Stanford cepa | Center for Education Policy Analysis

CEPA Working Paper No. 18-02

Has Income Segregation Really Increased? Bias and Bias Correction in Sample-Based Segregation Estimates

AUTHORS

Sean F. Reardon
Stanford University

Kendra Bischoff
Cornell University

Ann Owens
University of Southern California

Joseph B. Townsend
Stanford University

ABSTRACT

Several recent studies have concluded that residential segregation by income in the U.S. has increased in the decades since 1970, including a significant increase after 2000. Income segregation measures, however, are biased upwards when based on sample data. This is a potential concern because the sampling rate of the American Community Survey (ACS)—from which post-2000 income segregation estimates are constructed—was lower than that of the earlier decennial Censuses. This raises the possibility that the apparent increase in income segregation post-2000 simply reflects increased upward bias in the estimates from the ACS, and the estimated increase may therefore be inaccurate.

In this paper, we first derive formulas describing the approximate sampling bias in two measures of segregation. Next, using Monte Carlo simulations, we show that the bias-corrected estimators eliminate virtually all of the bias in segregation estimates in most cases of practical interest, although the correction fails to eliminate bias in some cases when the population is unevenly distributed among geographic units and the average within-unit samples are very small. We then use the bias-corrected estimators to produce unbiased estimates of the trends in income segregation over the last four decades in large U.S. metropolitan areas. Using these corrected estimates, we replicate the central analyses in four prior papers on income segregation. We find that the primary conclusions from these papers remain unchanged, although the true increase in income segregation among families after 2000 was only half as large as that reported in earlier work. Despite this revision, our replications confirm that income segregation has increased sharply among families with children in recent decades, and that income inequality is a strong and consistent predictor of income segregation.

VERSION

May 2018

Suggested citation: Reardon, S.F., Bischoff, K., Owens, A., & Townsend, J.B. (2018). Has Income Segregation Really Increased? Bias and Bias Correction in Sample-Based Segregation Estimates (CEPA Working Paper No.18-02). Retrieved from Stanford Center for Education Policy Analysis: http://cepa.stanford.edu/wp18-02

Has Income Segregation Really Increased?
Bias and Bias Correction in Sample-Based Segregation Estimates

Sean F. Reardon
Stanford University

Kendra Bischoff
Cornell University

Ann Owens
University of Southern California

Joseph B. Townsend
Stanford University

Version: May, 2018

Forthcoming in *Demography*

Direct correspondence to sean f. reardon at sean.reardon@stanford.edu, 520 CERAS Building, #526, Stanford University, Stanford, CA 94305. The research described here was supported by a grant from the UPS Foundation Endowment Fund at Stanford University and by fellowships to Bischoff and Owens from the National Academy of Education/Spencer Postdoctoral Fellowship Program. The opinions expressed here are ours and do not represent the views of UPS, Stanford University, the National Academy of Education, or the Spencer Foundation.

Has Income Segregation Really Increased?

Bias and Bias Correction in Sample-Based Segregation Estimates

Abstract

Several recent studies have concluded that residential segregation by income in the U.S. has increased in the decades since 1970, including a significant increase after 2000. Income segregation measures, however, are biased upwards when based on sample data. This is a potential concern because the sampling rate of the American Community Survey (ACS)—from which post-2000 income segregation estimates are constructed—was lower than that of the earlier decennial Censuses. This raises the possibility that the apparent increase in income segregation post-2000 simply reflects increased upward bias in the estimates from the ACS, and the estimated increase may therefore be inaccurate.

In this paper, we first derive formulas describing the approximate sampling bias in two measures of segregation. Next, using Monte Carlo simulations, we show that the bias-corrected estimators eliminate virtually all of the bias in segregation estimates in most cases of practical interest, although the correction fails to eliminate bias in some cases when the population is unevenly distributed among geographic units and the average within-unit samples are very small. We then use the bias-corrected estimators to produce unbiased estimates of the trends in income segregation over the last four decades in large U.S. metropolitan areas. Using these corrected estimates, we replicate the central analyses in four prior papers on income segregation. We find that the primary conclusions from these papers remain unchanged, although the true increase in income segregation among families after 2000 was only half as large as that reported in earlier work. Despite this revision, our replications confirm that income segregation has increased sharply among families with children in recent decades, and that income inequality is a strong and consistent predictor of income segregation.

Has Income Segregation Really Increased?

Bias and Bias Correction in Sample-Based Segregation Estimates

Introduction

Several recent studies have documented an increase in residential segregation by income in the U.S. over the last four decades (Bischoff and Reardon 2014; Jargowsky 1996; Owens 2016; Reardon and Bischoff 2011; Watson 2009). In particular, income segregation in the U.S. appears to have grown sharply in the 1980s and since 2000, indicating that American society is becoming more spatially polarized. Owens (2016) shows that the increase in income segregation is driven largely by the growing segregation of families with children. Given the importance of neighborhood socioeconomic conditions for young children's development and opportunities for economic mobility (Brooks-Gunn, Duncan and Aber 1997; Chetty and Hendren 2015; Chetty, Hendren and Katz 2015), the increasing economic segregation of children is of particular concern.

There is reason, however, to doubt the reported increase in economic segregation. The estimates of income segregation reported in the aforementioned papers are based on household or family income data reported on the "long form" of the U.S. decennial Census from 1970 to 2000 and on the American Community Survey (ACS) from 2005 onward. Only a sample of the U.S. population was asked to fill out the long form of the decennial Census from 1970 to 2000 (approximately one in six households); and an even smaller sample is asked to fill out the ACS in any 5-year window (approximately one in 12 households). As we describe below, segregation estimates based on random samples are generally biased upwards relative to the values that would be measured if the full population were observed. Moreover, the upward bias is inversely related to the sampling rate. This means that estimates of income segregation based on the Census and the ACS data are biased upwards, and the upward bias is larger for estimates based on the ACS than for estimates based on the Census. As a result, we would expect to see

an increase in estimated segregation between 2000 and later years, a period in which samples rates declined from approximately 17 to 8 percent, even if there were no true change in levels of income segregation. Without a method of accounting for (or eliminating) the bias in segregation estimates, we cannot make valid comparisons between estimates based on different sampling rates or sample sizes.

We are not the first to raise this concern. Most recently, Logan et al (forthcoming) question the reported increase in income segregation after 2000. They rightly note that the difference in sampling rates between the Census long form and the ACS bias estimates of the post-2000 trend in income segregation upwards. Napierala and Denton (2017) note that sampling variation leads to imprecision in segregation estimates and upward bias when samples are small. Several earlier papers also noted upward bias in segregation measures as a result of stochastic processes (Winship 1977) and small sample sizes (Reardon and Bischoff 2011). Some of these papers propose strategies for constructing unbiased sample-based estimates of segregation (Logan et al. forthcoming; Reardon and Bischoff 2011). The Reardon and Bischoff (2011) approach fails to remedy the problem, however (see Logan et al. forthcoming), while the methods proposed by Logan et al. (forthcoming) have not been validated across a wide range of data-generating conditions and require, in some cases, access to restricted micro-data. Our goals in this paper, therefore, are to develop and validate a method of eliminating the bias in sample-based segregation measures that does not rely on access to micro-data, and then to use this method to produce unbiased estimates of segregation patterns and trends over the last few decades.

We first derive formulas describing the approximate sampling bias in two measures of segregation. Our formulas allow us to quantify the bias in both binary measures of segregation between two mutually exclusive groups and in measures of rank-order segregation, such as the measures widely used to study income segregation. These formulas describe the approximate bias in segregation estimates as a function of the average unit (e.g., census tract) population and the harmonic mean of the sampling rate across units. If both unit population sizes and sampling rates are known, we show that they

can be used to construct bias-corrected segregation estimates without relying on access to sample micro-data. Using Monte Carlo simulations, we test these bias-corrected estimators across a wide range of realistic residential patterns and characterize the range of conditions under which they provide approximately unbiased estimates of segregation. We show that the bias-corrected estimators eliminate virtually all of the bias in segregation estimates in most cases of practical interest, although the correction fails to eliminate bias in some cases when the population is unevenly distributed among geographic units and the average within-unit samples are very small.

Second, we use the bias-corrected estimators to produce unbiased estimates of the trends in income segregation over the last four decades in large U.S. metropolitan areas. Using these corrected estimates, we replicate the central analyses in four prior papers on income segregation (Bischoff and Reardon 2014; Owens 2016; Reardon and Bischoff 2016; Reardon and Bischoff 2011) to examine whether the trends and patterns reported in those papers were an artifact of the biased estimators they relied on. We find that the primary results in these papers hold up, though the true increase in income segregation among families after 2000 was only half as large as that reported by Bischoff and Reardon (2014).

Sampling in the Decennial Census and American Community Survey

The Census is a decennial, housing unit-based survey that collects limited information on the full U.S. population, including housing tenure status and an enumeration of the age, sex, and race/ethnicity of each household member. These data are available for all geographic levels down to the census "block," the basic Census sampling unit that encompasses approximately one city block (though blocks may be larger in suburban and rural areas). All other sociodemographic information tabulated by the Census (including, in particular here, the household and family income data used to estimate income segregation) is collected from a sample of Americans from what was formerly called the Census "long form," also collected every ten years. These data are generally publicly available and tabulated in slightly

larger geographic units, such as "block groups" (an aggregation of several contiguous blocks) and tracts (an aggregation of several contiguous block groups). Tracts typically contain several thousand residents and have often been used in sociological research to approximate residential neighborhoods.

In Censuses up to and including 2000, the long-form data were collected from samples of approximately one in six households, or about 17 percent of the U.S. population. This amounted to approximately 18 million households contacted, and 16.4 million usable questionnaires (National Research Council 2007). Using unweighted sample counts and population estimates available from the Census Bureau, we estimate that on average in the 1970 through 2000 Censuses, 250-300 households were sampled in each tract and the household sampling rate ranged from 17 to 20 percent.

The ACS replaced the Census long form after 2000 with the promise of providing more frequent (annual) estimates of sociodemographic characteristics of the U.S. population. The cost of these annual estimates is a reduction in annual sample size. In addition, because segregation estimates rely on data for small geographies, it is necessary to use aggregate ACS data across 5-year windows to accumulate what the Census deems to be sufficient sample sizes in smaller geographies like tracts. In 2005, the first year that the ACS was fully implemented, the Census Bureau aimed to achieve a sampling rate of approximately 12 percent over a five-year period by sampling about three million unique addresses per year.[1] The ACS sample, however, is then reduced substantially by subsampling for in-person interviews.[2] In reality, the ACS sampled approximately 14.5 million housing units in the 2005-09 period, and conducted 9.7 million final interviews to be included in the usable data (approximately 7.5 percent of total housing unit addresses). The Census Bureau increased the sample size after 2011, resulting in an original sample of approximately 16.8 million housing units in the 2010-14 period, and a final sample of

[1] This plan aimed to survey 15 million housing units over 5 years out of the approximately 130 million housing unit addresses in the United States in 2005 (National Research Council 2007).
[2] The Census follow-up rate for non-response and unmailable addresses varies by the tract characteristics (Bureau of the Census n.d.).

nearly 11 million usable questionnaires. However, the Census Bureau samples housing units from small geographies, and target sampling rates vary by tract characteristics (Bureau of the Census 2014). Using available unweighted sample counts and population estimates, we estimate that in the ACS 5-year aggregate data from 2005-09 to 2012-16, the average tract sample size is 130-160 households and the average tract sampling rate ranged from 8 to 10 percent.

Clearly there has been a decline in the sample sizes and rates since the inception of the ACS, especially in small geographies. Knowing that small sample sizes would be problematic for researchers, in particular those interested in neighborhood analyses, the Census Bureau began publishing confidence intervals in ACS data tables to highlight the imprecision in the estimates. The Census Bureau was less transparent about sampling error in census long form data, though these small-geography data also suffered from substantial imprecision (to a lesser extent than data in the ACS) (National Research Council 2007: 65-74).

Prior research on sampling bias and corrections in income segregation estimates

Segregation measures are generally based on a decomposition of the population variation in income (or race or any other characteristic) into between- and within-neighborhood components (Reardon 2011; Reardon and Firebaugh 2002). Commonly-used measures differ in how they quantify variation: Jargowsky's Neighborhood Sorting Index (NSI) uses the variance of income (Jargowsky 1996); the rank-order information theory index (H^R) uses a measure of the entropy of income ranks; and the rank-order variance ratio index (R^R) uses the variance of income ranks (Reardon 2011; Reardon and Bischoff 2011), to name a few. The bias in sample-based segregation measures arises because the observed variation in a finite sample is generally a downwardly-biased estimator of the variation in the full population, and the magnitude of the downward bias is inversely related to sample size. This means that sample-based segregation estimators underestimate the true extent of within-neighborhood

variation much more than they underestimate population variation (because the population variation is estimated from a much larger sample than each neighborhood's). As a result, sample-based segregation estimates assign too much of the variation to the between-unit component of the decomposition: that is to say, they overestimate segregation.[3]

Several recent papers propose methods for correcting the sampling bias in income segregation estimates. Reardon and Bischoff (2011) noted that comparisons of segregation estimates across populations (across metropolitan areas, years, or racial/ethnic groups) are biased if the average within-tract sample sizes differ among populations. To address this, they randomly sampled 10,000 families from the Census-reported population in each metropolitan area-year-race/ethnic group cell, reasoning that by equalizing sample sizes they would equalize the amount of bias in each estimate. Several more recent papers use a modified version of this method, sampling a number of families equal to 50 times the number of census tracts in a metropolitan area, reasoning that it was preferable to hold the average sample size per tract constant rather than the total sample size (Bischoff and Reardon 2014; Owens 2016; Reardon and Bischoff 2016).[4]

Reardon and Bischoff (2011) argued that, while this approach would not yield unbiased estimates in any given population, the expected bias would be equal in each estimate, allowing for unbiased estimation of changes over time and differences between race/ethnic groups. Subsequent analyses of this method, however, indicate that subsampling from populations based on Census or ACS estimates does not yield comparable bias across populations (Logan et al. forthcoming).[5] The segregation trends

[3] To see this in a simple (extreme) case, suppose each neighborhood in a city were 50% poor and 50% rich, and we estimated income segregation by drawing a random sample of one person from each neighborhood. There would be no within-neighborhood variation in our sample, but considerable variation in the population as a whole, so we would (very wrongly) conclude that the city was completely segregated by income.

