Income Segregation: Up or Down, and for Whom?

John R. Logan, and Andrew Foster, Brown University

Hongwei Xu, Queens College, CUNY

Wenquan Zhang, University of Wisconsin, Whitewater

John Logan is the corresponding author, Department of Sociology, Box 1916, Brown University, Providence RI 02912; phone 401-863-2267; email john_logan@brown.edu.

Abstract

Reports of rising income segregation have been brought into question by the observation that post-2000 estimates are upwardly biased due to a reduction in the sample sizes on which they are based. Recent studies have offered estimates of this "sample-count" bias using public data. We show here that there are two substantial sources of systematic bias in estimating segregation levels: bias associated with sample size and bias associated with using weighted sample data. We rely on new correction methods using the original census sample data for individual households to provide more accurate estimates. Family income segregation rose markedly in the 1980s but only selectively after 1990. For some categories of families, segregation declined after 1990. There has been an upward trend for families with children, but not specifically for families with children in the upper or lower 10% of the income distribution. Separate analyses by race/ethnicity show that segregation was not generally higher among blacks and Hispanics than among white families, and evidence of segregation trends for these separate groups is mixed. Segregation increased for all three for families with children, and particularly for Hispanics (but not whites or blacks) in the upper 10%. Trends vary for specific combinations of race/ethnicity, presence of children, and location in the income distribution, offering new challenges for understanding the underlying processes of change.

Income Segregation: Up or Down, and for Whom?

Evidence of increasing income inequality in the United States has heightened interest in the degree to which different social classes are separating into different neighborhoods based on their incomes. Several recent studies focusing on the post-2000 period have reported that income segregation is trending upward. For example, Bischoff and Reardon (2014, p. 208) state "Socioeconomic residential sorting has grown substantially in the last forty years … and the bulk of that growth occurred in the 1980s and in the 2000s" (see also Florida and Mellander 2015, Fry and Taylor 2012). These reports have been questioned by the insight that the observed trends after 2000 are distorted by changes in census data collection. Logan et al. (2018) point out that the post-2000 income data on which all recent measures are based come from much smaller samples (less than 8%) in the American Community Survey (ACS) than were previously available from the decennial census (about 16%). They demonstrate that sampling at the census tract level results in an inherent upward bias in standard measures of income segregation, and that this bias is greater when the sample size is smaller.

Bias due to limited sample size has two main implications for past findings. First, income segregation may only have appeared to increase after 2000. Logan et al. (2018) offer a rough estimate that as much as half of the apparent increase in income segregation may be due to the Census Bureau's new, smaller samples, an estimate seconded by a subsequent reanalysis by Reardon et al. (2018). Second, the same bias has greater effects on estimates of income segregation in racial/ethnic subgroups of the population, because samples of African Americans and Latinos in particular are very limited in most census tracts. The much higher level and

substantially greater increases of income segregation that have been observed among minorities compared to whites since 2000 may be misleading as a result.

We confirm these insights but show that there is another equally important source of upward bias in segregation estimates stemming from the use of weighted sample data. Fortunately, more reliable measures can be calculated from the original unit-level household sample data collected by the Census Bureau, and we implement our proposed correction procedures for the decennial censuses of 1980, 1990, and 2000, and the American Community Surveys of 2007-11 and 2012-16. With these data we provide unbiased estimates for measures based on both original incomes and their rank-ordered transformation (H, R, and NSI as defined below). We report new findings on levels and trends in income segregation overall and for the top and bottom tenths of the population, with separate analyses for all families, families with children and for racial/ethnic subgroups. Specifically:

- Income segregation increased on all measures and for every category of families between 1980 and 1990. Most measures have been stable since 1990, and some have declined. Instead of explaining how increasing inequality translates into greater residential separation, researchers now need to understand why it may not.
- During this whole period income segregation was not consistently higher for black and Hispanic families than for white families. As true for the whole population, we find that income segregation among all three groups increased in the 1980s, but on most measures not after 1990.
- Looking specifically at families with children, there are increases after 1990 for the total population and for all three racial/ethnic groups. Based on measures that focus on the top and bottom of the income distribution, results vary by group. Segregation increased for

the lower 10% among white families with children but declined for the top 10%. Evidence is mixed for black families with children. For Hispanic families with children, there was no change for the lower 10% but a substantial increase for the top 10%.

These findings bear directly on past observations and interpretations of trends in income segregation. Research on the 1980-1990 decade found growing income segregation and attributed it in part to increasing poverty in central cities and middle-class flight to the suburbs (Jargowsky 1996, p. 996; see also Massey and Eggers 1993). Subsequent studies have emphasized how income inequality itself can translate into income segregation, building on evidence that income segregation was higher in metros with greater disparities in income (Mayer 2000, Watson 2009). Reardon and Bischoff (2011) theorize that this is a result of three kinds of processes that *motivate and enable affluent families* to seek more exclusive locations while *limiting options for the poor.* These include: 1) affluent residents' preference for living with neighbors of similar class standing, 2) the advantages in terms of public services that accrue to higher income communities with a stronger tax base, and 3) price competition in the housing market that raises prices and restricts access to such places. They reported that segregation of both affluent and poor families increased from 1970 through 2000 as income inequality rose (mainly in the 1980s). In later work they found that both low-income and high-income segregation also increased after 2000 (Bischoff and Reardon 2014). Our findings contradict these interpretations as they apply generally to family households. Income segregation did not increase for families overall, nor did the separation of the lower or upper decile of families from others increase.

Owens' (2016) calls attention more specifically to families with children, where she believes the main driver of change comes from the upper end of the class structure. She argues

that high-income and highly educated parents are becoming more conscious of the need to invest in their children's futures through more selective residential choices, leading to "increased willingness to pay to live in an expensive area associated with greater opportunities for children; and higher home prices associated with high-quality schools" (2016, p. 553). Consistent with this thesis, she reports that increasing income segregation since 2000 has occurred only among families with children. We do find increasing segregation among families with children, but not specifically for either the top or bottom tenths of the income distribution.

Several studies also have examined the role of racial segregation, mainly by studying income segregation trends separately for whites, blacks, and Hispanics (Jargowsky 1996, Massey and Fischer 2003, Watson 2009, Yang and Jargowsky 2006, Reardon and Bischoff 2011, Reardon, Fox and Townsend 2015). Most attention has been given to the situation of African Americans, though similar reasoning can apply to Hispanics. A highly segregated racial minority might tend to live in mixed-class neighborhoods due to obstacles to residential mobility for more affluent members. But if racial separation creates large black districts in urban areas, lower and higher income households may cluster in separate class-based neighborhoods within them. This has been the pattern in major cities like New York and Chicago since the early decades of the Great Migration (Logan et al. 2015). Would income segregation across tracts in this case be similar to, less than, or greater than income segregation among whites?

Jargowsky (1996) reported that income segregation among both blacks and Hispanics was greater and increased more over time in the 1970-1990 period. Using a different measure, Bischoff and Reardon (2014) reported that income segregation among blacks and Hispanics was lower than among whites in 1970, but (despite a decline in the 1990s) grew much faster over time, especially after 2000. These scholars attributed the rising segregation among blacks to the

modest relaxation of race-based segregation that has occurred since 1970. The logic is that the high segregation of blacks on the basis of race may restrict housing options even for more advantaged families, limiting the separation between higher and lower income group members. Reduction in race-based segregation might then result in an exodus of the black middle class from income-mixed neighborhoods and hence higher income segregation among African Americans. As stated by Jargowsky (1996, p. 993): "greater social distance between [race] groups constricts the housing options available to all members of the lower-status group … and leads to lower economic segregation within the group." But "[d]ecreases in racial segregation, whether spurred by changes in social distance, public policy, or other causes, should increase economic segregation as the artificial boundaries limiting housing options are removed." (This point is reminiscent of Wilson's [1987] discussion of the black middle-class exodus from poor inner city neighborhoods. See also Reardon and Bischoff 2011, p. 1106-1107.)

These interpretations also need to be reconsidered, because we find that income segregation among blacks and Hispanics was not generally higher than among whites. Further it is also only among Hispanics (especially Hispanic families with children) that we find increasing separation of the top 10% of the group from others.

Research design

We organize the description of our approach in three parts: first a definition of data sources and indicators, then a comparison of measures of segregation measures that we use here, and finally an analysis of the sources of error in estimation. In the final section we report corrected estimates of trends in segregation for all families and for different categories of families from 1980 through 2012-16.

Data sources and measures

We analyze confidential household records in a Research Data Center of the Census Bureau's Center for Economic Studies. Following the lead of past studies, we focus on family income segregation, leaving aside issues associated with single-person and non-family households. The original records are from samples: the one-in-six long form samples of the decennial census in 1980, 1990, and 2000, and the nearly 8% samples that result from pooling annual data from the American Community Survey (ACS) in 2007 through 2011 and again in 2012 through 2016. Family income is measured as the sum of income of all family household members from all sources. We apply household weights to these data as developed by the Census Bureau to correct for under- or over-representation of various population segments in the sampling process. The income data are not top-coded. In 1980 the Bureau protected privacy of personal information by suppressing income data in tracts with small populations, but there is no suppression in the files available to us. Since 1980 the Bureau has relied on data swapping to protect privacy. The general approach to swapping is to exchange the record for one person or household who have an uncommon set of personal characteristics with the record of a somewhat similar person or household in another nearby tract. The files available in the FSRDC, like those in the public data, include such swapped records.

Results of analyses with these data are released only after a disclosure review by census professionals. We conducted analyses in all years for all 384 metropolitan regions using constant 2010 metro boundaries. We have gained approval to report segregation measures for metros that had more residents than the smallest state in each study year (in 1980, for example, the smallest state was Alaska with 401,851 residents. These data are available from the "DATA" section of the Diversity and Disparities website at Brown University

(https://s4.ad.brown.edu/projects/diversity/Data/data.htm). Here we report average values of segregation measures for the 95 metros that met this criterion in every year. Averages are weighted by the number of families reporting income in a given metro, or by the number of white, black, or Hispanic families for group-specific measures. In addition, we impose a metro sample size threshold of 100 unweighted cases for each category of families studied here. We also omit cases for a given category of families and year if the estimate is below 0, which occurs occasionally with sample sizes that are only modestly above 100. This reduces the number of metros for analysis especially for Hispanic families and Hispanic families with children. Excluding cases with such small samples has little practical effect, since all results reported in our main findings are weighted averages.

For the calculation of race-specific measures, families are classified by the race and Hispanic origin of the household head. An advantage of access to the original sample data is that we are able to identify non-Hispanic black families in every data file (the published tables include Hispanic black families in the black category). In addition, while published tables are for families whose heads are "black alone" in 2000 and beyond (when multiple-race reporting was introduced), we are able to identify all who are "black alone or in combination with another race." This classifies African Americans in a way that is more consistent with the 1980 and 1990 reports (when only one race could be recorded).

We call attention to our use of information on unit-level family incomes, unlike studies that rely on tabulations of families in income categories, i.e., grouped data. Researchers have long been aware of the difficulties with using public data at the tract level. When income is reported in categories, the distribution of incomes within each category is unknown and has to be estimated. This estimation is more difficult for the top category (because it has no upper bound)

and the bottom category (where incomes may cluster close to zero). But it is problematic in any category, especially when samples are smaller, because incomes are not smoothly distributed within categories. Even careful approximation of the underlying income distribution can yield distorted estimates. Reardon and collaborators (e.g., Reardon and Bischoff 2011) simplify the problem by estimating segregation measures after converting incomes into percentiles. The value of their preferred measure (H) – which involves dividing the population into families above and below fixed points in the percentile distribution – can be calculated exactly for percentiles that coincide with the cutting points in the available grouped data. The value at other percentiles can be estimated by fitting a polynomial to the known points. If the full curve of values of estimated H at every percentile matches the estimated values form the unit-level household data, the overall value of H can be accurately estimated from it. Reardon et al. (2018, p. 2138) argue that because 'there is no theoretical reason to expect systematic bias related to the binning of income data," this procedure is unlikely to bias results.

