierarchical factor analysis of U.S. housing market dynamics. *Econometrics J*. 2011; 14(1): C1–C24.
74. Ryan CL, Siebens J. Educational attainment in the United States: 2009. Current Population Reports. U.S. Census Bureau. 2012; February 2012.
75. Heckman JJ, Lafontaine PA. The American high school graduation rate: trends and levels. *Rev Econ Stat*. 2010; 92(2): 244–262.
76. Duncan GJ, Brooks-Gunn J, Aber L. *Neighborhood Poverty: Context and Consequences for Children*. New York: Russell Sage Foundation; 1997.
77. Beaghen M, Weidman L. Statistical issues of interpretation of the American Community Survey's one-, three-, and five-year period estimates. 2008; http://www.census.gov/acs/www/Downloads/library/2008/2008_Beaghen_01.pdf. Accessed 9/23/2014.
78. Weden MM, Peterson C, Shih RA. Evaluating tract-level intercensal estimates of neighborhood demographics and socioeconomics for U.S. counties 2001-2009. *Popul Res Policy Rev*. 2015; 34(4): 541–559.
79. Steiger JH. Factor indeterminancy in the 1930s and the1970s: Some interesting parallels. *Psychometrika*. 1979; 44(2): 157–167.
80. Sokal RR, Rohlf FJ, Zang E, Osness W. Reification in factor analysis: A plasmode based on human physiology-of-exercise variables. *Multivar Behav Res*. 1980; 15(2): 181–202.

81. Galbraith SJ, Bowden J, Mander A. Accelerated longitudinal designs: an overview of modelling, power, costs and handling missing data. Statistical Methods in Medical Research. 2014;0962280214547150.
82. Curran PJ, Hussong AM. Integrative data analysis: the simultaneous analysis of multiple data sets. *Psychol Methods*. 2009; 14(2): 81–100.