[4] Logan et al (forthcoming) describe the approach used in these papers; the sampling procedure used is not well-documented in the published papers.

[5] We have replicated this Logan et al (forthcoming) finding in our own analyses (not shown). The Reardon and Bischoff (2011) approach does not eliminate differential bias in segregation estimates. The resampling is done from the estimated tract income distributions, not the actual income distributions. The income variation in these estimated distributions is more downwardly-biased in the ACS (because it is based on smaller samples) than in the

and patterns reported in the Reardon and Bischoff papers and the Owens paper are therefore potentially confounded by differential bias across time, place, and/or race/ethnic group.

Logan et al (forthcoming) propose several approaches for constructing unbiased sample-based estimates of segregation. First, they suggest an approach—which they term sparse-sampling variance decomposition (SSVD)—that uses Census or ACS micro-data and a finite sample correction to obtain an unbiased estimate of the average within-tract variance. This method is useful when one is using a variance-based segregation measure such as Jargowsky's (1996) Neighborhood Sorting Index (NSI) or Reardon's (2011) rank-order variance ratio index (R^R) and one has access to micro-data within a restricted-access Census Research Data Center (RDC). Second, they derive formulas describing the approximate bias in H^R, the rank-order information theory index, and in $H10$ and $H90$, the information theory index measures of segregation of poverty and affluence (Reardon and Bischoff 2011). They propose estimating H^R, $H10$, and $H90$ from micro-data and then subtracting the corresponding bias terms from these estimates. They show that both the SSVD approach and the bias-formula correction method yield approximately—but not perfectly—unbiased segregation estimates in the set of six cities they study using micro-data from the 1940 Census.

These methods rely on micro-data. When only aggregated data are available (as is the case with publicly available census data), Logan et al (forthcoming) suggest a multi-step approach: a) using the grouped data to estimate within-unit income distributions; b) generating repeated samples from these estimated distributions; and then c) applying the micro-data based bias-correction approaches to each of these simulated micro-data samples; and d) averaging the resulting estimates. Using simulations based on individual household Census data from Chicago in 1940, they show that this approach to estimating segregation from grouped data yields estimates that are generally less biased than the uncorrected

decennial Census. This differential bias is then carried into the (equally-sized) samples drawn in the resampling process.

estimates when the sampling rate is low (though the adjusted estimator of the NSI appears generally worse than the unadjusted estimator, at least for Chicago). The adjusted grouped-data estimators for H^R, $H10$, $H90$, and R^R appear to over-correct in some cases and under-correct in others, though are (at least for Chicago) less biased than the unadjusted estimators (Logan et al. forthcoming). The approaches they recommend for correcting estimates of segregation based on grouped data have not been validated across a range of data-generating scenarios, however. The approximations on which they are based may break down when samples are small or when sampling rates and/or sample sizes vary among units.

In this paper, we extend this literature in several ways. First, we derive formulas describing the approximate sampling bias in both binary and rank-order measures of segregation, focusing on the information theory index H and the variance ratio index R because these two measures have the most desirable mathematical properties in an index (James and Taeuber 1985; Reardon 2011; Reardon and Firebaugh 2002). Second, we use these formulas to derive bias-corrected segregation estimators that can be used with grouped data and do not require access to sample (or simulated) micro-data. Third, we use simulations to investigate the performance of the bias-corrected estimators over a wide range of data-generating models; we base these data-generating models on the spatial income distribution patterns found in U.S. metropolitan areas. And fourth, we use the bias-corrected estimators to produce corrected estimates of recent trends and patterns of income segregation in the U.S.

A review of binary and rank-order segregation measures

We focus in this paper on two binary measures of segregation, the information theory index, denoted H, and the variance ratio index, denoted R. These indices satisfy a set of important properties, including organizational equivalence, size invariance, organizational decomposability, and the principles of transfers and exchanges (James and Taeuber 1985; Reardon and Firebaugh 2002). The dissimilarity and

Gini indices do not satisfy all of these principles and so are less broadly useful.[6] In addition, both H and R can be used to construct measures of rank-order segregation, denoted H^R and R^R (Reardon 2011), which can be used to measure segregation along some ordered dimension such as income (Reardon and Bischoff 2011). The formulas and properties of these indices are described in detail elsewhere; we briefly review their formulas below.

To begin, we define some notation. We first are interested in computing (binary) segregation between two groups among a set of J units (e.g., census tracts). Let p denote the group proportion in a given unit. For values of $p \in [0,1]$, define the *Interaction index* (I) and *Entropy* (E):

$$I = p(1-p)$$
$$E = -[p \ln p + (1-p)\ln(1-p)], \tag{1}$$

where we define $0 \cdot \ln 0 = 0$. Note that both I and E are concave down functions of p, a feature that leads estimates of both to be biased when p is estimated from a sample (see Appendix section A1). The binary variance ratio and information theory segregation indices are respectively defined as

$$R = 1 - \frac{1}{I}\sum_{j=1}^{J}\frac{t_j}{T}I_j \tag{2}$$

and

$$H = 1 - \frac{1}{E}\sum_{j=1}^{J}\frac{t_j}{T}E_j, \tag{3}$$

where I and E are the values of I and E in the whole population; where I_j and E_j are the values of I and

[6] Moreover, while it is straightforward to show that sample-based estimates of other measures of segregation, such as the dissimilarity index and the Gini index, will also be biased upwards (see Appendix section A9), we do not have a tractable expression for the magnitude of the bias for these indices, because of the presence of the absolute value function in their formulas. For these reasons, we focus on the H and R indices in this paper. See Napierala and Denton (2017) for some discussion of sampling bias in the dissimilarity index.

E in unit j; and where t_j/T is the share of the population in unit j.

If y is an ordered variable such as income, the corresponding rank-order income segregation indices are

$$H^R = \frac{1}{\int_0^1 E(q)dq} \int_0^1 E(q)H(q)dq \; = 2\int_0^1 E(q)H(q)dq$$

$$R^R = \frac{1}{\int_0^1 I(q)dq} \int_0^1 I(q)R(q)dq \; = 6\int_0^1 I(q)R(q)dq \,, \tag{4}$$

where $I(q)$, $R(q)$, $E(q)$, and $H(q)$ are the values of I, R, E, and H when the population is divided into groups defined by whether y is above or below the $100 \times q^{th}$ percentile of y. For example $H(.5)$ is the value of H computed between those with above and below median values of y. The rank-order measures are weighted integrals of the binary indices over values of $q \in (0,1)$. In practice, when y is available only in coarsened form (such as when income data are coarsened into 16 income categories in U.S. census or ACS data), we estimate $H(q)$ or $R(q)$ by first computing H or R at the set of finite values of q that correspond to the percentiles of the thresholds used to coarsen the data; we then fit a polynomial function through the resulting points; and then using the fitted polynomial as an estimate of $H(q)$ or $R(q)$ in Equation (4) above (see Reardon 2011; Reardon and Bischoff 2011).

Bias in sample-based segregation estimates

The formulas above assume we observe p_j and t_j without error in each unit j. Instead, here we assume we know t with certainty but must estimate p from a sample. As we show in Appendix A, the assumption that t is known with certainty is not essential. More specifically, from each unit $j \in \{1, \ldots, J\}$ we observe a simple random sample of size n_j, drawn without replacement from the population in the unit, which is of known finite size t_j. Because p_j is estimated from a sample, $\hat{I}_j$ and $\hat{E}_j$ will be biased

downward (because, as we note above, I and E are concave down functions of p on the interval $(0,1)$). The formulas for R and H in Equations (2) and (3) indicate that downward bias in $\hat{I}_j$ and $\hat{E}_j$ will lead $\hat{R}$ and $\hat{H}$ to be biased upwards. This is the source of upward bias in sample-based estimates of segregation.[7]

In Appendix A, we show that the approximate biases in both $\hat{R}$ and $\hat{H}$ are functions of a bias term, B (defined below), which depends on the arithmetic and harmonic means of the unit populations and the harmonic mean of the sampling rates. The specific formulas are below.

Let t_j denote the population in unit j, and let $\bar{t}-1$ and $\widetilde{t-1}$ denote the arithmetic and harmonic means of t_j-1, respectively. Finally, let C denote the covariance of t_j and I_j:

$$C=\frac{1}{J}\sum_j\left(t_j-\bar{t}\right)\left(I_j-\bar{I}\right). \tag{5}$$

We show in Appendix A that the sampling biases in $\hat{R}$ and $\hat{H}$ are then approximately

$$bias\left(\hat{R}\right)\approx B(1-R)-B\left(\frac{C}{(\bar{t}-1)I}\right) \tag{6}$$

and

$$bias\left(\hat{H}\right)\approx\frac{B}{2E}. \tag{7}$$

In both cases, the bias is proportional to B, defined as

$$B=\frac{z}{\bar{t}-1}\cdot\frac{1-\tilde{r}}{\tilde{r}}, \tag{8}$$

where $\tilde{r}$ is the harmonic mean of the unit sampling rates, and where z is a function of the ratio of the arithmetic and harmonic means of t_j-1:

[7] Strictly speaking, I and E must also be estimated in Equations (2) and (3), and these estimates will be biased downward. But because I and E are estimated from the pooled sample over all units, rather than separately within each unit, the sampling bias in $\hat{I}$ and $\hat{E}$ is small compared to the bias in the $\hat{I}_j$'s and $\hat{E}_j$'s. As a result, $\hat{R}$ and $\hat{H}$ will be biased upwards in general.

$$z = 1 + \frac{1}{\bar{t}}\left(\frac{\bar{t}-1}{\widetilde{t-1}} - 1\right). \tag{9}$$

Note that z has a minimum of 1, obtained if $t_j = \bar{t}$ is constant across units, and grows larger with more variation in the t_j's. Moreover $z \approx 1$ unless $\bar{t}$ is small and the t_j's are highly variable.

Equation (8) shows that the bias term grows large when $\bar{t}$ is small, when $\tilde{r}$ is small, and when z is large. To get a sense of the absolute magnitude of B, note that if $r_j = r$ is constant, then

$$B = \frac{z(1-r)}{\bar{n}-r} \approx \frac{1-r}{\bar{n}}, \tag{10}$$

where $\bar{n} = \frac{1}{j}\sum n_j$ is the average sample size across units. So in the case where census tracts have an average of 1,000 families, and the sampling rate is 0.08 (a typical case with ACS data), $B \approx \frac{.92}{80} = 0.0115$. Note also how the bias factor changes as the sampling rate changes: changing the sampling rate by a factor of c changes B by a factor of $\frac{1-c\tilde{r}}{c-c\tilde{r}}$. So, for example, halving the sampling rate from 0.16 to 0.08 increases B by a factor of 2.19. Halving it again from 0.08 to 0.04 increases B by a factor of 2.09.

The bias in $\hat{R}$ described in Equation (6) has two components, each proportional to B. The first component is positive and proportional to $1 - R$, so that $\hat{R}$ is biased toward 1 by a proportion B. The second component of the bias in Equation (6) may be positive or negative, depending on the sign of C, the covariance between unit populations (t_j's) and unit diversities (I_j's); it is proportional to C and inversely proportional to $\bar{t} - 1$ and I. This bias term will be small relative to the first bias term unless $\bar{t}$ is small and I is small (which occurs if the proportion of individuals in one of the two binary categories is near 0 or 1) and C is large. Note that if $C = 0$ (which will occur by definition if either $t_j = \bar{t}$ or $I_j = I$ is constant across units), then

$$bias\left(\hat{R}\right) \approx B(1-R). \tag{11}$$

In practice, we will assume that the second component of the bias in $\hat{R}$ is 0, since in most cases it

will be very small relative to the first term. As we show below, this assumption is reasonable in many, but not all, cases we examine.

The bias in $\widehat{H}$ described in Equation (7) has a single component. The bias is positive, and proportional to B and inversely proportional to E. This bias term will be large when E is small (which occurs if the proportion of individuals in one of the two binary categories is near 0 or 1). In Appendix A, we note that the approximation in Equation (7) fails substantially when $\bar{t}$ is small and/or E is near 0.

The bias in the rank-order measures $\hat{R}^R$ and $\widehat{H}^R$ follow directly from the bias in the binary measures from which they are computed. We show in Appendix A that the biases are approximately

$$bias\left(\hat{R}^R\right) \approx B(1 - R^R) - \frac{6B}{(\bar{t} - 1)} \int_0^1 C(q)dq$$

$$\approx B(1 - R^R) \tag{12}$$

and

$$bias\left(\widehat{H}^R\right) \approx B. \tag{13}$$

Note that Equations (12) and (13) provide straightforward expressions of the bias in the rank-order segregation measures. These expressions depend on the arithmetic and harmonic means of the unit sizes (which are generally readily observable) and the harmonic mean of the sampling rate (also available from published Census data). In the case of the rank-order variance ratio index, the bias also depends on the true level of segregation. Note also that Equation (13) is a more general and more precise version of the bias formula derived in Logan et al. (forthcoming).[8]

The bias described in Equations (12) and (13) above results from the fact that income data are based on samples. One might additionally worry that the coarsening of income data may lead to

[8] Logan et al (forthcoming) use a formula that assumes both sampling with replacement (which is not the case with Census data) and a low sampling rate. Our formula in Equation (13) assumes sampling without replacement and accommodates heterogeneity in sampling rates (and unit sizes).

additional bias in rank-order segregation estimates (because even among sampled households, income is not known exactly). Because $E(q)$ and $I(q)$ in Equation (4) are known by definition, there will be error in $\widehat{H}^R$ or $\widehat{R}^R$ only to the extent that the estimated functions $\widehat{H}(q)$ and $\widehat{R}(q)$ are error-prone. Intuitively, the error in the functions $\widehat{H}(q)$ and $\widehat{R}(q)$ will depend on the number and location of the data points used to estimate them; these are determined by the number and location of the thresholds used to coarsen the income data. When there are fewer income categories, and when these income categories are not relatively evenly spaced across the income distribution, the uncertainty in $\widehat{H}(q)$ and $\widehat{R}(q)$ will be greater, leading to more uncertainty in the rank-order income segregation estimates. Although this coarsening may lead to imprecision in income segregation estimates, there is no reason to expect it to lead to systematic bias in rank-order segregation measures or error patterns that differ systematically over time or are related to sampling rates. Having fewer or differently located income thresholds would not be expected to systematically shift the fitted functions $\widehat{H}(q)$ and $\widehat{R}(q)$ upwards or downwards. Moreover, Reardon (2011) shows that choosing different numbers or locations of the income thresholds does not yield systematic differences in estimated segregation levels. Because there is no theoretical reason to expect systematic bias related to the coarsening of income data, we focus here on the bias that results from sampling.