Because we can replicate estimates using both grouped and ungrouped data, we are now able to assess how the use of grouped data can affect results. We present this analysis in Appendix A. We find that grouped and unit-level data may generate similar results but do not always do so. Distortions are most likely for measures of the separation of the top or bottom income groups from all others. For studies that must rely on grouped data, therefore, our advice is to proceed with caution. In our study, as we explain below, we must rely on unit-level household income data in order to carry out the correction procedures to compensate for the bias in standard income segregation estimates.

Measures of income segregation

We study several different measures of income segregation. These differ on whether they measure variation within and between tracts as entropy or as variance. Some of these are based on reducing the income distribution to a dichotomy and asking how segregated people in one income category are from all others. Reardon and Bischoff (2014) do this with a class of measures H_p and a related class of measures R_p. This is similar to the approach of studies that divide the population into three categories and calculate a standard segregation index (the Index of Dissimilarity) between the bottom and top categories (the rich and poor) as in Massey and Eggers (1993) and Massey and Fischer (2003). Having transformed incomes into rank order, Reardon and Bischoff dichotomize the income distribution at a given percentile (p), and compute the segregation between income ranks above and below this point. Both H_p and R_p can be calculated at multiple cutting points, and Bischoff and Reardon focus particularly on the segregation of those at the below the 10^{th} percentile from all others (H_{10} and R_{10}, segregation of the poorest) and those above the 90^{th} percentile (H_{90} and R_{90}, segregation of the most affluent). H denotes their use of an information theory measure of segregation between the two categories, where the entropy within census tracts is compared to the total entropy in the population. R is based instead on variance within tracts in comparison to the total variance.

Measures based on dichotomies do not make use of the full income distribution provided by the census. Four other measures do exploit the multiple and ordered category nature of the data. One, H^R, is built from the full set of rank-order measures H_p. As Bischoff and Reardon (2014, p. 228) describe it, "if we computed the segregation between those families above and below each point in the income distribution and averaged these segregation values, weighting the segregation between families with above-median income and below-median income the most, we

get the rank-order information theory index." Its equivalent based on analysis of variance is R^R, built from the full set of rank-order measures R_p. Another alternative is based on a partitioning of the variance in income, without recoding incomes to ranks. This measure is the correlation ratio, which Jargowsky (1996) refers to as the Neighborhood Sorting Index (NSI). It is simply the square root of the between-tract variance in income divided by the total variance of income, a familiar statistic in analysis of variance.

In this study we report estimates of both H and NSI. We also report an alternative version of R that we call R^F. R^F may be thought of as the NSI applied after the income data have been recoded to quantiles. It uses the same formulae as the R_p except that the {0,1} index of whether the quantile of income is less than or greater than *p* is replaced with the quantile itself. An attraction of this measure is that there is a convenient and intuitive way to construct small-sample bias corrections for it, as explained below. Henceforth, for notational convenience we will use H to denote H^R and R to denote R^F.

Biased estimates and their correction

Our findings rely on progress in identifying and correcting for biases and inaccuracies that have distorted prior studies without being recognized, and also on access to original sample data available only at a Restricted Data Center (FSRDC). We consider two issues: the fact that income data at the tract level are based on relatively small samples and the fact that the underlying unit-level data generally have sample weights.

1. Correcting bias related to sample size

Having demonstrated that measures of income segregation based on sample data are biased upwards, Logan et al. (2018) propose approximate methods using public data to correct for the upward bias in entropy-based measures (e.g., H, H_{10}, H_{90}) that draw solely on knowledge

of the tract-level sample sizes and tract population counts. The approximate bias for the entropy-based measure of income segregation in the case of an unweighted sample is

$$H_{cu} - H_{bu} = -\sum_j \frac{M_j}{M} \frac{1}{N_j}$$

where M_j and M are the tract-specific and total metro population and N_j is the tract sample size.[1] Here the subscript "bu" refers to the "uncorrected (biased), unweighted" estimate of H, and the subscript "cu" refers to the "count-corrected, unweighted" estimate. Recall that H_{bu} is a segregation index that depends on the percent of households in the sample from each tract sample that is below each percentile in the combined sample of all tracts. Let us call the difference between H_{cu} and H_{bu} the "count-based correction" because it depends only on the sample counts as proposed by Logan et al. (2018). Entropy estimates for points in the income distribution, such as H_{10}, the segregation of households in the lowest 10 percent of the distribution, have a closely related correction factor. [2]

---

[1] By taking a second-order expansion of the entropy function around the fraction p_j of households in tract *j* with income below some given level and taking expectations Logan et al. (2018) show that the expected bias in any given tract *j* is $\frac{E(\hat{p}_j - p_j)^2}{p_j(1-p_j)} = \frac{p_j(1-p_j)/N_j}{p_j(1-p_j)} = \frac{1}{N_j}$ where $\hat{p}_j$ is the corresponding fraction in the sample. This expression assumes the sample is done with replacement. Logan et al. also derive expressions for the case without replacement, which corresponds to the ACS procedure. This latter approach, however, complicates the resulting mathematical expressions and does not lead to a measurable improvement in performance in our simulated data.

[2] A somewhat similar correction was subsequently advanced by Reardon et al. (2018). Their correction applies to H and R, but not NSI. , It has two other limitations First, the derivation of their correction depends on the assumption that no systematic bias is introduced by grouped data. We show below (Appendix A) that there may in fact be distortions due to inability to model the uppoer and lower tails fo the income distribution. Second, in grouped data the sample weights have been applied but they are invisible to the researcher. Therefore it is not possible to correct

Logan et al. also propose an approach (termed Sparse Sample Variance Decomposition, SSVD) to correct the partitioning of variance within and between tracts using either the original interval-scale measure of income or a rank-order measure, which then allows for estimates of NSI or R. This is possible because 1) the income variance within tracts can be estimated from samples of any size without bias, and 2) the population-weighted average of the variance estimates for each tract from the sample converges to the within variation for the population as the number of tracts gets large. The total variance in the metropolitan area is estimated from a very large sample, and the between-tract variance is simply the difference between the total and within variance. We refer readers to the original article for details of the SSVD procedure (Logan et al. 2018).

None of these methods addresses the risk that when there is only one sample, it is subject to sampling variation that is inherently greater when samples are smaller. However, an analysis of many sample draws from a 100% transcription of incomes from the 1940 Census of the population in Chicago shows that these methods do yield unbiased estimates of income segregation, whether based on H, R, or NSI.

2. Correcting for weighted sample data

A final step that we take here is to show analytically and empirically that weighting of sample data by the Census Bureau also introduces bias. Then we offer an approach to estimate and correct for this weighting-induced bias. Unlike the data for the full population in 1940 on which Logan et al. (2018) relied to validate their sample-count bias correction, the contemporary data are weighted. This is problematic because, as we will show, heterogeneity in weights alters

for weighting, which we discuss below. A more useful tack is to turn to unit-level household data in the RDC, as we do here.

the precision of estimates of the dispersion in tract income. As a result, bias corrections for these measures must also account for weighting. In the following section we develop this point theoretically and present alternative measures that incorporate weighting.

Let us first consider entropy-based measures (H). In the case of unweighted observations, bias depends only on the sample size. But the *effective* sample size for the computation of variance of an estimator is smaller when weights are variable than when weights are uniform (e.g., all case have a weight of 1). To get some sense of this effect, suppose we have a population of 3200 with income variance v that is randomly divided into two equal-sized sub-populations A and B. Then the variance of the estimated mean for a 10 percent sample of 320 households is $v/320$. If the sampling rate for A is reduced to 6.25 percent (1/16) then the sampling rate for B must be raised to 25 percent to achieve the same variance. This change results in a total sample size of 500 rather than 320. [3] The sample must be 56% larger due to the heterogeneous weights.

To apply this insight to the entropy bias correction we need to compute the variance of the estimate $\hat{p}_j$ of the true fraction of households p_j in a given tract j with income below a given level using a weighted sample of given size:

$$E(\hat{p}_j - p_j)^2 = p_j(1-p_j)\sum w_{ij}^2$$

[3] We use the fact that if y_i is income and w_i the weight normalized so that $\sum w_i = 1$ then $Var(\sum w_i y_i) = \sum w_i^2 Var(y_i)$.

Here weights are normalized so that they sum to one within each tract, $\sum_j w_{ij} = 1$.[4] We also have to assume that household weights are independent of income. Without this assumption the bias correction will depend in general on the unknown true tract fraction p_j and thus not be feasible with sampled data. With this assumption, the bias in the entropy for a weighted sample is

$$H_{gw} - H_{bw} = -\sum_j \frac{M_j}{M} \sum_i w_{ij}^2$$

where we denote the "count-and-weighting-corrected" value of H as H_{gw}. Henceforth, unless further clarification is required, we will call H_{gw} the "corrected" estimate, and designate the estimate that only corrects for sample counts as the count-corrected estimate. H_{bw} is the uncorrected (biased) estimate calculated using weighted data.

In fact the Census Bureau generally assigns larger weights to lower income families. We have examined the impact of this correlation on our estimation procedure in two ways. First, we explored this issue analytically in a simplified two-strata sample population. This thought-experiment (available from the authors on request) suggests that our proposed expressions will be useful as long as the covariance of weight and income within tracts is small relative to the variation in income within tracts. Second, we validated our estimation procedures with 1940 data

[4] Reassuringly this expression reduces to $\frac{1}{N_j}$ in the case that all sampled families have equal weight, which would be the case with unweighted data, so that $w_{ij} = \frac{1}{N_j}$. Note that we are in effect disregarding differences in weights across tracts. This is reasonable because the published census data in 2000 provide the true tract sizes and total population (not their sample analogs) and the ACS data include adjustments based on Census 2010 full counts.

in which we have introduced weights and where the 100% population measure of segregation is known. We report these analyses in detail in Appendix B. We first assigned weights to the 1940 microdata in accordance with a multilevel model predicting weights in the Chicago metro in the ACS 2008-2012. This model shows that the relations of household income to weight is small, (b = -.0305) but statistically significant. We then compared the true value of H, R, and NSI to the estimated value based on our approach to correcting for sample counts and for weighting. These analyses demonstrate that estimates are affected by both sample-count and weight-related bias and also that our proposed alternative measures correct for both types of bias.