Bias-corrected segregation measures

Given the bias approximation formulas above, we can construct bias-corrected estimates of segregation. Let $\widehat{R}$ and $\widehat{H}$ denote the sample-based estimates of segregation. We show in Appendix A that we can construct bias-corrected estimates $\widehat{R}^*$ and $\widehat{H}^*$ as follows:

$$\widehat{R}^* = \frac{\widehat{R} - B}{1 - B}$$

$$\widehat{H}^* = \widehat{H} - \frac{B}{2E}.$$

(14)

From these we can estimate polynomial functions $\hat{R}^*(q)$ and $\hat{H}^*(q)$ (following Reardon 2011; Reardon and Bischoff 2011, we use fourth-order polynomials; Reardon 2011 shows that income segregation estimates are insensitive to the choice of higher-order polynomials) and use the fitted values to construct bias-corrected measures of rank-order segregation using the formulas in Equation (4):[9]

$$\hat{R}^{R*} = 6\int_0^1 I(q)\hat{R}^*(q)dq$$

$$\hat{H}^{R*} = 2\int_0^1 E(q)\hat{H}^*(q)dq.$$

(15)

Equation (15) provides a much more straightforward bias-correction to the rank-order indices than proposed by Logan et al (forthcoming): it requires no access to micro-data, no estimation of unit-specific income distributions from grouped data, and no simulated sampling from the estimated distribution. Likewise, because the bias-corrected version of the rank-order information theory index here uses a bias formula that applies under a wider range of conditions than that used by Logan et al (forthcoming) (see footnote 8 above), Equation (15) may also provide more accurate bias correction over a range of data generating scenarios than the Logan et al (forthcoming) approach.

[9] The corrections in Equation (15) rely on first correcting the binary measures and then constructing a rank-order measure from these estimates. An alternate approach would be to use the uncorrected binary segregation measures to construct a (biased) rank-order segregation estimate, and then to correct the rank-order measure, using the following formulas:

$$\hat{R}^{R*} = \frac{\hat{R}^R - B}{1 - B}$$
$$\hat{H}^{R*} = \hat{H}^R - B.$$

These formulas will yield identical estimates of $\hat{R}^{R*}$ as Equation (15), but will typically yield very slightly different estimates of $\hat{H}^{R*}$. We prefer the approach described by Equation (15) both because it yields rank-order estimates that are consistent with the binary estimates used to construct them, and because in simulations we conducted (not shown), Equation (15) generally produced very slightly better results (in terms of bias elimination) than this alternate approach.

Assessing the accuracy of the bias-corrected segregation measures

Equations (14) and (15) will yield unbiased estimates of segregation if the approximations used in deriving Equations (6) and (7) are accurate. To assess the validity of these approximations and the resulting formulas, we conduct a series of simulation analyses. In order to ensure that our simulations capture the range of data-generating conditions that arise in practical applications, we use observed tabulations from the 2005-2009 ACS in generating simulated data. Specifically, we do the following.[10]

First, we select all census tracts in a given metropolitan area (we use the 2003 Office of Management and Budget metropolitan area and division definitions). For each census tract, the ACS provides a tabulation of the estimated family income distribution, with income reported in 16 discrete ordered categories. This tabulation takes the form of a vector $\{\hat{t}_{1j}, \hat{t}_{2j}, \ldots, \hat{t}_{16j}\}$, where $\hat{t}_{kj}$ is the estimated number of families in income category k in tract j and where $t_j = \sum_k \hat{t}_{kj}$ is the total number of families in tract j. We also obtain the reported unweighted sample size n_j in tract j, and compute the tract-specific sampling rate $r_j = n_j/t_j$.[11]

Second, we construct a simulated population data file with $T = \sum_j t_j$ observations, where each observation represents a single family in tract j with income in category k, and where the grouped income distribution in each tract is defined by the ACS-reported tabulations. We treat this population as the `true' population of the metropolitan area; from it we compute the `true' binary and rank-order

[10] The data and code used in these simulations are available at https://cepa.stanford.edu/wp18-02.

[11] Tract-level unweighted sample sizes of persons and housing units are publicly available from the Census Bureau for each decennial census and ACS 5-year aggregate estimate. We downloaded them via Social Explorer. We estimate tract-level sampling rates as the ratio of the unweighted count of housing units to the population estimate of housing units reported by the Census Bureau. The population estimates are subject to margins of error, so the sampling rates we compute are subject to a small amount of error. This will tend to lead to very slight underestimates of the harmonic mean of sampling rates ($\tilde{r}$), which may in turn lead to very slight overcorrections of sampling bias in segregation estimates. However, any such overcorrection will generally be extremely small. Note that If tract-level sampling rates were not available, the overall sampling rate in the larger geographic unit of interest, here the metropolitan area, could be substituted; as long as the sampling rates do not vary substantially among tracts, the overall sampling rate will be a reasonable approximation of the harmonic mean of tract sampling rates.

segregation indices in this population.

Third, for a given sampling rate r, we draw without replacement samples of sizes $r \cdot t_j$ from each tract j. From this sample, we estimate the unadjusted binary and rank-order segregation measures and their bias-corrected analogs. We repeat this step 100 times, producing distributions of uncorrected and corrected segregation estimates.

Finally, we repeat this process for each of the 380 metropolitan areas (excluding Puerto Rico) and for sampling rates ranging from $r = 0.02$ to $r = 0.20$. For each metropolitan area and sampling rate (and for both H and R), we compute the average of the uncorrected estimates; the difference between this average estimate and the 'true' segregation level is the bias of the uncorrected estimate. We do the same with the bias-corrected estimates in order to assess the bias in the corrected estimator. For each metropolitan area and sampling rate, we compute the bias in these estimators of binary segregation at each of 15 income thresholds, and for both the corrected and uncorrected versions of each of the two rank-order indices.

Simulation results

We first examine the uncorrected and corrected binary segregation measures $\hat{R}$ and $\hat{H}$ as a function of the proportion of families in the metropolitan area below the income threshold used to define the binary measure. Figure 1 presents the bias in the uncorrected and corrected binary segregation measures at each income threshold in each metropolitan area when the sampling rate is $r = 0.08$. This results in $380 \cdot 15 = 5{,}700$ estimates of bias, spanning a wide range of income thresholds and levels of segregation.

[Figure 1 here]

Figure 1 shows that, at an 8 percent sampling rate, the bias in the binary $\hat{H}$ is roughly +0.007 when the income threshold is in the middle of the income distribution, but as much as three times larger

when it is in the tails of the income distribution. The bias in the binary $\hat{R}$ is roughly +0.01 at income thresholds across the full income distribution. The bias correction performs well in most cases for both measures of segregation, though it clearly overcorrects binary $\widehat{H}$ at low and high income percentiles.

For the binary $\hat{R}$ measure, there is a slight but discernible negative slope in both the uncorrected and corrected bias. In supplemental analyses, we find that this is driven by the fact that C (equation 5) is generally not zero (see Appendix B). Rather, within most metropolitan areas, tract size is often slightly positively correlated with median income. As a result, C is generally positive for high income thresholds and negative for low income thresholds. Because we assume $C = 0$, the bias-corrected estimator tends to slightly overcorrect the binary $\hat{R}$ at high income percentiles and undercorrect at low percentiles. In Appendix B, figures B1 and B2 show that when we artificially constrain all tracts in a metropolitan area to have the same population size in our simulations (thereby setting $C = 0$ by construction), the bias-corrected estimator performs equally well at all income thresholds. Nonetheless, the bias correction is generally very good; remaining bias in the corrected measures is generally a very small fraction of original bias of the uncorrected measures.

Figure 2 presents smoothed estimates of the bias in binary uncorrected and corrected $\hat{R}$ and $\widehat{H}$; the dashed lines for the uncorrected and corrected estimates at $r = 0.08$ are constructed by fitting a smoothed line through the points in Figure 1. The lines associated with $r = 0.04$ (solid) and $r = 0.16$ (dotted) are estimated in a similar fashion. For the uncorrected measures, the average bias at any income percentile decreases as the sampling rate increases. The bias corrections, however, appear to perform equally well across a range of sampling rates, except for the bias-corrected $\widehat{H}^*$ measure at income percentiles near the extremes of the income distribution.

[Figure 2 here]

We next examine the bias in the rank-order income segregation measures. Figure 3 describes the bias in the uncorrected and corrected rank-order $\widehat{H}^R$ and $\hat{R}^R$ as a function of sampling rate, where r

varies from two percent up to 20 percent. Bias in the uncorrected measures is clearly decreasing as sampling rate increases, and is roughly twice as large at eight percent as at 16 percent for both $\hat{H}^R$ and $\hat{R}^R$, suggesting that comparisons of Census- and ACS-based rank-order income segregation estimates will be biased. The bars on the figure represent the range in which the bias of 95 percent of metropolitan areas fall. For the uncorrected measures of rank-order income segregation, the bias varies considerably among metropolitan areas, particularly as the sampling rate gets small. This is partly because the bias term B depends on the average tract population, which varies across metropolitan areas. In the case of $\hat{R}^R$, the variation also emerges because the bias depends on the true value of R^R (see Equation 12), which varies across metropolitan areas.

[Figure 3 here]

The bias in the corrected measures, averaged across all metropolitan areas, is nearly zero at all sampling rates. Moreover, the bias is not only zero on average, but it is nearly zero in every metropolitan area, as seen by the very narrow 95 percent range bars (some of which are too narrow to be visible in the figure). At the very lowest sampling rate (two percent), the bias in the bias-corrected rank-order $\hat{R}^R$ is slightly negative, but still vastly smaller than the bias in the uncorrected estimator. The figure makes clear that the corrected estimators perform very well when used to estimate income segregation among the full population across the range of conditions present in U.S. metropolitan areas.

The derivations of the bias-correction formulas rely on several approximations that are valid when tract populations are relatively large and do not vary much across tracts. These conditions are met when considering the income segregation of all families, but may be less true for sub-populations. Consider income segregation among black families, for example. In many metropolitan areas, the average number of black families per tract is much smaller (one-tenth or even smaller) than the total number of families. Moreover, because of residential segregation, the number of black families per tract varies widely across tracts. Both of these conditions might lead to a failure of the simplifying assumptions used

to derive the bias and bias-correction formulas.

To assess the performance of the bias-corrected measures under such conditions, we compute uncorrected and bias-corrected measures of income segregation among black families from the simulations above.[12] Figure 4 presents the bias in rank-order income segregation among black families by the average number of black families per tract in a metropolitan area. This figure shows two things. First, the bias in estimates of income segregation among black families is very large, particularly in metropolitan areas with few black families. Second, the bias-corrected estimators tend to overcorrect the estimates, producing segregation estimates that are often much too low; this overcorrection is particularly pronounced when the average number of families in a tract is low. For metropolitan areas where the number of black families in the average tract is greater than 200, the bias-corrected estimates are generally less biased (in terms of absolute value of the bias) than the uncorrected estimates, though they are still somewhat negatively biased. Appendix Table B1 quantifies the average extent of overcorrection in $\widehat{H}^{R*}$ and $\hat{R}^{R*}$ at different sampling rates and average tract sizes. At an average tract size of 200, $\hat{R}^{R*}$ overcorrects by roughly 30 percent on average, regardless of sampling rate; $\widehat{H}^{R*}$ overcorrects by even more, particularly at low sampling rates. Below average tract sizes of 200, the bias-corrected estimators do not perform well. This is particularly true for $\widehat{H}^{R}$, where the cure is worse than the disease when average tract sizes are small.

[Figure 4 here]

It is important to note that the failure of the bias-correction formulas in Figure 4 results from the confluence of three conditions: the tract population sizes are a) very small on average, b) highly variable, and c) correlated with tract median income. Without all three of these conditions present, the bias-

[12] We do the same for white and Hispanic families, but we focus here on the simulation results for black families, because they represent the strictest test of the formulas. Failures of the bias-correction formulas are most likely to appear in the black family income segregation estimates, as the black population is both smaller and more segregated than the Hispanic or white population in most metropolitan areas.

correction formulas perform well across a wide range of conditions.

Bias-corrected estimates of trends in income segregation

In this section, we report and compare bias-corrected and uncorrected estimates of income segregation among different populations over the past several decades. We then compare bias-corrected estimates of income segregation to previously published estimates and replicate multivariate regression analyses from several key papers, using the corrected estimates of income segregation. Estimates rely on publicly available counts of families or households in 8 to 25 income categories, depending on the year and population, from the decennial Census or ACS.

We begin by estimating rank-order income segregation among four census-defined populations that have been the focus of recent published research—families, households, families with children under the age of 18, and households without children (Bischoff and Reardon 2014; Owens 2016; Reardon and Bischoff 2011, 2016).[13] Comparisons among estimates for these populations highlight important distinctions among the social contexts and patterns of inequality for children and adults. We present trends in average levels of income segregation within a set of large metropolitan areas used in previous work, those with populations over 500,000 as of 2007 (N=116).[14] We pay particular attention to trends between the 2000 decennial Census and the ACS years because of the change in the sampling rate between these surveys that resulted in reduced tract-level sample sizes.

[Table 1 here]

Table 1 presents uncorrected and bias-corrected estimates of both rank-order H^R and R^R using decennial Census data since 1970 and ACS data for two non-overlapping 5-year spans, 2005-09 (labeled

[13] Households are categorized by the Census as either family or non-family households. Family households consist of people related by marriage or parenthood; non-family households include single people living alone and non-related people living together. Data on income by the presence of children is not publicly available prior to 1990.
[14] Cape Coral is excluded from these estimates due to missing data in 1970.

as 2007, the middle year) and 2012-16 (labeled 2014). Results for families (top panel) show that uncorrected estimates are inflated in all years, with larger differences between uncorrected and bias-corrected estimates in the post-2000 ACS years. From 1970 to 2000, estimates of the average uncorrected H^R and R^R are 4 to 6 percent higher than the corresponding bias-corrected estimates; in 2007 and 2014, the uncorrected estimates are 8 to 10 percent higher than the bias-corrected estimates. Figure 5 presents trends in uncorrected (dashed line) and bias-corrected (solid line) estimates of H^R from 1970 to 2014, clearly showing that the uncorrected estimates are more upwardly biased after 2000, when the tract-level sampling rate declined, than in earlier years. Nonetheless, the corrected estimates show that average metropolitan area income segregation increased by roughly 4% between 2000 and 2014 ($p<0.001$; see Table 1).

[Figure 5 here]

The lower three panels of Table 1 present the average bias-corrected rank-order income segregation estimates among all households, families with children, and households without children. Similar to the findings for families, bias-corrected estimates for these populations are lower than uncorrected estimates in all years, with greater differences after 2000. Bias is smaller for all households (2 to 4 percent in Census years and 6 to 7 percent in ACS years) than families with children or households without children (5 to 8 percent in Census years and 11 to 15 percent in ACS years) because the total household population is larger than the subpopulations by the presence of children.

[Figure 6 here]

Figure 6 provides a summary of trends in rank-order income segregation among all four populations, presenting the mean bias-corrected estimates of H^R found in Table 1. Two key patterns are evident here. First, income segregation is substantially higher among families with children than among households without children. It is slightly higher among family households than among all households after 1980, because a larger share of families have children than do households and segregation is highest

among families with children. Second, there has been no substantial change in income segregation among all households since 1990, though this masks divergent trends among families with children and childless households. Income segregation has risen by 20 percent since 1990 among families with children, with most of the increase occurring between 2000 and 2014 ($p<0.001$). Among households without children, income segregation declined by 10 percent in the 1990s and has been stable since then. The higher segregation levels and divergent trends mean that income segregation among families with children is now more than twice as high as among households without children.[15] Table 1 shows that these patterns are the same if segregation is measured using the rank-order variation ratio index, R^R.

In addition to describing average levels of income segregation, binary H and R can be estimated at any percentile in the income distribution by fitting a polynomial through the binary estimates at each income threshold defined by census or ACS categories and interpolating to every percentile in the income distribution (for details about this method, see Reardon and Bischoff 2011). Figure 7 and Appendix Table C1 present trends in family income segregation at the 90th (*H*90 and *R*90), 50th (*H*50 and *R*50), and 10th (*H*10 and *R*90) percentiles of the income distribution from 1970 to 2014. Figure 7 presents bias-corrected estimates of *H10* (dotted line), *H50* (solid line), and *H90* (dashed line). Uncorrected and bias-corrected estimates for *H* and *R* are presented in Appendix Table C1. The difference between uncorrected and bias-corrected estimates are again larger after 2000.

[Figure 7 here]

Figure 7 shows that segregation of affluent families (*H90*) declined in the 1970s, rose in the 1980s, and declined modestly in the 1990s. Since 2000, *H90* has increased modestly and then declined again. The difference from 2000 to 2014 is not statistically significant. Segregation between families in the top and bottom halves of the income distribution (*H50*) has increased since 1980, with levels significantly

[15] As described above, our bias-correction method performs best when the average tract population in a metropolitan area is at least 200. The average tract population of both families with children and households without children is greater than 200 in all metropolitan areas in our analysis sample.

higher in 2014 than in 2000. Segregation of poor families from others (*H10*) rose in the 1970s and 1980s, declined in the 1990s and has been stable since 2000; the difference between the 2014 and 2000 estimates is not statistically significant. Overall, income segregation has increased only very modestly since 2000 at the top and bottom of the income distribution, with a slightly larger increase between the top and bottom halves of the income distribution. Notably, segregation of affluent families is much higher than segregation of poor families in all years: in 2014, segregation of affluent families was approximately 30 percent higher than segregation of poor families.

Examining trends in income segregation among different racial/ethnic groups is important for understanding the changing relationship between race/ethnicity, socioeconomic status, and residential attainment. Estimates for specific racial/ethnic groups may be particularly prone to bias because the size of these populations is small relative to the total population and, due to racial segregation, unevenly distributed across tracts in many areas of the U.S. As previously noted, our bias-correction method performs best when the average tract population in a metropolitan area is at least 200 families of a particular group (though even then, our method tends to modestly over-correct the bias). Therefore, we estimate rank-order income segregation among white, black, and Hispanic families in only the metropolitan areas in our sample that meet this criterion (116 metropolitan areas for white families, 22 for black families, and 20 for Hispanic families).[16] We discuss R^R for the race-specific results because R^R provides better bias adjustments than H^R for small populations, as shown in Figure 4 above, though trends in H^R and R^R (presented in Appendix Table C2) are consistent with one another.

The estimated trends in average within-race segregation are shown in Appendix Table C2 and Appendix Figure C1. In 2000, income segregation levels were similar among white and black families

[16] For comparison with prior research (Bischoff and Reardon 2014), we focus on estimates that include all white and black families without regard to Hispanic ethnicity. Trends for non-Hispanic white families, also presented in Appendix C, are similar to those for white families, though levels of segregation are 6 to 14% lower among non-Hispanic white families. Income data on non-Hispanic black families are not publicly available.

(0.107 and 0.105, respectively) and slightly lower among Hispanic families (0.095). Income segregation among all three groups has risen since 2000, but segregation rose much more among black and Hispanic families (by approximately 25 percent) than it did among white families (approximately 4 percent). Compared to the 2000 distribution of segregation levels (which had standard deviations of roughly 0.025 for white and black income segregation, and 0.015 for Hispanic segregation; see Appendix Table C2), the increases in black and Hispanic average income segregation were very large—about one standard deviation—while the increase in white income segregation was less than one quarter of a standard deviation. These results are consistent with Logan et al (forthcoming), whose bias-adjusted estimates show that income segregation among black families grew substantially, and faster than income segregation among the population as a whole. Given that the bias-correction will tend to overcorrect the black and Hispanic segregation estimates (and will do so more when the sampling rate is lower), the estimated 25% increase in segregation among black and Hispanic families likely underestimates the true trend modestly. The trends for different racial/ethnic groups are not strictly comparable, however, because they reflect a different set of metropolitan areas. In Appendix Table C2, we report estimated average income segregation among white families in the sets of 22 and 20 metropolitan areas for which we estimate black and Hispanic family income segregation, respectively. These analyses yield the same conclusion: average income segregation among black and Hispanic families increased more than among white families from 2000 to 2014.

Replications of previously published research on income segregation

Do our bias-corrected estimates alter the conclusions reached in previously published research on the trends and correlates of income segregation? Table 2 presents published estimates of rank-order income segregation (H^R) among families (Bischoff and Reardon 2014; Reardon and Bischoff 2016), households, families with children, and households without children (Owens 2016). Recall that the

authors of these papers adjusted the simple income segregation estimates in an attempt to address small sample bias concerns. Their adjustment methods generally inflated the estimates but did not eliminate the bias. Estimates of R^R have not been previously published.

[Table 2 here]

Bischoff and Reardon (2014) concluded that income segregation among families declined modestly in the 1970s, rose sharply in the 1980s, was stable in the 1990s, and increased after 2000. As presented in the top panel, bias-corrected estimates of H^R for the same years and metropolitan area sample support these conclusions. Of particular interest are the trends from 2000 onward, when sampling rates declined. The bias-corrected estimates indicate that rank-order income segregation increased by about four percent from 2000 to 2012, compared to an eight percent increase reported by Reardon and Bischoff (2016). Therefore, about half of the previously-reported increase after 2000 was due to bias induced by the changing sampling rate between the Census and ACS. The bias-corrected estimates indicate that from 1970 to 2012, income segregation among families increased by about 25 percent, slightly less than the previously reported estimates.

Bischoff and Reardon (2014) also report estimates of income segregation by race. We do not compare our results to these previously published estimates because our sample restriction (to metropolitan areas where there are an average of 200 or more black or Hispanic families per tract) differs from that used in earlier work and because previous work did not publish estimates of R^R, which we prefer for the race-specific results. Our bias-corrected estimates in Appendix Table C2 indicate that income segregation among white families in 2000 was more similar to that of black and Hispanic families than Bischoff and Reardon (2014) reported. Our findings are, however, consistent with Bischoff and Reardon's conclusion that income segregation increased more quickly among black and Hispanic families than among white families.

The lower panels of Table 2 compare published and bias-corrected estimates of rank-order

income segregation among all households and by the presence of children. Owens (2016) showed that income segregation did not increase substantially among all households but did increase among families with children from 1990 to 2010. The bias-corrected estimates for the same years and metropolitan areas support this conclusion—H^R changed negligibly from 1990 to 2010 among all households and actually declined by about nine percent among households without children. The previously-reported increase in income segregation, H^R, among families with children from 2000 to 2010 was 17 percent; the bias-corrected increase is about 14 percent. Therefore, our bias-corrected income segregation estimates generally confirm the trends reported in past research, though the magnitude of changes since 2000 is slightly smaller than previously reported.

Past research has also documented a relationship between income segregation and income inequality, demonstrating that rising income inequality is associated with increasing residential sorting by income (Reardon and Bischoff 2011; Watson 2009). We replicate multivariate results from several published papers using bias-corrected H^R and $R^R$and find no substantial change in the results.

[Table 3]

Table 3 presents the key coefficients from models estimating the relationship between income inequality and rank-order income segregation (replications of the full tables are presented in Appendix C, tables 3-5). Each paper measures income inequality using the Gini coefficient. The top panel of Table 3 replicates Reardon and Bischoff (2011), predicting income segregation among large metropolitan areas from 1970 to 2000; the second panel replicates Bischoff and Reardon (2014), which extends the model to 2009. Models include metropolitan area and year fixed effects and metropolitan area-year covariates. As shown in column 1, the published results indicated that a one-point increase on the Gini index corresponded to approximately a one-half-point increase in family income segregation between neighborhoods. Columns 2 and 3 present results from the same model using bias-corrected H^R and R^R; the coefficients in these models are similar in magnitude and statistical significance to the corresponding

previously published estimates. The magnitude of the differences between published and corrected coefficients ranges from -1 percent to +14 percent.

The bottom panel of Table 3 presents results from Owens (2016), who focused on differences between income segregation among households with and without children. Replications of regression analyses from this paper using bias-corrected H^R and R^R produce very similar results to those published. The coefficient for income inequality is positive and of similar magnitude to published results, indicating that changes in income inequality from 1990 to 2010 were positively associated with changes in income segregation among childless households. The positive and significant interaction term between income inequality and families with children across all estimates indicates that the relationship between income inequality and income segregation was more than twice as large among families with children as among childless households. Owens (2016) also investigated whether school district fragmentation (the degree to which a metropolitan area was split up between many school districts) contributed to higher residential segregation among families with children. The results predicting bias-corrected H^R and R^R are consistent with the published results: income segregation among families with children is higher in metropolitan areas that are more fragmented (the coefficient for fragmentation x families with children is significant and positive).

In summary, our replications of previously published regression models using bias-corrected measures of income segregation do not alter any of the substantive conclusions of prior research. The relationships between income segregation and both income inequality and school district fragmentation remain positive, large, and statistically significant when the bias-correction methods for income segregation are used.

Discussion

Our investigations of potential bias in recent income segregation trends demonstrate several

important facts. First, we confirm that sample-based segregation measures are biased upwards. The bias is often moderately large relative to the magnitude of observed differences and changes in segregation. Ignoring the bias may therefore lead to erroneous inferences. Second, we show that it is possible to compute bias-corrected segregation measures that largely eliminate the small sample bias using publicly available Census and ACS tabulations; the correction does not require access to restricted census micro-data nor does it require repeated simulation of micro-data. Third, bias-corrected estimates indicate that roughly half of the increase in family income segregation between 2000 and later ACS years reported in recent research is due to increased upward bias resulting from the lower sampling rate of the ACS relative to the 2000 Census. Nonetheless, the bias-corrected trend indicates that income segregation did rise after 2000, albeit more slowly than has been reported. The increase in family income segregation is largely due to trends among families with children, which we find did increase substantially after 2000, consistent with Owens' (2016) findings. We also find that income segregation among black and Hispanic families increased much more than it did among white families. Furthermore, replications of multivariate analyses from previously published papers confirm that rising income inequality is a primary predictor of increases in residential sorting by income.

The bias in segregation measures that we investigate here is not limited to the study of income segregation. All standard segregation indices (binary, multigroup, ordinal, and rank-order measures) are biased upwards when computed from sample data. Sample-based measures of gender segregation within firms, for example, will be biased upwards—and will be more biased in firms or divisions in which the organizational units are smaller. Likewise, sample-based measures of racial/ethnic segregation among organizations (such as churches, schools, clubs, or firms) will be subject to upward bias. Moreover, the bias we describe is not limited to the two segregation measures we focus on here; similar bias is present in the dissimilarity and Gini indices of segregation, though we do not derive formulas for their biases here. Segregation measures are not subject to the form of bias we describe here, however, when

computed from full population data. Racial segregation measures computed from the decennial Census, for example, are based on race counts from the full population enumeration, and so are not subject to the sample-induced bias we describe here. Likewise, racial/ethnic school segregation measures computed from population-based racial/ethnic enrollment counts—such as in the Common Core of Data (CCD)—are not subject to sampling bias.

We show that the sampling bias is inversely related to the average unit size and the sampling rate. This has implications not just for the measurement of income segregation trends, but for any comparison involving either different sampling rates or units of different average size. For example, a comparison of income segregation in different countries will be biased if the countries use a different sampling rate or if the geographic units (the equivalents of census tracts in each country) differ in size. A comparison of between-tract to between-block group segregation will be biased because the average tract population is roughly three times larger than the average block group population. It follows from this that any sample-based decomposition of segregation into within- and between-unit components will overstate the within-unit component. And a comparison of segregation in two different subpopulations—such as income segregation among individuals over age 65 and among those younger than 65, for example—will be biased if the subpopulations have different average within-tract sizes.

The biases in $\hat{R}$ and $\hat{R}^R$ are also inversely related to the true level of segregation. This means, for example, that if we wish to compare the change in segregation between two metropolitan areas, one highly segregated and one much less segregated, the estimated change will be biased upwards more in the less segregated metropolitan area than the more segregated one. And finally, the bias in binary $\widehat{H}$ is inversely related to the entropy E. So an estimate of the segregation between the bottom 10 percent of earners and all others will be more upwardly biased than an estimate of the segregation between the bottom and the top half of the income distribution.