Note that the size of the bias must be larger in absolute value than the bias term in a sample without heterogeneous weights ($1/N_j$). This follows from the fact that because the weights sum to one and are not the same we may define $u_{ij} = w_{ij} - \frac{1}{N_j}$ with at least some u_{ij} nonzero and $\sum u_{ij} = 0$. Thus

$$\sum w_{ij}^2 = \frac{1}{N_j} + \sum u_{ij}^2 > \frac{1}{N_j}.$$

A different approach is required for bias correction in the case of variance-based estimates. Consider the NSI. It is defined as the square root of the across variance divided by the total variance in the measure of income and can be written in the presence of weights as:

$$NSI_{bw} = \left(\frac{\sum_j \frac{M_j}{M} (\sum_i w_{ij} y_{ij} - \bar{y})^2}{\sum_j \frac{M_j}{M} \sum_i w_{ij} \left(y_{ij} - \bar{y} \right)^2} \right)^{1/2}$$

where $\bar{y} = \sum_j \frac{M_j}{M} \sum_i w_{ij} y_{ij}$ is the metro-weighted mean. The subscript w indicates that this measure of the NSI uses weights and the subscript b indicates it is not corrected for sampling

bias. To construct an unbiased estimator, we use the fact that the within and across variation sum to the total variance. Again, incorporating the assumption that the weights are uncorrelated with income (see footnote 4) for the purpose of bias adjustment, the unbiased estimate of the within variance is:

$$WI = \sum_j \frac{M_j}{M} \frac{1}{1 - \sum_i w_{ij}^2} \sum_i w_{ij} (y_{ij} - \bar{y}_j)^2$$

and the total variance is

$$TO = \sum_j \frac{M_j}{M} \sum_i \left(w_{ij} y_{ij} - \bar{y} \right)^2$$

where $\bar{y}_j = \sum_i w_{ij} y_{ij}$. Thus, the unbiased estimate of the NSI with weighted data is

$$NSI_{gw} = \left(1 - \frac{WI}{TO} \right)^{1/2}$$

where the subscript g indicates, as before, that this measure has been count-and-weight corrected. As with $H_{gw,}$ this will be called the "corrected" estimate, unless there is a need for further clarification. Note that in the absence of variation in household weights $w_{ij} = \frac{1}{N_j}$ and both expressions reduce to the corresponding expressions in Logan et al. (2018) so the count-corrected and corrected estimates will be the same.

But the NSI estimates in Logan et al. (2018) in fact correspond neither to NSI_{bw} nor to NSI_{gw} . They are based on grouped data, which implicitly incorporate sample weights, but correct only for the sample size. They do not correct for the effective sample size, which is lower than the actual sample size due to the differential weights. Formally, those estimates are

$$NSI_{cw} = \left(1 - \frac{\sum_j \frac{M_j}{M} \frac{1}{1-(1/N_j)} \sum_i w_{ij}(y_{ij} - \bar{y}_j)^2}{\sum_j \frac{M_j}{M} \sum_i \left(w_{ij} y_{ij} - \bar{y} \right)^2} \right)^{1/2}$$

where the subscript *c* indicates the estimate is corrected for counts but not for the effective sample size. Note that the only difference between NSI_{gw} and NSI_{cw} is the replacement of the $\sum w_{ij}^2$ term in the numerator in the former with $1/N_j$ in the latter. Because $1/N_j < \sum w_{ij}^2$ as before it follows that $NSI_{cw} > NSI_{gw}$. The corrected *NSI* should be lower than the corresponding figures that only correct for sample size counts. The same approach can be used to produce corrected measures of R by replacing y_{ij} with its percentile in the population distribution, and to estimate, for example, R_{10} by replacing y_{ij} with an indicator of whether income is above the 10th percentile in the population.[5]

4. Consequences of count and weight correction

A way to summarize the impact of the two forms of bias discussed above is to show how they affect estimates of change in income segregation over time. We do this in Figure 1, which plots the estimated change in one segregation measure for all families, H, between 2000 and 2007-2011 using household-level income data in the FSRDC and applying our methods of correction. The figure displays three estimates for every one of the 95 metros studied here. For every metro it shows the change in the uncorrected estimate of H, the count-corrected estimates, and the final estimate that incorporates corrections for both sample counts and weighting. The horizontal axis arrays metros according to the change in the final estimate. Thus, the dots along

[5] Note that a fully corrected estimate of R^R could be constructed by weight-correcting R_p for every value of p and then computing the p(1-p) weighted average of the weight-corrected values.

the 45-degree straight line represent the corrected estimates of the change. The average values of H (see Table 1 below) are around .123 with a standard deviation of .026. Figure 1 shows that most metros experienced change within a range of -.01 and +.01, averaging change closer to zero.

Figure 1 about here

We compared corrected and uncorrected estimates in the following way. The plus signs in the figure represent the original uncorrected estimates of change in H. The vertical distance between a plus sign and the corresponding corrected value reveals the total bias from both sources for this metro. Note that in every case the bias is positive – the uncorrected estimates show more increase in H than do the corrected estimates. In many cases where H actually declined, the uncorrected value of H increased. Where H increased, the uncorrected value increased more.

The figure also shows (as hollow circles) the estimates after correcting only for the reduced sample count in the post-2000 data. These values are intermediate between the corrected and uncorrected estimates, but they are also in a positive direction in every case.

We make three observations about these results. First, the count-alone corrections only address about 60% of the bias in the raw estimates. The (unweighted) mean bias correction only for counts is .0068 while the mean total bias is .0114. Thus, estimates of the change in segregation by Logan et al. (2018) and Reardon et al. (2018) that corrected only for sample counts still overstated the growth in income segregation over this interval. Second, the three groups of points are roughly parallel. This indicates that the ranking of changes in segregation estimates are not substantially affected by the process of bias correction. Third, the fraction of estimates lying above zero *is substantially affected* by the process of bias correction. While 93%

of the uncorrected observations lie above zero (the dotted line) only 56% of the count-corrected observations do. Put another way, all the uncorrected observations in the northwest corner of the graph are misclassified as having growing income segregation estimates even though the corrected-estimates show decreased segregation.

Results: uncorrected and corrected estimates

Let us now summarize our methodological conclusions. Relying on grouped income data introduces errors in estimation of several standard measures of income segregation. For this reason it is preferable to work directly with the original unit-level household data that are accessible in the FSRDC. There is systematic bias associated with the size of samples and with reliance on weighted data (noting that all census or ACS sample data are weighted). These biases can support incorrect conclusions about trends in segregation, but they can be reliably estimated for every one of the income segregation measures that we consider here. We have implemented these corrections using the unit-level income data in the FSRDC. Here we present the results for all years between Census 1980 and ACS 2012-2016 for family households on different types.

The average values (weighted by the number of households) of the largest metropolitan regions are reported in Tables 1-6.[6] Each table includes the uncorrected values calculated from unit-level family-household data followed by the corrected values, in order to gauge how the bias corrections have altered the observed results. Although we would expect sampling variation to affect estimates for any given metro, we are confident that the average across all large metros is close to the true value. Tables 1 and 2 present results separately for all families and for families

[6] We have also calculated changes in the unweighted averages, yielding very similar patterns. We prefer to weight by the number of families, so that the statistic reflects the experience of the average family, the average family with children, or the average white, black, or Hispanic family with or without children, in large metropolitan regions.

with children, providing a test of the influence of children on locational choices. Table 1 presents the overall summary measures across the entire income distribution (H, R, and NSI). Table 2 presents the measures corresponding to the separation of the bottom tenth (H_{10} and R_{10}) and top tenth (H_{90} and R_{90}) of families. Tables 3 and 4 offer parallel sets of results for white, black, and Hispanic families. Finally, Tables 5 and 6 report results for families with children of each specific racial/ethnic group.

All families and families with children

Table 1 replicates findings in previous studies that showed a spike in income segregation between 1980 and 1990 for all measures and both types of families. In these decades, when the decennial census provided a full one-in-six sample of income data in every year, the uncorrected estimates are higher than the corrected estimates, but both increased substantially. Note that if we relied on the uncorrected measures, it would appear that income segregation for all families increased again between 2000 and 2007-11 and then stabilized. The corrected values show that neither H nor R increased after 1990, while NSI vacillated (down by 2000, then up, then down again). By these measures, the general rise in income segregation that has previously been reported did not occur.

Table 1 about here

However, a different result is found for families with children. For these families, the corrected measures show that income segregation continued in each interval through 2012-16. This result is consistent with the trend reported by Owens (2016), although the magnitude of these gains is much reduced after correction. For example, the uncorrected H for families with children increased from .170 in 1990 to .215 in 2012-16 (up .045), while the corrected H rose far more slowly from .156 to .176 (up .020, about half as much).

Table 2 focuses on the upper and lower ends of the income distribution, relying on the dichotomies of the upper (or lower) 10% of the population vs all others to provide more detail about the patterns of change. Let us focus first on the actual trends as reflected in the corrected values. In the 1980s, when overall income segregation was rising strongly, segregation of the poor and segregation of the affluent both rose substantially as measured by either H or R. Levels of segregation and increases were higher for families with children than for all families. After 1990 the levels stabilize or decline.

- For all families H_{10} and R_{10} (segregation of the lower tenth) dropped during 1990-2000. H_{10} and R_{10} then stabilized or continued to decline through 2012-16. At the end of these years, these measures were actually lower than they had been in 1980.
- Again looking at all families, H_{90} and R_{90} (segregation of the upper tenth) stabilized or declined slightly through 2012-16, but the final levels remained higher than in 1980.
- Trends are somewhat different for families with children. H_{10} and R_{10} both declined steadily after 1990. But H_{90} and R_{90} rose again during 1990-2000, then stabilized.

These patterns of change in Tables 1-2 challenge recent interpretations. Based on the uncorrected estimates, one could describe a fairly steady rise in overall income segregation (Table 1) that coincides with rising income inequality. The upward trend appeared to be most striking and consistent from decade to decade for families with children (also Table 1). Then turning to Table 2, rising segregation seems especially clear for affluent families with children. These trends could be interpreted in terms of the motivations and behaviors of parents whose locational decisions increasingly seek advantaged communities for their children – especially affluent parents – which is Owens' interpretation. However, the corrected results do not fit this narrative as well. After 1990 there was a continuing increase in H, R and NSI for families with

children, though smaller than previously reported. Yet this post-1990 trend does not appear either for segregation of the poor or segregation of the affluent families with children. From these results we infer that the locational shifts evident in Table 1 were occurring more toward the middle of the income distribution.

Table 2 about here

Race-specific patterns

We turn now to findings for whites, blacks, and Hispanics. Note that in the previous tables, some portion of income segregation was due to racial/ethnic segregation since black and Hispanic families have lower incomes than white families. The race-specific measures in the following tables consider each group separately, so they measure the degree to which white (or black or Hispanic) families are segregated by income from other white (or black or Hispanic) families. In these tables the mean values and standard deviations of segregation estimates are group-specific, and the measures of segregation of affluence and poverty refer to the top and bottom tenths of that group's income distribution.

Because many census tracts have few black or Hispanic residents even in metros with large minority populations, we expected bias corrections for these groups to be especially large, particularly after 2000. An example is provided in Figure 2, which displays the uncorrected and corrected estimates of H for whites and blacks. The uncorrected estimates are upwardly biased, much more so for blacks than for whites. After 2000 the corrected value of H for whites declines modestly, while the uncorrected estimate increases. Among blacks the corrected value of H remains nearly unchanged, while the uncorrected value spikes remarkably from .128 to .173, equivalent to nearly than 2.0 standard deviations. This discrepancy leads to widely divergent understanding of these trajectories. From the uncorrected data it appears that income segregation

among blacks was much higher than among whites, and while there was a mild increase after 2000 among whites the jump after 2000 among blacks was enormous. After correction, we conclude that income segregation among blacks was only modestly higher than among whites, and both remained rather stable post-2000 after rising in the 1980-1990 decade.