The bias-corrected estimators we describe here provide a method of obtaining approximately

unbiased estimates in many cases of practical interest. Researchers may use these bias-corrected estimators to make valid comparisons in the types of cases listed above. In the substantive cases of interest in this paper, the methods we describe here allow comparisons of income segregation across years with different sampling rates, metropolitan areas with different overall levels of segregation, and subpopulations of different size. These methods do not require access to micro data. Instead, one need only know the total unit populations and the harmonic mean of the unit sampling rates, which can be estimated with publicly available data. With these, it is straightforward to implement the bias-corrected estimators we describe. We have written a set of Stata commands, "seg" and "rankseg," which perform the bias correction. These are publicly available via the Boston College Statistical Software Components (SSC) archive.

As we show, however, there are some cases where the bias-corrected estimators fail to provide accurate estimates. When sample sizes are small on average and highly variable, the estimators may fail to provide unbiased estimates. We advise caution in such cases; Figure 4 and the derivations in Appendix A may provide researchers with some guidance regarding the potential magnitude of the bias and extent to which our estimators eliminate this bias.

Note also that the methods and estimates we describe here provide bias correction due to bias that arises in the case of sampling without replacement in each geographic unit, where the sampling rate in each tract is known. The approach can accommodate variation in sampling rates across units. Simple adjustments to the formulas can accommodate cases where sampling is done with replacement. But our methods do not address several other factors that complicate the estimation of income segregation. Our method does not explicitly address cases where sampling probabilities vary within a unit (as is implicitly the case when sampling weights are used to account for various forms of non-response) or where missing income data are imputed, reducing the effective sample size. In such a case, however, if the effective sampling rate in each unit is known, it can be used in place of the simple sampling rate in the bias-

correction formulas. Our discussion here also does not address bias that arises from data suppression. For example, Logan et al (forthcoming) report that the black household income distribution is not reported in public ACS tabulations for 17,000 census tracts that together contain 1.2 million black individuals. Missing income data for black families in these tracts may bias estimates of income segregation among black families, though it is likely that the bias is small given that these households make up only 3% of the total US black population.

Finally, our concern in this paper is the bias in sample-based estimates of segregation. We have not addressed the issue of sampling variability in segregation estimates. Because the substantive focus of this paper is on average trends in segregation among many metropolitan areas, the error in each individual metropolitan area's estimated income segregation that arises from sampling variability is a secondary concern. In our estimates of average trends, uncertainty due to sampling variability is captured in the standard errors of the regression models. But in other contexts, one might want to know whether income segregation changed in one particular place. In that case, it is essential to quantify the sampling variance in segregation estimates. We know of no published research that provides formulas describing the sampling variability of segregation measures, however (though Reardon (2011) does provide formulas for the error that arises due to model (polynomial order) uncertainty). In part, this is because most methodological research on segregation indices has focused on racial segregation, where full population data have been widely available, obviating concerns about sampling variability. Sample-based segregation estimates, however, may have substantial sampling variance. Future work should therefore aim to construct standard errors for sample-based segregation estimators.

We conclude with some practical guidelines for estimating segregation. First, if the data are based on samples rather than full populations, segregation estimates will be subject to bias. In general, lower sampling rates and smaller unit populations will yield larger bias. Second, the formulas above (particularly Equations 7, 11, 12, and 13) allow one to estimate the degree of bias. If the expected bias is

large enough to impact the inferences of interest, the bias-correction methods we propose will be useful. Third, if unit-specific sampling rates are not available, we recommend using instead the overall or average sampling rate in the larger population; unless sampling rates are highly variable, this will yield bias-corrected estimates that are very close to what would be obtained using unit-specific sampling rate information. Fourth, when estimating binary segregation measures, if the group proportions are near 0 or 1, we recommend using R instead of H (see Figure 1), given the poor performance of the bias-correction of H in such cases. Fifth, when estimating rank-order segregation measures, the bias correction work well for both R^R and H^R under most conditions. Finally, we advise caution when estimating (binary or rank-order) segregation in cases when the unit populations are small on average (our rule of thumb here is 200), highly variable, and correlated with the interaction or entropy indices, as is the case when estimating within-race income segregation under conditions of substantial racial and economic segregation. With the exception of these conditions, the bias-correction methods appear to satisfactorily address the concern about bias in sample-based segregation estimates.

References

Bischoff, Kendra, and Sean F. Reardon. 2014. "Residential Segregation by Income, 1970-2009." in *Diversity and Disparities: America Enters a New Century*, edited by John Logan. New York: The Russell Sage Foundation.

Brooks-Gunn, Jeanne, Greg J. Duncan, and J. Lawrence Aber (Eds.). 1997. *Neighborhood poverty: Context and consequences for children*. New York: Russell Sage Foundation.

Bureau of the Census. 2014. "American Community Survey Design and Methodology (January 2014): Chapter 4: Sample Design and Selection." Washington, DC: U.S. Census Bureau.

—. n.d. "American Community Survey Accuracy of the Data (2015)." Washington, DC: U.S. Census Bureau.

Chetty, Raj, and Nathaniel Hendren. 2015. "The impacts of neighborhoods on intergenerational mobility: Childhool exposure effects and county-level estimates." National Bureau of Economic Research.

Chetty, Raj, Nathaniel Hendren, and Lawrence F. Katz. 2015. "The effects of exposure to better neighborhoods on children: New evidence from the Moving to Opportunity experiment." Harvard University.

James, David R., and Karl E. Taeuber. 1985. "Measures of segregation." *Sociological Methodology* 14:1-32.

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. forthcoming. "The Uptick in Income Segregation: Real Trend or Random Sampling Variation?" *American Journal of Sociology*.

Napierala, Jeffrey, and Nancy Denton. 2017. "Measuring Residential Segregation With the ACS: How the Margin of Error Affects the Dissimilarity Index." *Demography* 54(1):285-309.

National Research Council. 2007. *Using the American community survey: Benefits and challenges*. Washington, DC: National Academies Press.

Owens, Ann. 2016. "Inequality in Children's Contexts Income Segregation of Households with and without Children." *American Sociological Review* 81(3):549-74.

Reardon, Sean F, and Kendra Bischoff. 2016. "The Continuing Increase in Income Segregation, 2007–2012." *CEPA working paper series*.

Reardon, Sean F. 2011. "Measures of Income Segregation." in *CEPA Working Papers*. Stanford, CA: Stanford Center for Education Policy Analysis.

Reardon, Sean F., and Kendra Bischoff. 2011. "Income Inequality and Income Segregation." *American Journal of Sociology* 116(4):1092-153.

Reardon, Sean F., and Glenn Firebaugh. 2002. "Measures of multi-group segregation." *Sociological Methodology* 32:33-67.

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

Winship, Christopher. 1977. "A reevaluation of indexes of residential segregation." *Social Forces* 55(4):1058-66.

Table 1. Mean Uncorrected and Bias-Corrected Estimates of Rank-Order Income Segregation, 116 Largest Metropolitan Areas, by Household Type and Year

Year	Uncorrected H^R	Bias-Corrected H^R	Uncorrected R^R	Bias-Corrected R^R
		All Families		
1970	0.097	0.093	0.110	0.106
	(0.027)	(0.027)	(0.031)	(0.031)
1980	0.094 *	0.088 ***	0.106 *	0.101 **
	(0.027)	(0.027)	(0.030)	(0.030)
1990	0.116 ***	0.111 ***	0.129 ***	0.124 ***
	(0.029)	(0.029)	(0.032)	(0.032)
2000	0.116	0.111	0.130	0.126
	(0.027)	(0.026)	(0.030)	(0.030)
2007	0.125 ***	0.114 *	0.138 ***	0.128
	(0.028)	(0.027)	(0.031)	(0.031)
2014	0.126 ***	0.115 **	0.140 ***	0.130 *
	(0.027)	(0.027)	(0.030)	(0.030)
		All Households		
1970	0.100	0.097	0.114	0.112
	(0.024)	(0.024)	(0.027)	(0.271)
1980	0.092 ***	0.089 ***	0.105 ***	0.102 ***
	(0.022)	(0.022)	(0.025)	(0.025)
1990	0.101 ***	0.097 ***	0.113 ***	0.110 ***
	(0.023)	(0.023)	(0.026)	(0.026)
2000	0.098 **	0.094 **	0.110 ^	0.107 *
	(0.020)	(0.020)	(0.023)	(0.023)
2007	0.103 ***	0.096	0.115 **	0.108
	(0.021)	(0.021)	(0.023)	(0.023)
2014	0.103 ***	0.096	0.115 **	0.108
	(0.021)	(0.021)	(0.023)	(0.023)
		Families with Children		
1990	0.157	0.146	0.173	0.164
	(0.038)	(0.037)	(0.041)	(0.041)
2000	0.161 **	0.150 **	0.179 ***	0.169 ***
	(0.034)	(0.034)	(0.038)	(0.038)
2007	0.188 ***	0.165 ***	0.204 ***	0.184 ***
	(0.038)	(0.036)	(0.040)	(0.040)
2014	0.200 ***	0.175 ***	0.216 ***	0.195 ***
	(0.037)	(0.037)	(0.040)	(0.041)
		Households without Children		
1990	0.088	0.082	0.097	0.092
	(0.020)	(0.020)	(0.022)	(0.022)
2000	0.080 ***	0.074 ***	0.089 ***	0.084 ***
	(0.017)	(0.017)	(0.019)	(0.019)
2007	0.085 ***	0.074	0.093 ***	0.083
	(0.017)	(0.016)	(0.019)	(0.018)
2014	0.084 ***	0.074	0.092 ***	0.083
	(0.016)	(0.016)	(0.018)	(0.018)

Notes: Cells report mean estimated segregation; standard deviations are in parentheses below the means. Sample is 116 metropolitan areas with over 500,000 residents as of 2007 (Cape Coral is excluded due to missing data in 1970). 2007 refers to 2005-09 ACS; 2014 refers to 2012-16 ACS. ^p≤0.10; *p≤0.05; **p≤0.01; ***p≤0.001; statistical significance tests come from regression models with metropolitan area fixed effects that compare the estimate in each decade to the prior decade. Note that 2007 and 2014 are both compared to the 2000 estimate.

Table 2. Comparison of Published and Bias-Corrected Estimates of Rank-Order Income Segregation, Large Metropolitan Areas, by Household Type and Year

Year	Published H^R	Bias-Corrected H^R
	All Families	
1970	0.115	0.093
	(0.027)	(0.027)
1980	0.112 *	0.088 ***
	0.027	0.027
1990	0.134 ***	0.111 ***
	(0.029)	(0.029)
2000	0.135	0.111
	(0.027)	(0.026)
2007	0.143 ***	0.114 *
	(0.028)	(0.027)
2012	0.146 ***	0.116 ***
	(0.027)	(0.027)
	All Households	
1990	0.123	0.101
	(0.021)	(0.021)
2000	0.120 **	0.098 ***
	(0.018)	(0.018)
2010	0.126 ***	0.101 **
	(0.019)	(0.019)
	Families with Children	
1990	0.171	0.153
	(0.031)	(0.034)
2000	0.179 ***	0.157 *
	(0.029)	(0.030)
2010	0.209 ***	0.179 ***
	(0.031)	(0.033)
	Households without Children	
1990	0.107	0.085
	(0.017)	(0.018)
2000	0.105 *	0.077 ***
	(0.015)	(0.015)
2010	0.107 **	0.078
	(0.015)	(0.015)

Notes: Cells report means; standard deviations are in parentheses below the means. ^p≤0.10; *p≤0.05; **p≤0.01; ***p≤0.001; statistical significance tests come from regression models with metropolitan area fixed effects that compare the estimate in each decade to the prior decade. Note that 2007 and 2012 are both compared to the 2000 estimate. Top panel: Sample is 116 metropolitan areas with over 500,000 residents in 2007 (Cape Coral is excluded due to missing data in 1970). Published estimates from Bischoff and Reardon (2014) and Reardon and Bischoff (2016). 2007 refers to 2005-09 ACS and 2012 refers to 2010-14 ACS. Bottom three panels: Sample is 100 most populous metropolitan areas in 2010. Published estimates from Owens (2016). 2010 refers to 2008-12 ACS.

Table 3. Replications of Models Estimating the Relationship between Income Inequality and Rank-Order Income Segregation

	Published Estimates (using published H^R)	Estimates Using Bias-Corrected H^R	Estimates Using Bias-Corrected R^R
	Reardon and Bischoff (2011)		
Income Inequality (Gini)	0.561 ***	0.480 ***	0.528 ***
	(0.085)	(0.124)	(0.142)
	Bischoff and Reardon (2014)		
Income Inequality (Gini)	0.443 ***	0.448 ***	0.500 ***
	(0.090)	(0.083)	(0.098)
	Owens (2016)		
Income Inequality (Gini)	0.232 ***	0.194 **	0.230 **
	(0.069)	(0.075)	(0.081)
Income Inequality x Families with Children	0.223 ***	0.260 ***	0.275 ***
	(0.057)	(0.061)	(0.066)
Fragmentation x Families with Children	0.028 **	0.031 **	0.034 **
	(0.010)	(0.011)	(0.012)

Notes: Top panel: Models include metropolitan area fixed effects, year indicators, and metropolitan-year covariates (population, unemployment rate, proportion under age 18, proportion over age 65, proportion with high school diploma, proportion foreign born, proportion female headed families, per capita income, proportions employed in manufacturing, construction, financial and real estate, professional and managerial jobs, and proportions of housing built within ten, five, and one years). Bootstrapped standard errors in parentheses. Reported results published in Reardon and Bischoff 2011, Table 4, "All Families" model, predicting H^R. Sample includes 100 largest metropolitan areas in 2000; data from 1970, 1980, 1990, and 2000.

Middle panel: Models include metropolitan area fixed effects, year indicators, and metropolitan-year covariates (log population, age composition, residents' educational attainment, unemployment rate, proportion employed in manufacturing, per capita income, racial composition, foreign-born composition, and female-headed family rate). Bootstrapped standard errors in parentheses. Reported results published in Bischoff and Reardon 2014, Table 5, predicting H^R. Sample includes 117 metropolitan areas with population >500,000 in 2007 (minus Cape Coral in 1970); data from 1970, 1980, 1990, 2000, and 2007-11.

Bottom panel: Models include metropolitan area fixed effects, district fragmentation in 1990 x year, group (families with children) x fragmentation x year, group and year indicators and their interaction, and group-metro-year covariates and their interaction with group (log population, age composition, female-headed household rate, racial composition, racial segregation, foreign born composition, residents' educational attainment, unemployment rate, proportion employed in manufacturing, and private school enrollment rate). Reported results published in Owens 2016, Table 4, Model 2, predicting H^R. Sample includes 95 largest metropolitan areas in 2010 with more than 1 school district; data from 1990, 2000, and 2008-2012.