Figure 2 about here

Table 3 presents the full set of values of H, R, and NSI for the three groups. The upward bias in uncorrected estimates of H found in Figure 1 for whites and blacks is replicated for R and NSI, as is the post-2000 spike for blacks. In these respects, the results for Hispanics follow the same pattern. Now let us focus on the trends revealed by the corrected measures. For every group and every measure (with a small inconsistency for Hispanic NSI) there were substantial increases between 1980 and 1990. This is what we found previously for the total population. If we then compare the 1990 value to the final value in 2012-16, we do not find consistent increases:

- For whites, H declined from .095 to .090. R declined from .170 to .158. NSI declined from .136 to .132.
- For black families, H declined from .108 to .100. R declined from .186 to .175. NSI declined from .127 to .108. (In this case, however, NSI fluctuated, rising in 2007-2011 before dropping again. We cannot account for this inconsistency.)
- For Hispanic families, H remained at .091, after dropping from 1990 to 2000, then rising back to the 1990 level. R remained at .159, also after dropping from 1990 to 2000, then rising back to the 1990 level. NSI also fluctuated, but this is the one case where NSI ended up higher in 2012-16 than it had been in 1990.

From these findings we can conclude that previously reported results for these groups overstated the differences between whites and blacks/Hispanics. Income segregation was somewhat higher among blacks than among either whites or Hispanics. Previous reports also overstated the tendency for income segregation to rise for any of them. In fact, income segregation among white and black families declined after 1990, while income segregation among Hispanic families was the same in 2012-16 as in 1990 for H and R.

Table 3 about here

Table 4 repeats our analysis of segregation of the affluent and of the poor for all families in each group, reporting trends for the lower income (H_{10} and R_{10}) and upper income (H_{90} and R_{90}) segments. Because there are so many comparisons to make in this table, we will not discuss the uncorrected measures, including them here only for reference. Consistent with Table 3 for the overall income segregation measures, the segregation of both poverty and affluence increased from 1980 to 1990 for all three groups. After 1990:

- The average levels of all these measures at the ends of the income distribution were stable (for the bottom 10%) or declining (for the upper 10%) for white families.
- For black families there was some decline for lower income families, more clearly for H_{10} than for R_{10} and also for affluent families, more clearly for H_{90} than for R_{90}.
- For Hispanic families, there is little trend for poor families, but there was a substantial increase in income segregation of the affluent as measured by H_{90} and R_{90}.

In relation to previous reports, the main conclusion from Table 4 is that instead of a generalized increase in segregation of either the affluent or the poor after 1990, there was a decline for whites and blacks and an increase only for the higher-income segment of Hispanic families.

Table 4 about here

As a final step we report group-specific results for families with children in Tables 5-6. Again we focus only on the corrected measures. Recall that we found evidence of increasing income segregation for families with children in Table 1 based on measures for the full income distribution (H, R, and NSI) but not for the upper and lower segments (($H_{10,}$ R_{10}, H_{90} and R_{90}). Is there, however, a tendency for increasing segregation for families within racial/ethnic groups?

As shown in all the tables up to this point, segregation rose from 1980 to 1990. With respect to the full income distribution after 1990 (measured by H, R, and NSI), the answer is mixed:

- After 1990 there is little trend among white families with children. H rose in the 1990s (from .118 to .127), but then stabilized. R also rose in the 1990s (from .207 to .218), then again to .223 in 2012-16. NSI rose in the 1990s (from .166 to .180), but then declined to .177 by 2012-16.
- There is a more substantial upward trend for black families with children, shown most clearly in the increases after 2000. For example, NSI rose from .124 in 2000 (after declining in the 1990s to .170 in 2012-16.
- There is a similar upward trend for Hispanic families with children, reaching its lowest level in 2000 and then rising strongly after that time.

Table 5 about here

Trends for the poorest and most affluent families with children are reported in Table 6. For all groups and measures there was a strong increase in the 1980s. Here again the clearest evidence of rising segregation among families with children is among Hispanics, and specifically the separation of the most affluent Hispanics from others.

- For whites, H10 and R10 both declined slightly in the 1990s but then rose moderately after 2000. H90 and R90, in contrast, were both on the decline after 2000.
- For blacks, H10 declined after 2000 while R10 remained stable. H90 and R90 changed little after rising in the 1990s.
- For Hispanics, H10 and R10 were stable after 1990, ending at about the same level as they began. But H90 and R90 both had strong upward trajectories through this whole period, starting in 1980 and continuing through 2012-16. H90 rose from .092 in 1980 consistently through to .137 in 2012-16, while R90 rose from .073 to .117.

Table 6 about here

Discussion and conclusion

This study contributes to two kinds of goals. One is substantive, to document trends in income segregation in U.S. metropolitan areas since 1980. We compare patterns between all families and families with children, and between white, black, and Hispanic families, and look separately at trends for the upper and lower tails of the income distribution. The other purpose is methodological. We call attention to the effects of stratified sampling on measures of spatial inequality, and especially to the problems associated with the shift in data sources from the long-form samples of decennial censuses to the smaller samples of the American Community Survey.

Substantive findings

We find that social scientists cannot rely on published tract-level data to discover the real levels and trends in income segregation. After recognizing and seeking to take into account the upward bias associated with smaller samples after 2000, two research teams (Logan et al. 2018 and Reardon et al. 2018) estimated that the post-2000 increase was about half of what had previously been reported. We take advantage of confidential census data files at the individual

family level, obviating the need to interpolate income distributions within the categories that are used in public tract data and allowing us to account for variance in sampling probabilities. The results show that not only have increases in income segregation been overstated in past studies, but for several categories of families there was no change or actual declines after 1990.

An important caveat is that removing the bias associated with sample size does not solve all concerns with sampling variability. Current ACS data come from a single sample that can be very limited in many census tracts, especially for subgroups of the population. Any income segregation measure aggregated up from tract-level distributions can be no more than an estimate of the actual population value. Researchers should be cautious in interpreting results for any single metropolitan area, since there is a possibility that the estimate in a given year is too large or too small and that observed changes over time reflect sampling variability rather than real change. Nevertheless we have high confidence in the average value of segregation over many metros, because in that case random sampling errors, positive and negative, can be expected to cancel each other out.

After surging in the 1980s, family income segregation has undergone some ups and downs but not increased, and it has declined for families headed by non-Hispanic whites and for affluent white families. Segregation for families with children who have been described as especially conscious of the advantages of moving to places with more resources continued the trend toward higher levels on H, R, and NSI through 2012-2016. However, segregation of affluent families with children was very stable. This result undermines the interpretation of changes for families with children that they result from the most advantaged parents seeking special place-based advantages for their children. The finding that this measure of segregation of

affluent families with children increased only for Hispanics but not for whites (who presumably have the most locational options) points to a more group-specific process.

Summary statistics like these do not reveal who is moving to more separate neighborhoods at each time point, and we cannot draw strong conclusions from them about the processes at work within metropolitan neighborhoods. The general conclusion is that rather than focusing on why income segregation seems to be rising in parallel with growing income inequality, scholars need to give more attention to why it may not. There are many directions to look. In the post-2000 period one might consider the possible effects of the Great Recession and foreclosure crisis that occurred in the middle of the 2007-11 ACS period. As income inequality continued to rise, many people lost jobs, many lost their homes, many were forced to postpone moves by changing mortgage requirements, and there was a temporary steep decline in the value of non-home assets held by the most affluent households. We are not in a position to fit these pieces together into a coherent narrative, and this remains a challenge for future research.

Our findings for black and Hispanic families are intriguing in light of expectations that even a modest opening up of opportunities in the housing market might motivate and enable some minority families to seek more advantaged neighborhoods. We find such a pattern for Hispanics, but not so clearly for blacks. The hypothesis of an exodus of more affluent minorities from income-diverse neighborhoods after 1980 needs a more direct test through analyses of residential mobility.

Implications of bias from smaller samples

This research adds to concerns that others have expressed about the use of tract-level data from the American Community Survey. The Census Bureau has made efforts to educate users on the potentially large sampling variation in point estimates (such as the median value of

income or the percent of residents born abroad) for census tracts, and it now routinely disseminates measures of standard errors around these estimates. Fortunately researchers have begun to notice these standard errors, and fortunately the point estimates are unbiased. That is, they may be far from the population value in a given tract, but they will cluster randomly around the true value. We draw attention to a different phenomenon associated with sampling variation. Standard measures of spatial inequality such as the measures of income segregation analyzed here have an inherent upward bias when based on samples and the bias is greater when the sample size is smaller and where sampling is stratified. This is why income segregation was observed to increase again for all families after 2000 after seeming to moderate in the 1990s. It is also why differences in levels and trends between whites and minorities were especially exaggerated after 2000.

Measures of income segregation are often included in multivariate analyses of other outcomes. In a cross-sectional study the previously reported metro-level estimates may perform well. In supplementary analyses not reported here, we found very high cross-sectional correlations between uncorrected and corrected measures for the whole population ($r > .95$). This indicates that studies of the correlates of income segregation in a given year are likely to be only slightly affected by biased measures. Studies of specific segments of the population, however, should be attentive to the average sample size for a given subgroup, which may vary greatly across metros. We also found lower correlations in change over time between the corrected and uncorrected measures (in the range of .75 to .85), suggesting that there is greater potential for error in longitudinal analyses.

For the 95 largest metros we recommend the use of the corrected estimates analyzed here, whether for measures based on entropy (H) or variance (NSI, R). These different measures

typically trend in the same direction, but it is prudent not to rely only on one of them. Data for smaller metros may also be approved for disclosure in the future. Finally, for researchers who are able to gain access to the original sample data in the Census Bureau's Federal Statistical Restricted Data Center (FSRDC) system, the programs used to calculate measures and implement corrections will available from the authors. There are significant obstacles to FSRDC use, including their geographic location (they are spread unevenly around the country), their cost (sometimes free to faculty of hosting institutions but with fees of as much as $20,000 per year to others), the time required for an individual to gain special sworn status and for a proposed research project to be approved (sometimes six months to a year), the learning process of how to find documentation and use confidential data sets through the FSRDC's computing system, the difficulty of evaluating interim findings that cannot be printed but only viewed on a terminal screen, and a learning process associated with disclosure reviews. There is a clear rationale for every one of these obstacles and therefore no simple solution. Nevertheless, as we discover that some kinds of studies that rely extensively on census data can no longer be carried out in familiar ways, scholars will increasingly need to learn how to make effective use of this data resource.

Acknowledgements. This research was supported by the Sociology Program of the National Science Foundation (grant 1756567) and National Institutes of Health (1R21HD078762-01A1). The Population Studies and Training Center at Brown University (P2CHD041020) provided general support. We thank Todd Gardner of the U.S. Bureau of the Census for his assistance in working with census data through the FSRDC network. Any opinions and conclusions expressed herein are those of the author(s) and do not necessarily represent the views of the U.S. Census Bureau. All results have been reviewed to ensure that no confidential information is disclosed.

Authors' Contributions. All authors contributed to the study concept and design. Data preparation in the FSRDC was performed by Todd Gardner, Charles Zhang, and Hongwei Xu. Methodological innovations and programming were developed primarily by Andrew Foster. Analysis of income segregation patterns was carried out primarily by John Logan. The first draft of the manuscript was written by John Logan, and all authors commented on subsequent versions of the manuscript. All authors read and approved the final manuscript.

Data Availability. All data sets used are held within the FSRDC under strict confidentiality rules. Segregation indices have been approved for disclosure by the Census Bureau and are posted on the Brown University website of the Initiative on Spatial Structures in the Social Sciences developed by John Logan: https://s4.ad.brown.edu/projects/diversity/Data/data.htm.

Compliance with ethical standards

Ethics and Consent The authors report no ethical issues.

Conflict of Interest The authors declare no conflicts of interest.

References

Bischoff, Kendra and Sean F. Reardon. 2014. Residential Segregation by Income, 1970–2009. Pp. 208-233 in John R. Logan (editor), *Diversity and Disparities: America Enters a New Century*. New York: Russell Sage Foundation.