^p≤0.10; *p ≤0.05; **p≤0.01; ***p≤0.001

Figure 1: Bias in Uncorrected and Bias-Corrected Binary H and R at 8% Sampling Rate, 380 Metropolitan Areas, by Income Percentile

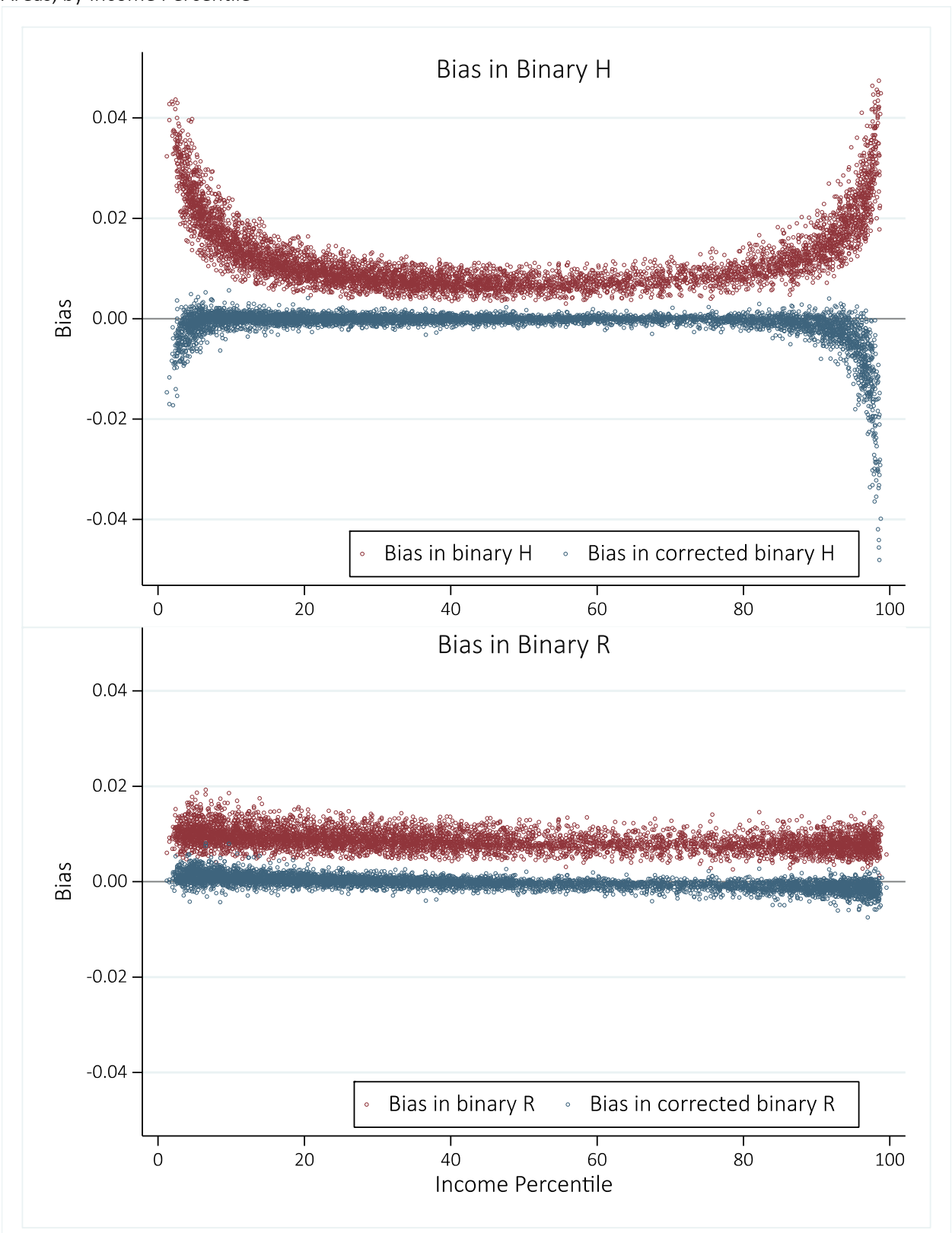


Figure 2: Average Bias in Uncorrected and Bias-Corrected Binary H and R, 380 Metropolitan Areas, by Income Percentile and Sampling Rate

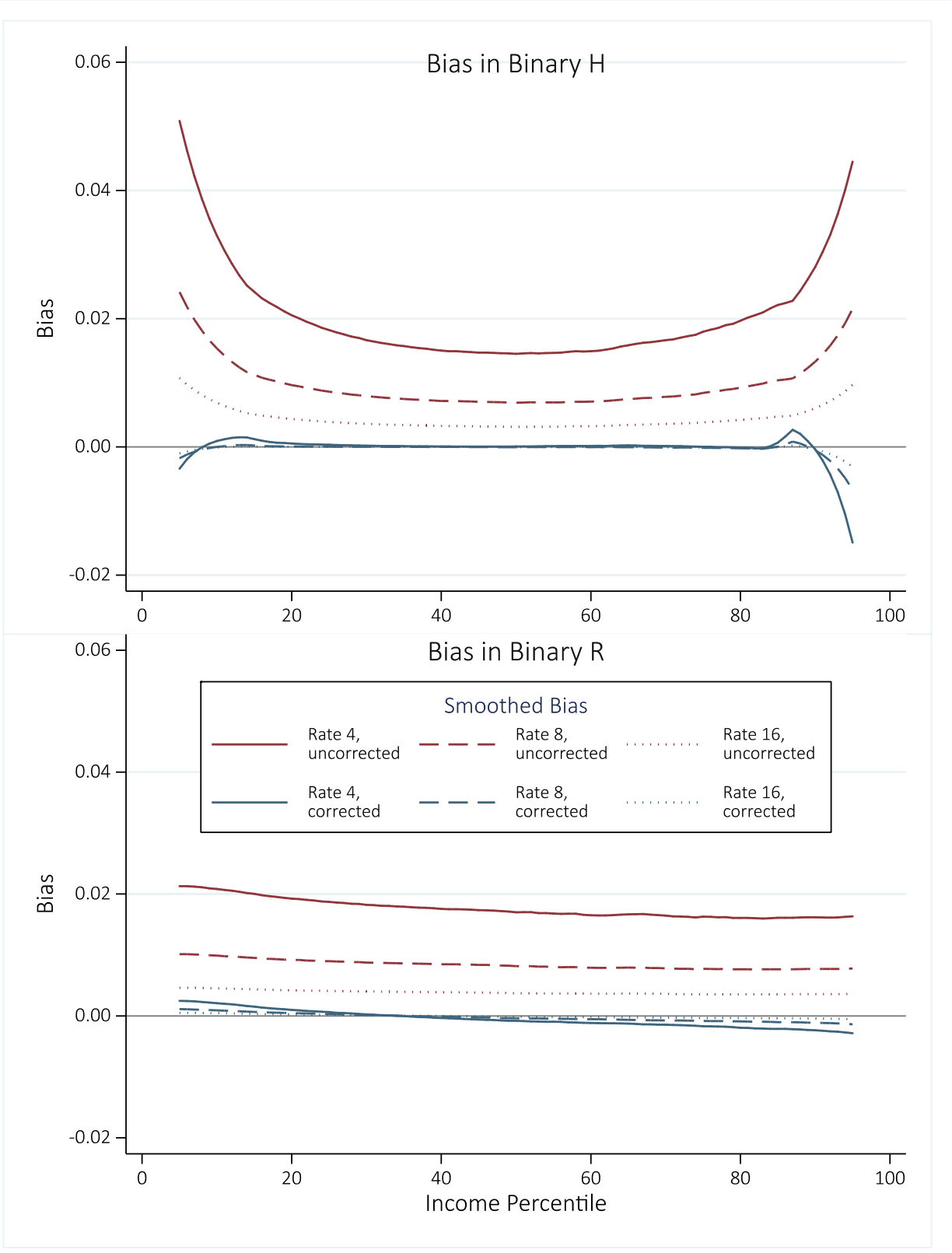


Figure 3: Average Bias in Uncorrected and Bias-Corrected Rank-Order H^R and R^R, 380 Metropolitan Areas, by Sampling Rate

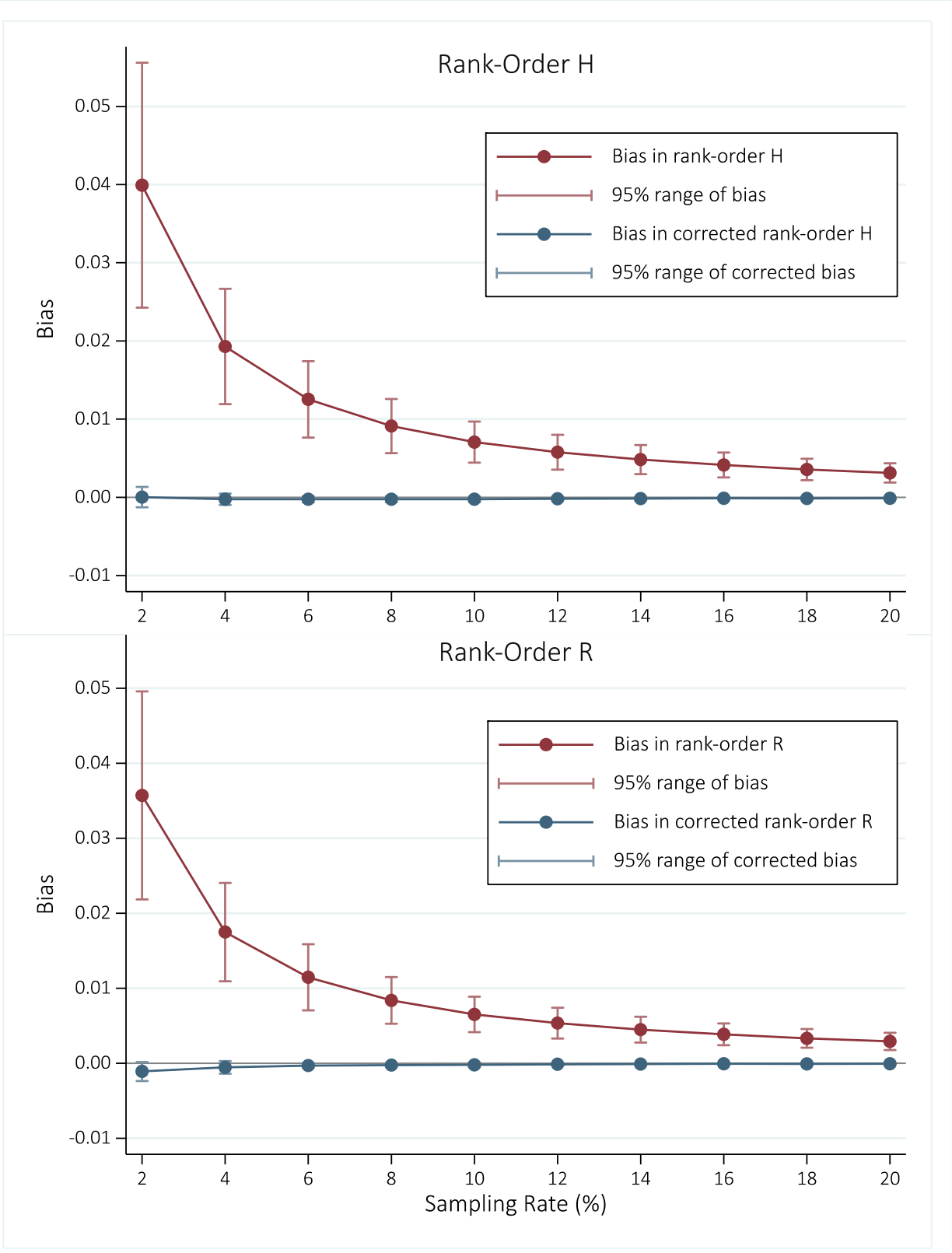


Figure 4: Bias in Uncorrected and Bias-Corrected Rank-Order Measures of Income Segregation at 8% Sampling Rate, Among Black Families, 380 Metropolitan Areas, by Mean Tract Size

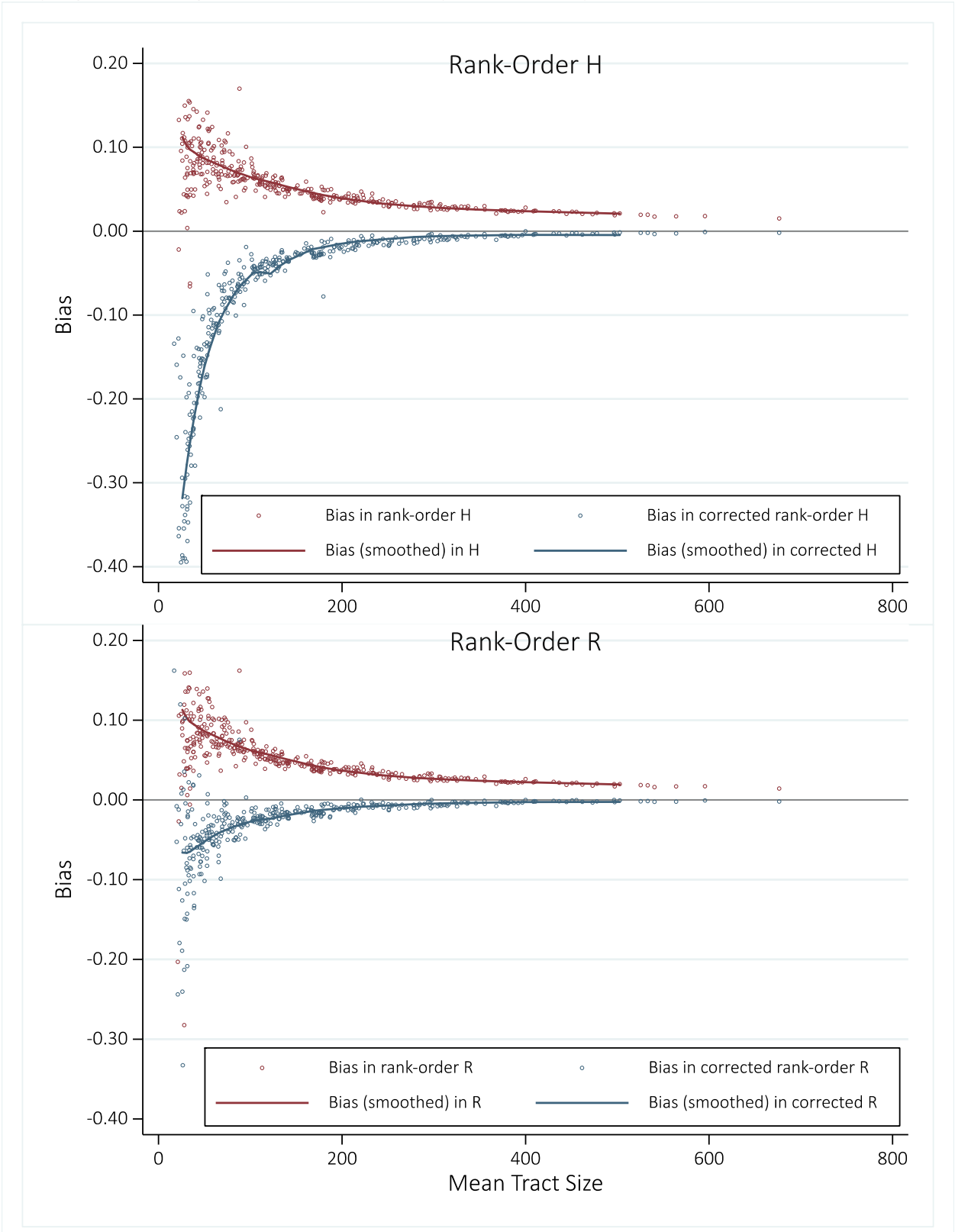
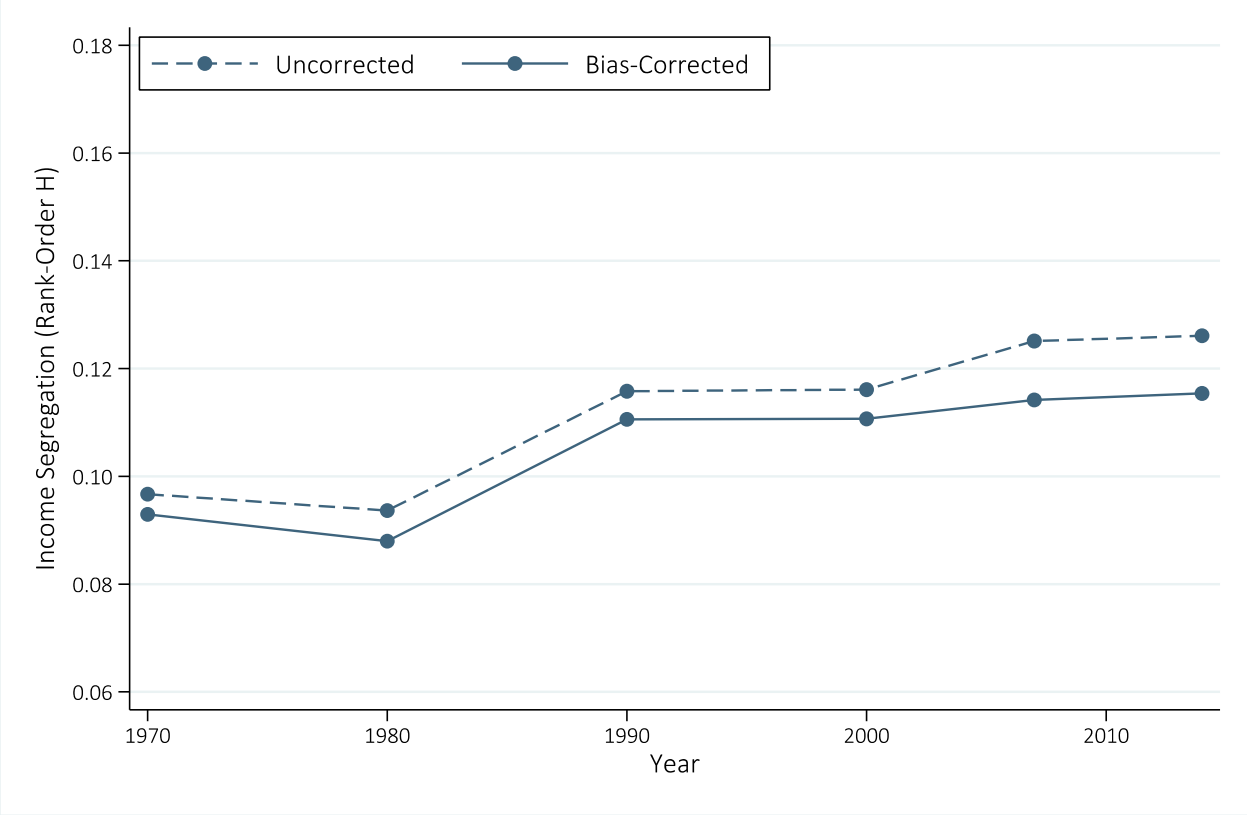
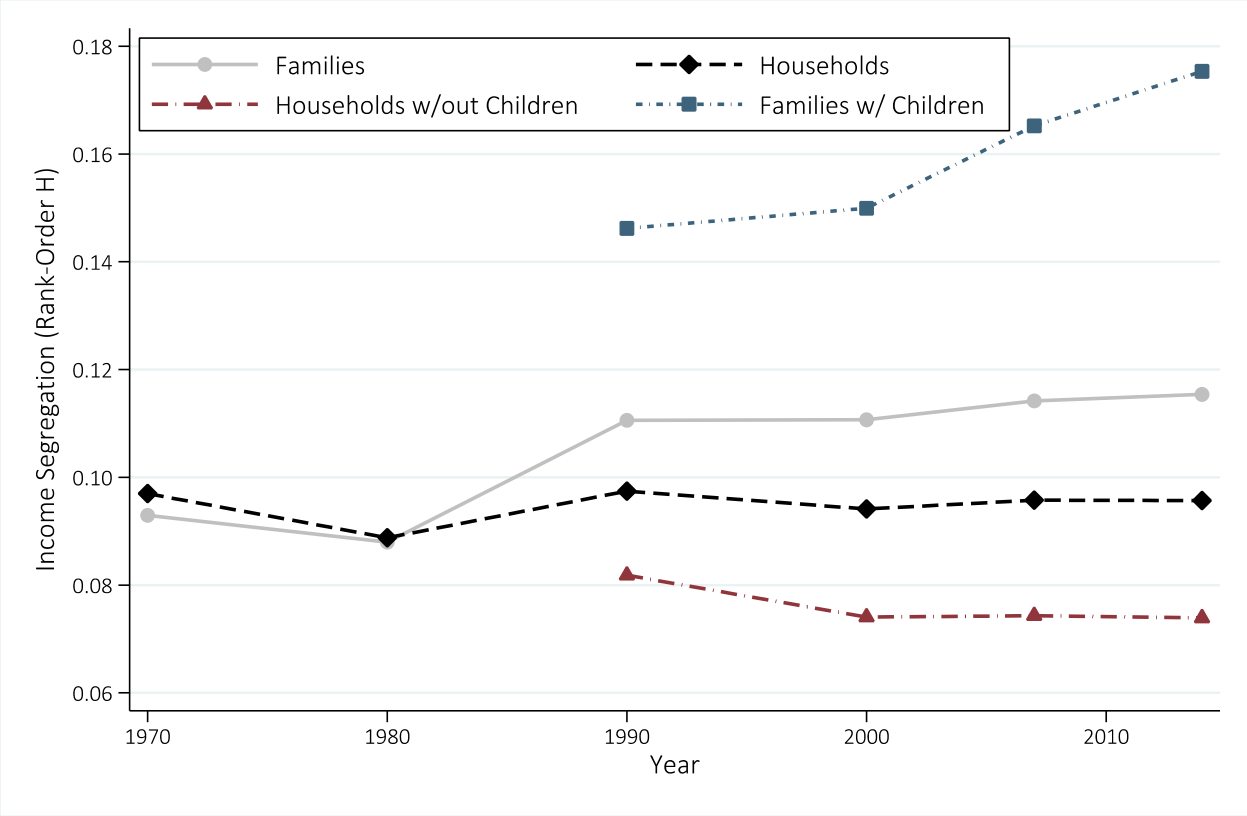


Figure 5. Mean Uncorrected and Bias-Corrected Estimates of Income Segregation (Rank-Order *H*) among Families, 116 Largest Metropolitan Areas, 1970 to 2014



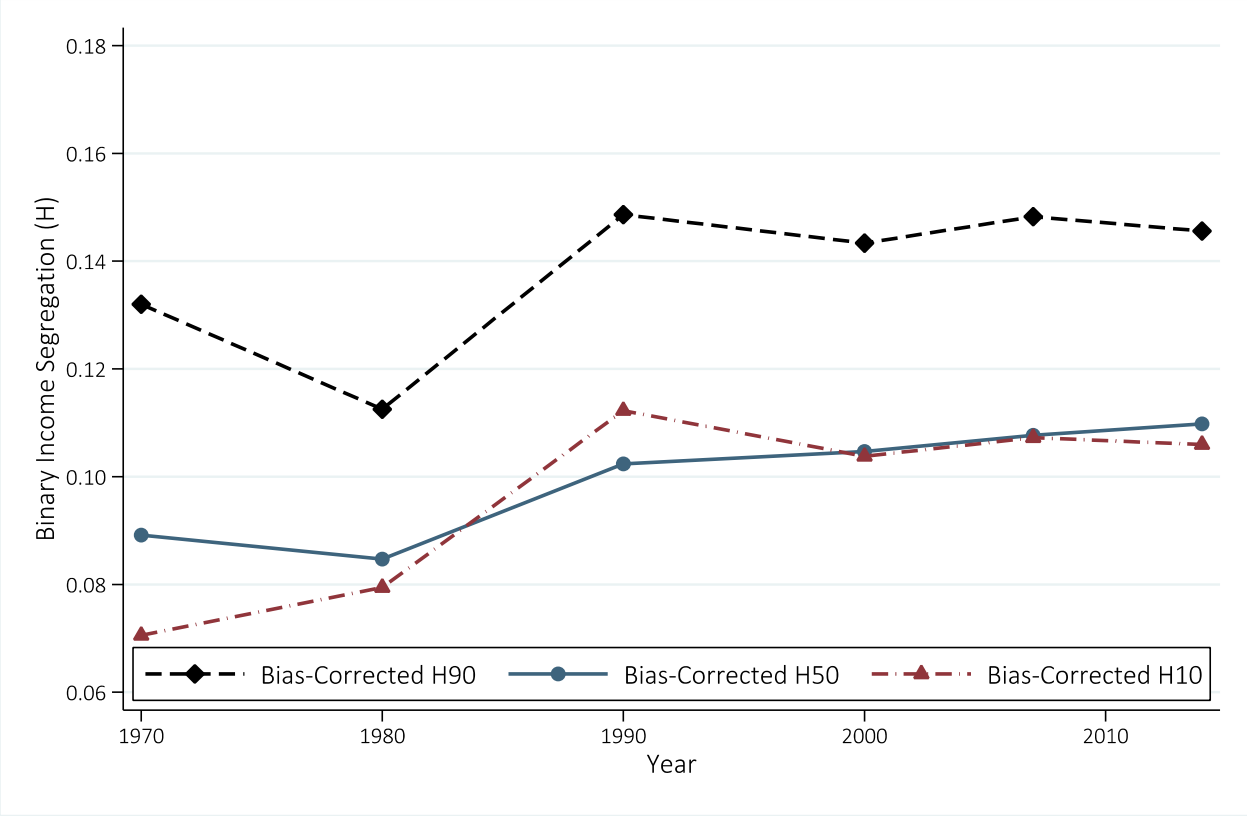
Note: Estimates are from Table 1, which reports statistical significance of changes over time.

Figure 6. Mean Bias-Corrected Estimates of Income Segregation (Rank-Order *H*) among Households, Families, Families with Children, and Households without Children, 116 Largest Metropolitan Areas, 1970 to 2014



Note: Estimates are from Table 1, which reports statistical significance of changes over time.

Figure 7. Mean Bias-Corrected Estimates of *H10*, *H50*, and *H90* among Families, 116 Largest Metropolitan Areas, 1970 to 2014



Note: Estimates are from Appendix Table C1, which reports statistical significance of changes over time.

Appendix A: Derivations of sampling bias in segregation measures

A.1. Definitions

To start, we define some notation. We first are interested in computing segregation between two groups among a set of J units (e.g., census tracts). Let p denote the group proportion in a given unit. For values of $p \in [0,1]$, define the *Interaction index* (I) and *Entropy* (E):

$$I = p(1-p)$$

$$E = -[p \ln p + (1-p)\ln(1-p)],$$

where we define $0 \ln 0 = 0$. Note that both I and E are concave down functions of p, a feature that leads estimates of both to be biased when p is estimated from a sample.[17]

The binary variance ratio and information theory segregation indices are respectively defined as

$$R = 1 - \frac{1}{I}\sum_{j=1}^{J}\frac{t_j}{T}I_j$$

and

$$H = 1 - \frac{1}{E}\sum_{j=1}^{J}\frac{t_j}{T}E_j\,,$$

where I and E are the values of I and E in the whole population; where I_j and E_j are the values of I and E in unit j; and where t_j/T is the share of the population in unit j (Reardon 2011).

If y is some ordered variable such as income, the corresponding rank-order income segregation indices are

$$H^R = \frac{1}{\int_0^1 E(r)dr}\int_0^1 E(r)H(r)dr \;= 2\int_0^1 E(r)H(r)dr$$

[17] To see this, note that if $\hat{p}$ is a random variable with $E[\hat{p}] = p$, then the second-order Taylor expansion of a concave down function $f(\hat{p})$ yields $E[f(\hat{p})] = f(p) + \frac{1}{2}f''(p)Var(\hat{p}) \leq f(p)$, since f'' is negative everywhere.

$$R^R = \frac{1}{\int_0^1 I(r)dr} \int_0^1 I(r)R(r)dr \; = 6 \int_0^1 I(r)R(r)dr \,,$$

where $I(r)$, $R(r)$, $E(r)$, and $H(r)$ are the values of I, R, E, and H when the population is divided into groups defined by whether y is above or below the $100 \times r^{th}$ percentile of y. For example, $H(.5)$ is the value of H computed between those with above and below median values of y. The rank-order measures are weighted integrals of the binary indices over values of $r \in (0,1)$ (Reardon 2011).