Florida, Richard and Charlotta Mellander. 2015. "Segregated City: The Geography of Economic Segregation in America's Metros" Report by the Martin Prosperity Institute, University of Toronto. Accessed 11/24/15 at http://martinprosperity.org/media/Segregated%20City.pdf.

Fry, Richard and Paul Taylor. 2012. "The Rise of Residential Segregation by Income" Report by the Pew Research Center. Accessed 11/24/15 at http://www.pewsocialtrends.org/files/2012/08/Rise-of-Residential-Income-Segregation-2012.2.pdf.

Jargowsky, Paul A. 1996. "Take the Money and Run: Economic Segregation in U.S. Metropolitan Areas." *American Sociological Review* 61(6): 984–98.

Logan, John R., Andrew Foster, Jun Ke, And Fan Li. 2018. "The Uptick in Income Segregation: Real Trend or Random Sampling Variation?" *American Journal of Sociology* 124:

Logan, John R., Weiwei Zhang, and Miao Chunyu. 2015. "Emergent Ghettos: Black Neighborhoods in New York and Chicago, 1880-1940" *American Journal of Sociology* 120(4):1055-1094.

Massey, Douglas S. and Mitchell L. Eggers. 1993. "The Spatial Concentration of Affluence and Poverty during the 1970s." *Urban Affairs Quarterly* 29:299-315.

Massey, Douglas S., and Mary J. Fischer. 2003. "The Geography of Inequality in the United States, 1950–2000." *Brookings-Wharton Papers on Urban Affairs*, 1–40.

Mayer, Susan E. 2001. "How Growth in Income Inequality Increased Economic Segregation," Joint Center for Poverty Research Working Paper 230.

Owens, Ann. 2016. "Inequality in Children's Contexts: Income Segregation of Households with and without Children" *American Sociological Review* 81(3) 549–574.

Reardon, Sean F. and Kendra Bischoff. 2011. "Income Inequality and Income Segregation Source" *American Journal of Sociology*, 116:1092-1153.

Reardon, Sean F., Kendra Bischoff, Ann Owens, and Joseph B. Townsend. 2018. "Has Income Segregation Really Increased? Bias and Bias Correction in Sample-Based Segregation Estimates" *Demography* 55:2129-2160.

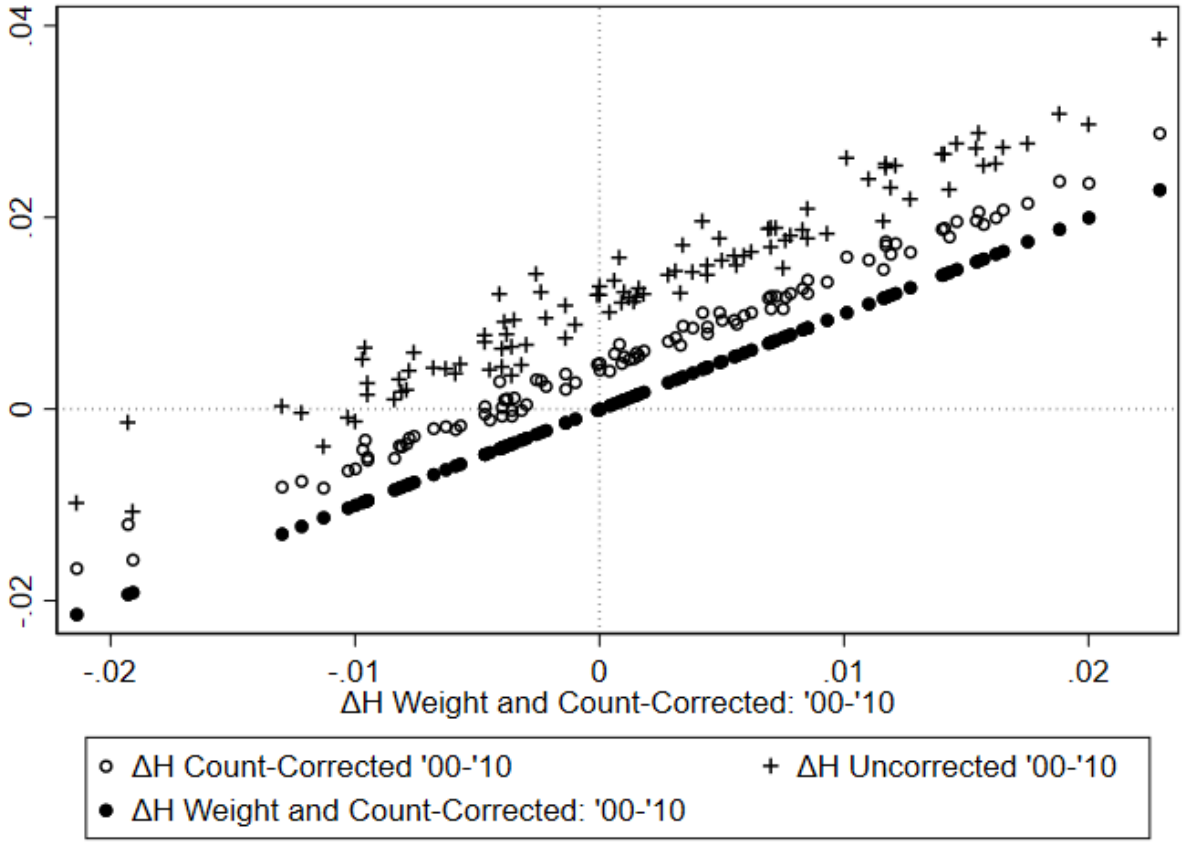
Reardon, Sean F., Lindsay Fox and Joseph Townsend. 2015. "Neighborhood Income Composition by Household Race and Income, 1990-2009" *Annals of the American Academy of Political and Social Science* 660: 78-97.

Watson, Tara. 2009. "Inequality and the Measurement of Residential Segregation by Income." *Review of Income and Wealth* 55(3): 820–44.

Wilson, William J. 1987. *The Truly Disadvantaged: The Inner City, the Underclass, and Public Policy*. Chicago: University of Chicago Press.

Yang, Rebecca, and Paul A. Jargowsky. 2006. "Suburban Development and Economic Segregation in the 1990s." *Journal of Urban Affairs* 28:253–73.

Figure 1: Estimates of change in H for 95 metros between 2000 and 2007-2011: uncorrected, corrected for sample count, and corrected



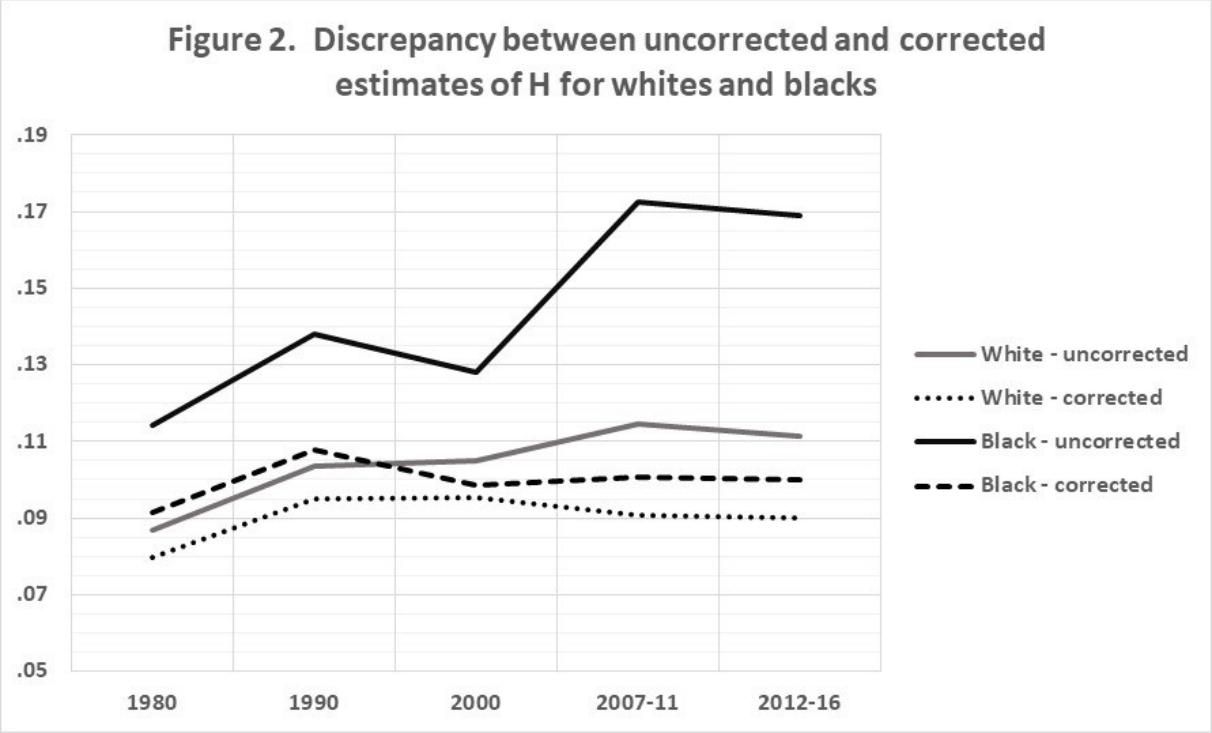


Figure 2. Discrepancy between uncorrected and corrected estimates of H for whites and blacks
.19
.17
.15
.13
.11
.09
.07
.05
1980
1990
2000
2007-11
2012-16
White - uncorrected
White - corrected
Black - uncorrected
Black - corrected

Table 1. Income segregation over time, all families and families with children.										
	1980		1990		2000		2007-11		2012-16	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
I. Uncorrected										
All families										
H	0.112	*0.027*	0.132	*0.028*	0.131	*0.026*	0.142	*0.027*	0.140	*0.026*
R	0.198	*0.047*	0.229	*0.048*	0.228	*0.045*	0.239	*0.045*	0.236	*0.045*
NSI	0.156	*0.037*	0.165	*0.039*	0.163	*0.037*	0.186	*0.039*	0.180	*0.039*
Families with children										
H	0.140	*0.033*	0.170	*0.033*	0.178	*0.031*	0.212	*0.033*	0.215	*0.033*
R	0.241	*0.055*	0.285	*0.053*	0.298	*0.050*	0.334	*0.052*	0.339	*0.053*
NSI	0.190	*0.045*	0.211	*0.046*	0.218	*0.042*	0.258	*0.043*	0.261	*0.045*
II. Corrected										
All families										
H	0.106	*0.027*	0.125	*0.028*	0.123	*0.026*	0.123	*0.026*	0.124	*0.026*
R	0.192	*0.047*	0.223	*0.048*	0.221	*0.045*	0.224	*0.046*	0.223	*0.045*
NSI	0.150	*0.037*	0.159	*0.039*	0.156	*0.038*	0.171	*0.040*	0.167	*0.039*
Families with children										
H	0.129	*0.032*	0.156	*0.032*	0.164	*0.030*	0.169	*0.032*	0.176	*0.032*
R	0.231	*0.055*	0.274	*0.053*	0.286	*0.051*	0.303	*0.054*	0.311	*0.055*
NSI	0.180	*0.045*	0.198	*0.046*	0.205	*0.042*	0.227	*0.045*	0.232	*0.046*

Table 2. Segregation of poverty and affluence over time , all families and families with children.										
	1980		1990		2000		2007-11		2012-16	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
I. Uncorrected										
All families										
H_{10}	0.114	*0.036*	0.137	*0.042*	0.121	*0.033*	0.136	*0.029*	0.126	*0.028*
R_{10}	0.090	*0.032*	0.109	*0.038*	0.093	*0.029*	0.101	*0.025*	0.094	*0.024*
H_{90}	0.144	*0.036*	0.177	*0.039*	0.172	*0.035*	0.192	*0.039*	0.185	*0.036*
R_{90}	0.115	*0.032*	0.140	*0.034*	0.135	*0.032*	0.142	*0.034*	0.136	*0.032*
Families with children										
H_{10}	0.141	*0.041*	0.172	*0.051*	0.156	*0.041*	0.196	*0.034*	0.186	*0.034*
R_{10}	0.111	*0.037*	0.134	*0.047*	0.117	*0.037*	0.140	*0.031*	0.132	*0.030*
H_{90}	0.167	*0.042*	0.213	*0.044*	0.225	*0.042*	0.264	*0.046*	0.260	*0.045*
R_{90}	0.134	*0.037*	0.172	*0.041*	0.178	*0.040*	0.196	*0.043*	0.191	*0.042*
II. Corrected										
All families										
H_{10}	0.104	*0.036*	0.126	*0.041*	0.110	*0.032*	0.107	*0.028*	0.102	*0.027*
R_{10}	0.083	*0.032*	0.101	*0.038*	0.085	*0.028*	0.081	*0.024*	0.077	*0.023*
H_{90}	0.134	*0.036*	0.166	*0.038*	0.161	*0.035*	0.163	*0.037*	0.161	*0.035*
R_{90}	0.110	*0.032*	0.134	*0.034*	0.129	*0.032*	0.127	*0.034*	0.123	*0.031*
Families with children										
H_{10}	0.124	*0.040*	0.151	*0.049*	0.134	*0.040*	0.131	*0.034*	0.127	*0.033*
R_{10}	0.099	*0.037*	0.120	*0.046*	0.103	*0.037*	0.096	*0.030*	0.092	*0.028*
H_{90}	0.150	*0.041*	0.192	*0.042*	0.203	*0.041*	0.199	*0.043*	0.200	*0.044*
R_{90}	0.124	*0.037*	0.161	*0.040*	0.166	*0.040*	0.164	*0.043*	0.162	*0.042*

Table 3. Income segregation over time, by race and Hispanic origin (all families)										
	1980		1990		2000		2007-11		2012-16	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
I. Uncorrected										
White										
H	0.087	*0.018*	0.104	*0.020*	0.105	*0.020*	0.115	*0.023*	0.111	*0.023*
R	0.150	*0.030*	0.178	*0.033*	0.178	*0.033*	0.181	*0.036*	0.176	*0.036*
NSI	0.133	*0.029*	0.143	*0.033*	0.147	*0.032*	0.154	*0.032*	0.148	*0.031*
Black										
H	0.114	*0.022*	0.138	*0.027*	0.128	*0.029*	0.173	*0.040*	0.169	*0.040*
R	0.182	*0.034*	0.212	*0.038*	0.199	*0.039*	0.235	*0.048*	0.233	*0.049*
NSI	0.145	*0.042*	0.160	*0.045*	0.116	*0.035*	0.205	*0.046*	0.177	*0.048*
Hispanic										
H	0.120	*0.037*	0.128	*0.034*	0.111	*0.026*	0.160	*0.039*	0.154	*0.039*
R	0.176	*0.036*	0.191	*0.034*	0.169	*0.025*	0.216	*0.036*	0.214	*0.037*
NSI	0.170	*0.042*	0.167	*0.041*	0.110	*0.035*	0.218	*0.041*	0.205	*0.046*
II. Corrected										
White										
H	0.080	*0.017*	0.095	*0.019*	0.095	*0.019*	0.091	*0.020*	0.090	*0.020*
R	0.144	*0.029*	0.170	*0.033*	0.169	*0.032*	0.161	*0.035*	0.158	*0.035*
NSI	0.126	*0.029*	0.136	*0.033*	0.139	*0.032*	0.136	*0.032*	0.132	*0.031*
Black										
H	0.091	*0.019*	0.108	*0.023*	0.099	*0.022*	0.101	*0.025*	0.100	*0.025*
R	0.161	*0.033*	0.186	*0.038*	0.173	*0.038*	0.174	*0.044*	0.175	*0.044*
NSI	0.123	*0.037*	0.127	*0.039*	0.086	*0.029*	0.131	*0.039*	0.108	*0.041*
Hispanic										
H	0.083	*0.018*	0.091	*0.016*	0.079	*0.013*	0.090	*0.017*	0.091	*0.017*
R	0.143	*0.029*	0.159	*0.028*	0.141	*0.024*	0.154	*0.031*	0.159	*0.031*
NSI	0.123	*0.030*	0.121	*0.031*	0.074	*0.024*	0.135	*0.033*	0.130	*0.037*

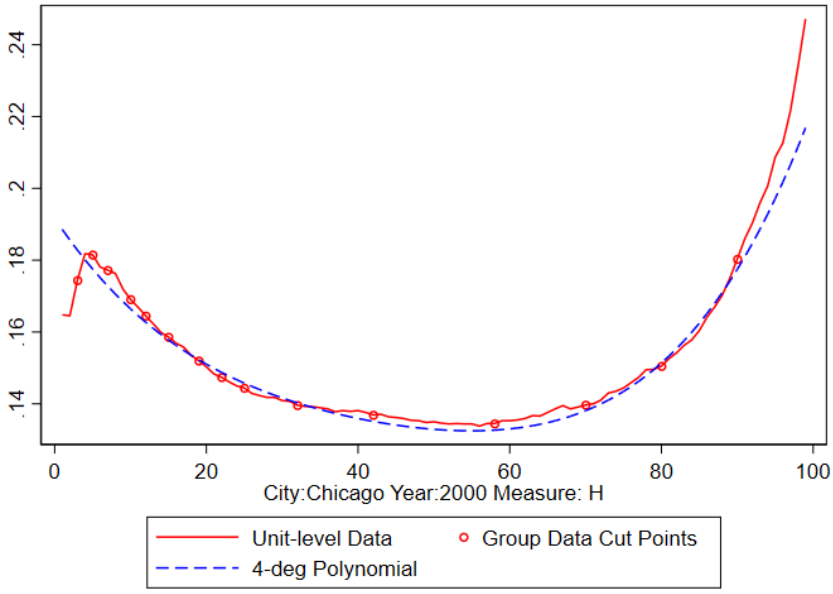
Table 4. Segregation of poverty and affluence over time, by raceand Hispanic origin, all families										
	1980		1990		2000		2007-11		2012-16	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
I. Uncorrected										
White families										
H_{10}	0.069	*0.014*	0.085	*0.020*	0.082	*0.021*	0.104	*0.024*	0.100	*0.024*
R_{10}	0.051	*0.012*	0.063	*0.017*	0.060	*0.018*	0.076	*0.020*	0.074	*0.020*
H_{90}	0.135	*0.034*	0.165	*0.035*	0.164	*0.033*	0.176	*0.034*	0.166	*0.031*
R_{90}	0.108	*0.029*	0.130	*0.030*	0.126	*0.027*	0.126	*0.027*	0.118	*0.025*
Black families										
H_{10}	0.091	*0.027*	0.130	*0.034*	0.126	*0.034*	0.185	*0.044*	0.182	*0.042*
R_{10}	0.059	*0.020*	0.087	*0.026*	0.083	*0.025*	0.122	*0.032*	0.120	*0.031*
H_{90}	0.130	*0.029*	0.152	*0.030*	0.145	*0.034*	0.206	*0.044*	0.205	*0.047*
R_{90}	0.098	*0.025*	0.115	*0.026*	0.111	*0.026*	0.153	*0.038*	0.153	*0.040*
Hispanic families										
H_{10}	0.100	*0.043*	0.111	*0.043*	0.098	*0.034*	0.155	*0.049*	0.144	*0.046*
R_{10}	0.059	*0.026*	0.068	*0.026*	0.058	*0.020*	0.096	*0.032*	0.088	*0.029*
H_{90}	0.152	*0.042*	0.170	*0.042*	0.153	*0.035*	0.224	*0.043*	0.222	*0.046*
R_{90}	0.121	*0.034*	0.135	*0.033*	0.124	*0.028*	0.175	*0.038*	0.174	*0.039*
II. Corrected										
White families										
H_{10}	0.058	*0.013*	0.072	*0.019*	0.067	*0.019*	0.067	*0.019*	0.068	*0.020*
R_{10}	0.043	*0.011*	0.053	*0.016*	0.049	*0.016*	0.049	*0.016*	0.050	*0.017*
H_{90}	0.124	*0.032*	0.151	*0.033*	0.149	*0.031*	0.139	*0.030*	0.134	*0.029*
R_{90}	0.102	*0.029*	0.123	*0.030*	0.118	*0.027*	0.108	*0.027*	0.102	*0.025*
Black families										
H_{10}	0.056	*0.021*	0.084	*0.030*	0.080	*0.028*	0.074	*0.028*	0.076	*0.023*
R_{10}	0.038	*0.018*	0.061	*0.026*	0.058	*0.024*	0.057	*0.025*	0.058	*0.022*
H_{90}	0.095	*0.025*	0.106	*0.022*	0.099	*0.022*	0.096	*0.030*	0.099	*0.035*
R_{90}	0.073	*0.022*	0.084	*0.021*	0.078	*0.020*	0.081	*0.028*	0.081	*0.032*
Hispanic families										
H_{10}	0.042	*0.017*	0.053	*0.016*	0.048	*0.013*	0.047	*0.016*	0.047	*0.014*
R_{10}	0.027	*0.013*	0.037	*0.014*	0.032	*0.011*	0.035	*0.014*	0.033	*0.012*
H_{90}	0.095	*0.026*	0.112	*0.025*	0.103	*0.024*	0.116	*0.031*	0.125	*0.030*
R_{90}	0.076	*0.024*	0.090	*0.022*	0.085	*0.022*	0.096	*0.028*	0.101	*0.027*

Table 5. Income segregation over time, by race and Hispanic origin, families with children										
	1980		1990		2000.000		2007-11		2012-16	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
I. Uncorrected										
White										
H	0.106	*0.022*	0.136	*0.025*	0.147	*0.024*	0.177	*0.029*	0.182	*0.030*
R	0.176	*0.034*	0.222	*0.039*	0.235	*0.037*	0.260	*0.041*	0.265	*0.043*
NSI	0.157	*0.034*	0.182	*0.040*	0.195	*0.034*	0.213	*0.034*	0.215	*0.036*
Black										
H	0.131	*0.025*	0.161	*0.031*	0.166	*0.035*	0.243	*0.048*	0.244	*0.047*
R	0.203	*0.037*	0.237	*0.043*	0.242	*0.049*	0.304	*0.064*	0.302	*0.064*
NSI	0.161	*0.045*	0.188	*0.050*	0.168	*0.050*	0.276	*0.064*	0.271	*0.056*
Hispanic										
H	0.127	*0.035*	0.135	*0.032*	0.127	*0.028*	0.195	*0.037*	0.193	*0.034*
R	0.179	*0.035*	0.194	*0.032*	0.182	*0.029*	0.248	*0.041*	0.249	*0.038*
NSI	0.171	*0.042*	0.168	*0.040*	0.121	*0.040*	0.253	*0.047*	0.248	*0.048*
II. Corrected										
White										
H	0.092	*0.020*	0.118	*0.023*	0.127	*0.022*	0.123	*0.024*	0.129	*0.026*
R	0.164	*0.033*	0.207	*0.038*	0.218	*0.036*	0.217	*0.043*	0.223	*0.045*
NSI	0.144	*0.033*	0.166	*0.039*	0.180	*0.034*	0.174	*0.037*	0.177	*0.038*
Black										
H	0.102	*0.021*	0.121	*0.028*	0.122	*0.030*	0.135	*0.038*	0.135	*0.037*
R	0.177	*0.037*	0.204	*0.045*	0.205	*0.051*	0.216	*0.068*	0.215	*0.067*
NSI	0.134	*0.041*	0.149	*0.047*	0.124	*0.046*	0.175	*0.065*	0.170	*0.057*
Hispanic										
H	0.084	*0.019*	0.090	*0.016*	0.086	*0.016*	0.103	*0.022*	0.105	*0.023*
R	0.140	*0.030*	0.155	*0.028*	0.146	*0.031*	0.169	*0.043*	0.173	*0.045*
NSI	0.120	*0.033*	0.118	*0.032*	0.076	*0.034*	0.154	*0.046*	0.150	*0.054*

Table 6. Segregation of poverty and affluence over time, by race and Hispanic origin, families with children										
	1980		1990		2000		2007-11		2012-16	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
I. Uncorrected										
White families										
H_{10}	0.094	*0.021*	0.124	*0.029*	0.124	*0.027*	0.187	*0.033*	0.191	*0.031*
R_{10}	0.072	*0.018*	0.093	*0.026*	0.092	*0.024*	0.136	*0.028*	0.139	*0.025*
H_{90}	0.156	*0.040*	0.204	*0.043*	0.214	*0.037*	0.240	*0.036*	0.234	*0.036*
R_{90}	0.125	*0.034*	0.162	*0.038*	0.163	*0.032*	0.172	*0.031*	0.165	*0.030*
Black families										
H_{10}	0.101	*0.031*	0.141	*0.037*	0.153	*0.039*	0.244	*0.046*	0.238	*0.044*
R_{10}	0.064	*0.023*	0.091	*0.028*	0.099	*0.030*	0.162	*0.037*	0.156	*0.034*
H_{90}	0.147	*0.032*	0.176	*0.034*	0.193	*0.041*	0.279	*0.049*	0.283	*0.047*
R_{90}	0.110	*0.028*	0.130	*0.029*	0.146	*0.034*	0.207	*0.046*	0.211	*0.044*
Hispanic families										
H_{10}	0.108	*0.041*	0.118	*0.042*	0.111	*0.034*	0.189	*0.049*	0.181	*0.043*
R_{10}	0.065	*0.025*	0.071	*0.027*	0.066	*0.021*	0.117	*0.034*	0.112	*0.030*
H_{90}	0.159	*0.041*	0.176	*0.037*	0.174	*0.038*	0.264	*0.042*	0.272	*0.043*
R_{90}	0.124	*0.032*	0.140	*0.031*	0.141	*0.032*	0.205	*0.040*	0.214	*0.039*
II. Corrected										
White families										
H_{10}	0.073	*0.019*	0.097	*0.027*	0.093	*0.024*	0.103	*0.026*	0.110	*0.026*
R_{10}	0.056	*0.017*	0.073	*0.024*	0.069	*0.021*	0.075	*0.024*	0.079	*0.023*
H_{90}	0.135	*0.037*	0.177	*0.039*	0.183	*0.034*	0.156	*0.034*	0.153	*0.035*
R_{90}	0.113	*0.033*	0.148	*0.037*	0.147	*0.032*	0.132	*0.032*	0.128	*0.031*
Black families										
H_{10}	0.056	*0.025*	0.080	*0.032*	0.086	*0.036*	0.077	*0.037*	0.071	*0.033*
R_{10}	0.038	*0.021*	0.056	*0.027*	0.061	*0.032*	0.062	*0.037*	0.058	*0.033*
H_{90}	0.103	*0.027*	0.115	*0.027*	0.125	*0.032*	0.112	*0.042*	0.116	*0.044*
R_{90}	0.079	*0.024*	0.089	*0.024*	0.101	*0.029*	0.104	*0.041*	0.104	*0.043*
Hispanic families										
H_{10}	0.042	*0.018*	0.049	*0.019*	0.048	*0.016*	0.047	*0.021*	0.046	*0.020*
R_{10}	0.026	*0.014*	0.033	*0.017*	0.031	*0.013*	0.035	*0.020*	0.034	*0.019*
H_{90}	0.092	*0.027*	0.108	*0.026*	0.111	*0.031*	0.122	*0.043*	0.137	*0.042*
R_{90}	0.073	*0.023*	0.088	*0.026*	0.092	*0.029*	0.105	*0.040*	0.117	*0.037*

Appendix A. Comparison of results from grouped and household data

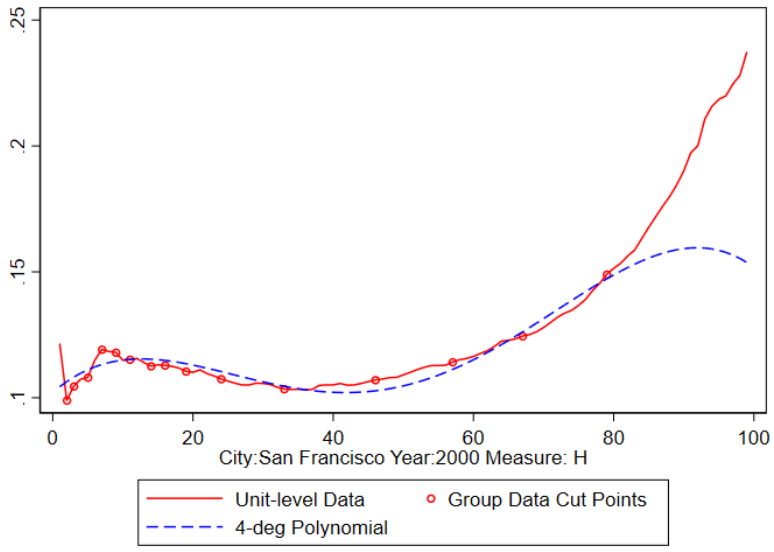
In this Appendix we show how estimates of the income segregation measure H can be distorted by the use of grouped data. To do this we have selected three large U.S. metropolitan regions as test cases: Chicago, New York, and San Francisco. In each case we aggregated unit-level family-household income data from Census 2000 to match the publicly available categories. We then calculated H_p at the 15 points in the income distribution that coincide with the cutting points in those categories. From these points we fitted a 4^{th}-order polynomial resulting in a smooth curve from which the values of values of H_p at every percentile from 1 to 100 could be derived. Finally we calculated an estimate of the overall H from these points. For this purpose we calculated estimates with no bias corrections, because the estimates with grouped data cannot be fully corrected with the available group-level data. The polynomial curve is represented by a dotted line in Appendix Figures A1 and A2 for Chicago and San Francisco (we omit New York because findings are very similar to those for Chicago). For comparison, we constructed a curve of corrected values of H_p at every centile from 1 to 99 using the confidential household-level data (referred to here as unit data). This curve is displayed as the solid line in the figures.



Appendix Figure A1: Estimating H from grouped versus unit-level data, Chicago 2000

For example, Appendix Figure A1 presents an estimate of the uncorrected H_p for every centile of p based on the household-level estimates (solid line) for Chicago in 2000. The circles along this line denote the values of p that correspond to cut points in the grouped data. These are the points on which the polynomial is based. Thus, values to the right of the last circle (at approximately the 90th percentile of the income distribution) are for levels of income above $200,000 (reflecting the top category of $200,000 and above). H is an average (weighted by entropy so either tail receives lower weight) of these centile estimates (from grouped data) or values (measured directly from unit-level household data). Overall, the solid and dashed lines are close except at the two extremes, with the solid line being above the dashed one at high centiles and below it at low centiles. Consequently, the corresponding H measures are quite similar. The H estimate based on the solid line that comes from unit-level data leads to an H of .146. H based on the polynomial curve is .145.

San Francisco (Appendix Figure A2) tells a different story. San Francisco had a relatively high fraction of households in the top income category in 2000. Consequently the highest available cutting point ($200,000 and above) falls only at the 79th percentile of the income distribution. This results in a best-fitting polynomial that is unable to project values accurately above p=.79. But in San Francisco, the "true" values of H_p rose rapidly after that point. In this case, as in Chicago, the overall estimate of H is not very different using the two data sources -- .119 versus .122. But at the 90th percentile the value of segregation (in this case, we can refer to it as H_{90}) is estimated to be .159 from the grouped data but .190 from individual household data. We suspect more generally that estimates of H from grouped data are especially vulnerable to error at the top and bottom ends of the income distribution (i.e., H_{10} and H_{90}), depending on the overall income range in the metropolitan population (shifted upwards in places like San Francisco but downwards in less advantaged locales).



Appendix Figure A2: Estimating H from grouped versus unit-level data, San Francisco 2000

Appendix B: Effects of estimation from weighted data

The formulae in the section on estimates with weighted data suggest the sign of the bias introduced by weighting and provide a method to correct for it. The resulting estimates rely on three main assumptions: 1) there is a sufficient number of tracts so that the estimated metro-level variance or rank-variance has minimal sampling variation, 2) tract samples are big enough so we can ignore third-order or higher terms in the expansion of estimated entropy around its true value, and 3) in the computation of bias corrections the weights are not correlated with household incomes within tracts.

Assumptions (1) and (2) are discussed in Logan et al. (2018) and in Reardon et al. (2018). Assumption (2) is sensitive to the degree of segregation. In highly segregated tracts, for example, the entropy function is not well approximated by a quadratic function and thus the proposed bias correction will not fully eliminate the bias. Assumption (1) will work for larger metros but may be a problem in metros with a small number of tracts.

Assumption (3) allows us to derive formal expressions for bias that depend only on sampled (versus population) data, but is at variance with census sampling practice. To better understand the implication of this simplification we constructed a thought experiment with two populations and derived analytic expressions relating the bias correction used in this paper and the bias that would arise if weights were correlated with income within tract. Under reasonable conditions we found the difference was a small percentage of the bias (a supplementary appendix showing this derivation is available from the authors on request). However, because the derived expressions require information on sampling stratification that is not available to us and rely on population measures that would not be available in sampled data, we could not construct a feasible bias correction that accounted for this correlation.

In this Appendix, we therefore use a simulation procedure to explore how well our estimates work in a realistic data set in which we have control over the sampling process as in Logan et al. (2018). Recall that the problem with weighted data arises when weights are correlated with people's incomes. The full weighting scheme of the Bureau of the Census is confidential, but for our purpose all we need to know is how much weights vary and how weights are correlated with income. We examined these relationships using the confidential family-level income data in the FSRDC for the city of Chicago from the 2008-2012 ACS. In particular, using these ACS sample data, we estimated a multi-level model in which the left-hand side variable is the log of the household weight w_{ij} constructed by the Census. The underlying structure of the equation is

$$\ln(w_{ij}) = \beta_0 + \beta_1 y_{ij} + \beta_2 y_{ij}^2 + \beta_3 \bar{y}_j + \beta_4 \sigma_j + \beta_5 N_j + \mu_j + \varepsilon_{ij}$$

where y_{ij} is the family income of household *i* in tract *j* divided by the mean income of all households in the metro area, $\bar{y}_j$. σ_j are the tract-level mean and variance of y_{ij}, N_j is the number of households in the tract, and μ_j and ε_{ij} are tract and household level random effects. We also carry out a similar set of estimates for the sub-population of African-Americans.

The results appear in Appendix Table B1. As is evident from this table there is substantial predictable as well as unpredictable variation in weights both within and across tracts. The income relationship with weights is quadratic with small but significant coefficients on both the linear and squared terms, though the relationship is negative over the full range of the income measure. The overall correlation between the weights and income is -.037 for all households and -.0091 for black households. Thus, while we do observe that poor households tend to receive higher weights, the relationship is not strong.

Appendix Table B1. Multilevel model predicting household weight (ln), Chicago MSA, ACS 2008-2012, for all households and non-Hispanic black households				
	All Households		Black Households	
	b	SE	b	SE
Level 1 (household)				
Household income	-0.0305	0.0016	-0.0133	0.0037
Household income-squared	0.0013	0.0001	0.0001	0.0002
Level 2 (tract)				
Tract-level mean of household income	-0.1465	0.0163	0.0914	0.0176
Tract-level variance of household income	0.0103	0.0375	-0.0051	0.0021
N of households in the tract (logged)	0.3684	0.0100	-0.0097	0.0076
Constant	-0.1082	0.0723	2.5560	0.0514
Variance of tract effect	0.039	0.001	0.081	0.004
Variance of household effect	0.280	0.001	0.330	0.003
Log likelihood *	-179,000		-31,680	
Intraclass Correlation (ICC)	0.123		0.197	
N of households *	225,000		35,500	
N of tracts *	2,000		1,400	
* Rounding is required by Census Disclosure Review Board				

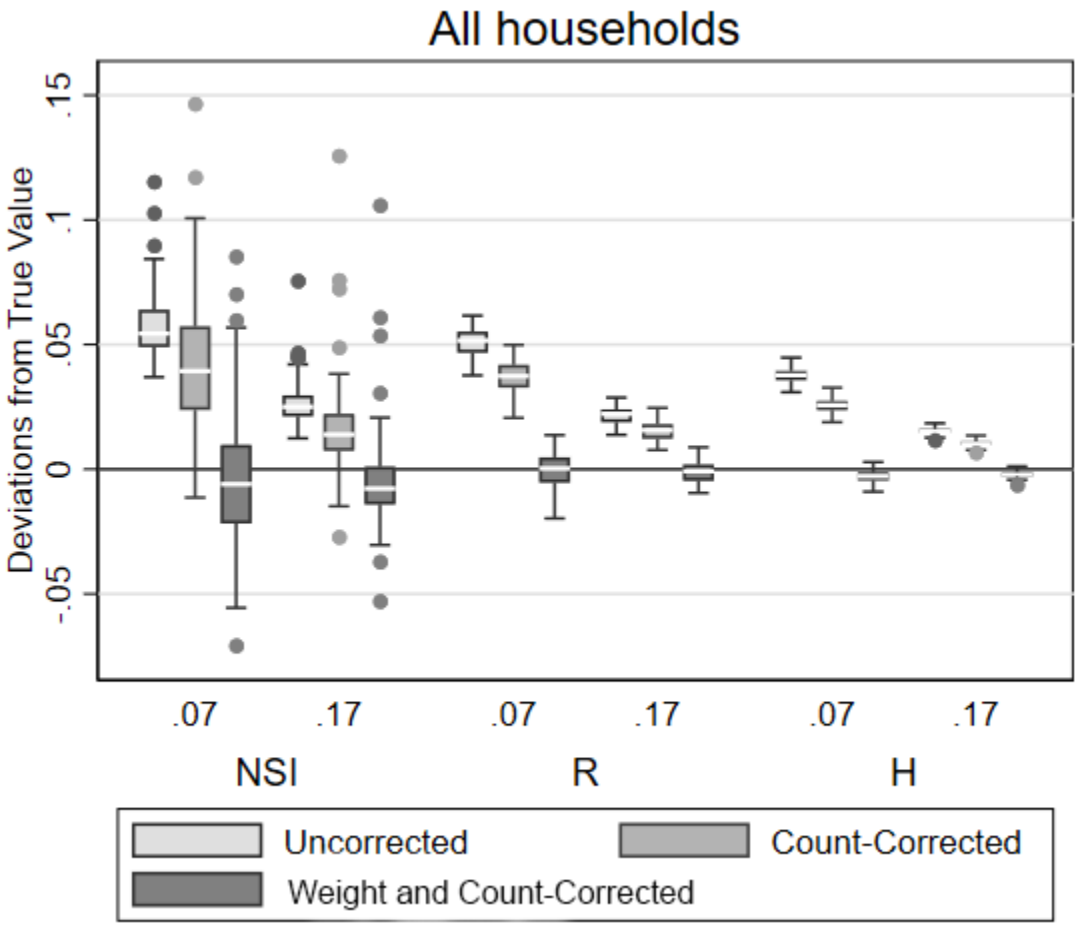
These regression equations are then applied to comparable measures for 1940 households in order to create household weights. We also add in values of μ_j and ε_{ij} drawn from independent normal distributions with variance equal to the estimated variance from the multi-level model. We then invert the simulated weights to obtain relative sampling probabilities and then scale these relative probabilities so that the overall sampling probability corresponds to the approximate sampling probabilities for the long-from and the ACS. Finally, we draw 100 different samples from the 1940 Chicago data for every sampling rate and segregation measure that we study and estimate the corrected and uncorrected measures using the estimated weights.[7]

[7] While the 1940 census data are not top-coded, preliminary analysis indicated that the NSI is very sensitive to extreme values of income when sampling rates are low. In the simulations, we therefore recoded all income values above 99% to the 99% level. We also used the 99th percentile to top-code the RDC data.

Figure B1 presents boxplots of the bias in estimates of H, R and the NSI for sampling rates of seven and seventeen percent, which reflect roughly the difference between the long form Census 2000 and the ACS. We include the uncorrected estimates of the bias, the estimates corrected for sample count, and the estimates where we also corrected for weighting. We report biases rather than the estimates themselves so estimates of different measures may be presented on the same graph. For both sampling rates and for all three measures, the mean uncorrected estimate has a strong upward bias, the bias is reduced by correction for the sample count, and there is almost no bias for the final estimate that is corrected for both counts and weighting. The mean value for estimates corrected only for sample count is about 2/3-3/4 of the total bias, so correcting for sample count alone leaves considerable bias when applied to weighted data with low sampling rates, as in the current ACS. Put another way, if the true level of income segregation between two points in time had not changed but the sample size was reduced from .17 to .07 then the weight-corrected estimates of H would (correctly) yield essentially no change in income segregation, the count corrected estimates would yield an average increase of .004 and the uncorrected estimate would yield an average increase of .009.

Figure B1 also illustrates the variability that is intrinsic to estimating income segregation using sampled data. In any given city at any point of time we would likely get a somewhat different estimate if we drew a new sample. When we find differences across metropolitan areas in the final corrected estimate (i.e., the standard deviations in Table 1) do these represent true variation in income segregation across cities or whether it results simply from sampling variation? The simulated sampling variance in Figure B1 for H with a 7 percent sample is only 0.0012, or less than 5 percent of the reported standard deviation for the 2012-2016 corrected measure of .027 in Table 1. Thus, we infer that most of the variation reported in Table 1 reflects

true differences in income segregation across large metros. The comparable figures for R and NSI are .0040 and .0037, respectively, in Figure B1 versus a standard deviation across metros in Table 1 of .050 and .028, respectively, so again sampling variability plays a secondary role.

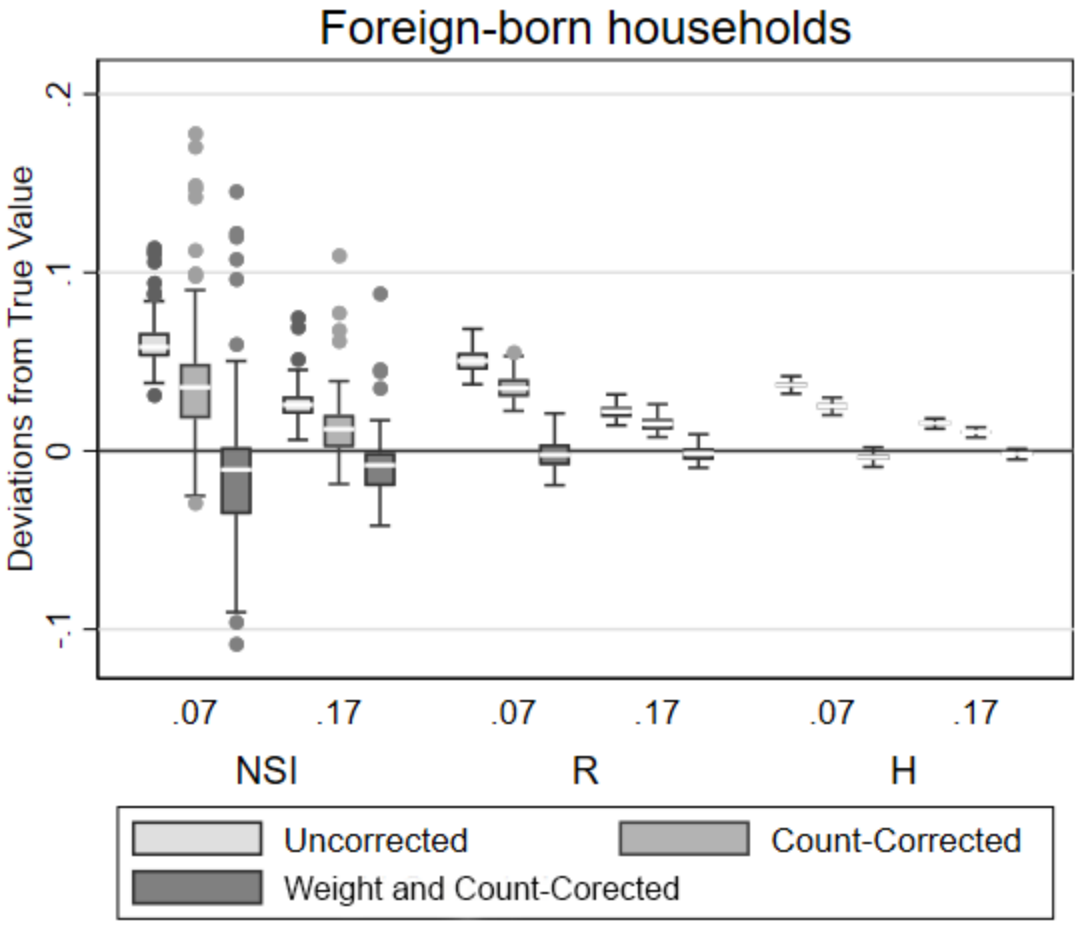


Appendix Figure B1. All households, Chicago 1940. Sampling variation of estimates of income segregation with weighted data, comparing estimates that are uncorrected, estimates corrected only for sample count, and estimates corrected for both sample count and weighting.

As noted in Logan et al. (2018) the problem of smaller samples in the measurement of income segregation is likely to be even more acute in subsamples of the population because what matters is the count of households in the sample for every tract. To examine this phenomenon with 1940 data we select the foreign-born population. We expect results for these households to

be informative for minorities like African Americans in contemporary data, given that both are a similar fraction of the total population and both have lower than average incomes. To reinforce this similarity we applied results from the weights for African Americans in the ACS to create weights for the foreign born population in the 1940 census.

Figure B2 reports the results. Overall, we see that biases in the uncorrected estimates are about three times as high for this subgroup as for the total population shown in Figure B1. The count-corrected estimates are always above the corrected estimates but these latter estimates in five of the six cases are just below the average value. But even for the NSI where the estimates tend to be low, there is essentially no difference in the bias by sampling rate for the weight-corrected estimates, a .013 difference for the count-corrected estimates and a .038 difference for the uncorrected estimate.



Appendix Figure B2. Foreign-born households, Chicago 1940. Sampling variation of estimates of income segregation with weighted data, comparing estimates that are uncorrected, estimates corrected only for sample count, and estimates corrected for both sample count and weighting.