The formulas above assume we observe p and t without error in each unit. Instead, here we assume we know t with certainty but must estimate p from a sample.[18] More specifically, from each unit $j \in \{1, \ldots, J\}$ we observe a simple random sample of size n_j, drawn without replacement from the population in the unit, which is of known finite size t_j. Let $T = \sum_{j=1}^{J} t_j$ denote the total population across the J units, and let $\bar{t} = \frac{T}{J}$ denote the average unit population size. Let $r_j = n_j / t_j$ denote the sampling rate in unit j. Let $\widetilde{t-1} = \left[\frac{1}{J}\sum_{j=1}^{J} \frac{1}{t_j - 1}\right]^{-1}$ denote the harmonic mean of $t_j - 1$. Let $\tilde{r} = \left(\frac{1}{J}\sum_{j=1}^{J} \frac{1}{r_j}\right)^{-1}$ indicate the harmonic mean of the units' sampling rates.

Now define the bias factor

$$B = \frac{z}{\bar{t} - 1} \cdot \frac{1 - \tilde{r}}{\tilde{r}},$$

where z is a function of the ratio of the arithmetic and harmonic means of $t_j - 1$:

$$z = 1 + \frac{1}{\bar{t}} \left(\frac{\bar{t} - 1}{\widetilde{t-1}} - 1 \right).$$

Note that if the sampling rate $r_j = r$ is constant across units, the bias factor will be

$$B = z \cdot \frac{1 - r}{\bar{n} - r}.$$

[18] As we show below, this is actually a stronger assumption than we need; we require only that the estimates of t_j be unbiased and that the error in the estimate of $\hat{t}_j$ be uncorrelated with the error in the estimated diversity $\hat{I}_j$ or $\hat{E}_j$. Error in $\hat{t}_j$ will contribute to sampling variance in estimated segregation, but will not contribute to bias, so long as the error in $\hat{t}_j$ is independent of the error in $\hat{I}_j$ or $\hat{E}_j$.

Finally, note that, assuming at least one person is sampled per unit, $0 \leq B \leq z$.[19]

We show below that the sampling bias in segregation measures is approximately proportional to B when the unit populations t_j are moderately large. B is a decreasing function of both the average unit size $\bar{t}$ and the harmonic mean of the sampling rate across units, $\tilde{r}$. That is, we show that bias in segregation measures is smaller as unit population increases and/or as sampling rates increase.

A.2. Expected value of $\hat{I}$:

The complete Taylor expansion of $\hat{I}$ around p is:

$$\begin{aligned}\hat{I} &= \hat{p}(1-\hat{p}) \\ &= p(1-p) + (1-2p)(\hat{p}-p) - (\hat{p}-p)^2.\end{aligned}$$

In addition, note:

$$E[(\hat{p}-p)] = 0$$

$$E[(\hat{p}-p)^2] = bp(1-p),$$

where $b = \frac{1-r}{r(t-1)}$ under sampling without replacement.[20] From this it follows that

$$\begin{aligned}\mathrm{E}\left[\hat{I}\right] &= I - bp(1-p) \\ &= (1-b)I.\end{aligned}$$

A.3. Expected value of $\hat{E}$:

The second-order Taylor expansion of $\hat{E}$ is:

$$\hat{E} = -p\ln(p) - (1-p)\ln(1-p) + \ln\left(\frac{1-p}{p}\right)(\hat{p}-p) - \frac{1}{2p(1-p)}(\hat{p}-p)^2 + e$$

So

[19] Note that if we assume sampling with replacement, we define the bias factor instead as $B = \frac{1}{\bar{t}\tilde{r}}$. In this case, if $r_j = r$ is constant across units, then $B = \frac{1}{n}$. Using this definition of B, all the derivations below hold under sampling with replacement. We do not show these derivations in the interest of space.

[20] When sampling with replacement, $b = 1/n$.

$$\begin{aligned}\mathrm{E}\left[\hat{E}\right] &= E - \left[\frac{1}{2p(1-p)}\right] bp(1-p) + E(e) \\ &= E - \frac{b}{2} + E(e) \\ &\approx E - \frac{b}{2}.\end{aligned}$$

Simulations (not shown) show that this approximation for $E\left[\hat{E}\right]$ is inaccurate for very large or small values of p (generally when $p < \frac{b}{4}$ or $p > 1 - \frac{b}{4}$), and is only approximate for other values of p. Nonetheless, approximations based on higher-order Taylor expansions did not perform better in our simulations, so we use the second-order approximation.

A.4. Expected value of $\hat{R}$

Here we assume T and J are large, so that $E\left[\frac{1}{\hat{I}}\right] \approx \frac{1}{I}$ and $Cov\left(\hat{I}, \hat{I}_j\right) \approx 0$. We also assume that the sampling rate is independent of unit size or diversity, so that $\frac{1-r_j}{r_j} \perp I_j$ and $\frac{1-r_j}{r_j} \perp t_j$. We denote the population covariance of t_j and I_j as $C = \frac{1}{J}\sum_{j=1}^{J}\left(t_j - \bar{t}\right)\left(I_j - \bar{I}\right)$. The expected value of $\hat{R}$ will then be:[21]

$$\begin{aligned}E\left[\hat{R}\right] &= 1 - E\left[\frac{1}{\hat{I}}\sum_{j=1}^{J}\frac{t_j}{T}\hat{I}_j\right] \\ &= 1 - E\left[\frac{1}{\hat{I}}\right] \cdot E\left[\sum_{j=1}^{J}\frac{t_j}{T}\hat{I}_j\right] - Cov\left(\frac{1}{\hat{I}}, \sum_{j=1}^{J}\frac{t_j}{T}\hat{I}_j\right) \\ &\approx 1 - \frac{1}{I}\sum_{j=1}^{J}\frac{t_j}{T}E\left[\hat{I}_j\right] \\ &= 1 - \frac{1}{I}\sum\frac{t_j}{T}\left(1 - b_j\right)I_j \\ &= R + \frac{1}{I}\left[\sum\frac{t_j - 1}{T}\left(\frac{1-r_j}{r_j\left(t_j - 1\right)}\right)I_j + \sum\frac{1}{T}\left(\frac{1-r_j}{r_j\left(t_j - 1\right)}\right)I_j\right]\end{aligned}$$

[21] Note that if we do not assume that t_j is known, but instead assume that we observe $\hat{t}_j$, an unbiased estimate of t_j, then the third line of the derivation below will include a term that includes the summation $\sum_{j=1}^{J} cov\left(\hat{t}_j, \hat{I}_j\right)$; under the assumption that sampling error in $\hat{t}_j$ is independent of the sampling error in $\hat{I}_j$, this term will be zero. So the assumption that t_j is known with certainty is not essential to the derivation. The same is true in the derivation below of the expected value of $\hat{H}$.

$$
\begin{aligned}
&\approx R + \frac{1-\tilde{r}}{\tilde{r}\bar{t}I}\frac{1}{J}\sum I_j + \frac{1-\tilde{r}}{\tilde{r}\bar{t}I}\frac{1}{J}\sum \frac{1}{t_j - 1} I_j \\
&\approx R + \frac{1-\tilde{r}}{\tilde{r}\bar{t}I}\bar{I} + \frac{1-\tilde{r}}{\tilde{r}\bar{t}I}\left[\frac{1}{\widetilde{t-1}}\bar{I} - \frac{1}{J(\bar{t}-1)^2}\sum\left(t_j - \bar{t}\right)\left(I_j - \bar{I}\right)\right] \\
&= R + \frac{1-\tilde{r}}{\tilde{r}\bar{t}I}\left(1 + \frac{1}{\widetilde{t-1}}\right)\bar{I} - \frac{1-\tilde{r}}{\tilde{r}\bar{t}I(\bar{t}-1)^2}C \\
&= R + \frac{1-\tilde{r}}{\tilde{r}\bar{t}I}\left(1 + \frac{1}{\widetilde{t-1}}\right)\left[I(1-R) - \frac{1}{\bar{t}}C\right] - \frac{1-\tilde{r}}{\tilde{r}\bar{t}I(\bar{t}-1)^2}C \\
&= R + \frac{1-\tilde{r}}{\tilde{r}\bar{t}}\left(1 + \frac{1}{\widetilde{t-1}}\right)(1-R) - \frac{1-\tilde{r}}{\tilde{r}\bar{t}^2 I}\left(\left(1 + \frac{1}{\widetilde{t-1}}\right) + \frac{\bar{t}}{(\bar{t}-1)^2}\right)C \\
&= R + B(1-R) - \frac{B}{\bar{t}I}\left(1 + \frac{\bar{t}}{\left(1 + \frac{1}{\widetilde{t-1}}\right)(\bar{t}-1)^2}\right)C \\
&= R + B(1-R) - \frac{B}{\bar{t}I}\left(1 + \frac{1}{z(\bar{t}-1)}\right)C \\
&= R + B(1-R) - \frac{B}{(\bar{t}-1)I}\left(1 + \frac{1-z}{\bar{t}z}\right)C.
\end{aligned}
$$

When $\bar{t}$ and $\widetilde{t-1}$ are large, $z \approx 1$ and $\frac{1-z}{\bar{t}z} \approx \frac{1}{\bar{t}^2} \approx 0$, so we have

$$E\left[\hat{R}\right] \approx R + B(1-R) - \frac{B}{(\bar{t}-1)I}C.$$

Conditional on $\tilde{r}$, the first bias term is approximately inversely proportional to $\bar{t} - 1$. The second is approximately inversely proportional to $(\bar{t}-1)^2 I$. Moreover, if t_j does not vary much among units, C will be small. In general, then, the first bias term is more important that the second, except when I is close to zero and/or the unit populations are small and variable.

A.5. Expected value of $\widehat{H}$

Again, we assume T and J are large, so that $E\left[\frac{1}{\hat{E}}\right] \approx \frac{1}{E}$ and $Cov\left(\hat{E}, \hat{E}_j\right) \approx 0$. We also assume $r_j \perp t_j$. Let e_j^* be the error in the approximation for the bias in $\hat{E}_j$: $e_j^* = E\left[\hat{E}_j\right] - E_j + \frac{b_j}{2}$. The expected value of $\widehat{H}$ will be:

$$E\left[\widehat{H}\right] = 1 - E\left[\frac{1}{\hat{E}}\sum_{j=1}^{J}\frac{t_j}{T}\hat{E}_j\right]$$

$$
\begin{aligned}
&= 1 - E\left[\frac{1}{\hat{E}}\right] \cdot E\left[\sum_{j=1}^{J} \frac{t_j}{T} \hat{E}_j\right] - Cov\left(\frac{1}{\hat{E}}, \sum_{j=1}^{J} \frac{t_j}{T} \hat{E}_j\right) \\
&\approx 1 - \frac{1}{E} \sum_{j=1}^{J} \frac{t_j}{T} E\left[\hat{E}_j\right] \\
&= 1 - \frac{1}{E} \sum_{j=1}^{J} \frac{t_j}{T} \left[E_j - \frac{b_j}{2} + e_j^*\right] \\
&= H + \frac{1}{E} \sum_{j=1}^{J} \frac{t_j}{T} \frac{b_j}{2} - \frac{1}{E} \sum_{j=1}^{J} \frac{t_j}{T} e_j^* \\
&= H + \frac{1}{2E\bar{t}} \cdot \frac{1}{J} \sum_{j=1}^{J} \frac{t_j(1 - r_j)}{(t_j - 1) r_j} - \frac{\bar{e}^*}{E} \\
&= H + \frac{1 - \tilde{r}}{2E\tilde{r}} \frac{1}{\bar{t}} \left(1 + \frac{1}{\widetilde{t - 1}}\right) - \frac{\bar{e}^*}{E} \\
&= H + \frac{B}{2E} - \frac{\bar{e}^*}{E} \\
&\approx H + \frac{B}{2E}
\end{aligned}
$$

The last approximation depends on the assumption that the average error in the approximation of the expected value of the $\hat{E}_j$'s is small relative to E. This will be true if t_j is large, but when t_j is small the second-order Taylor expansion is not a good approximation of $\hat{E}_j$, particularly when E_j is small (when p is near 0 or 1). In such cases—when H measures the segregation of one group that makes up a small proportion of the population from another, and when the units are small—the expression above may not be a good approximation of the expected value of $\widehat{H}$.

A.6. Expected value of $\widehat{R}^R$

The expected value of $\hat{R}^R$ will be:

$$
\begin{aligned}
E\left[\hat{R}^R\right] &= E\left[6 \int_0^1 I(r) \hat{R}(r) dr\right] \\
&= 6 \int_0^1 I(r) E\left[\hat{R}(r)\right] dr \\
&\approx 6 \int_0^1 I(r) \left[R(r) + B\left(1 - R(r)\right) - \frac{B}{(\bar{t} - 1) I(r)} \left(1 + \frac{1 - z}{\bar{t} z}\right) C(r)\right] dr \\
&= R^R + B(1 - R^R) - \frac{6B}{(\bar{t} - 1)} \left(1 + \frac{1 - z}{\bar{t} z}\right) \int_0^1 C(r) dr
\end{aligned}
$$

$$\approx R^R + B(1 - R^R) - \frac{6B}{(\bar{t} - 1)} \int_0^1 C(r)dr\,.$$

A.7. Expected value of $\widehat{H}^R$

The expected value of $\widehat{H}^R$ will be:

$$\begin{aligned} E\left[\widehat{H}^R\right] &= E\left[2\int_0^1 E(r)\widehat{H}(r)dr\right] \\ &= 2\int_0^1 E(r)E\left[\widehat{H}(r)\right]dr \\ &\approx 2\int_0^1 E(r)\left[H(r) + \frac{B}{2E(r)}\right]dr \\ &= H^R + B. \end{aligned}$$

A.8. Correcting segregation measures for small sample bias.

We can compute bias-corrected measures of segregation as follows:

$$\widehat{H}^* = \widehat{H} - \frac{B}{2E}$$

$$\widehat{R}^* = \frac{\widehat{R} - B\left[1 - \frac{1}{(\bar{t} - 1)I}C\right]}{1 - B}$$

These formulas assume we know E, I, and C. Because E and I are estimated from the full sample, they will be estimated precisely and with little bias as long as the total sample is large. We assume $C \approx 0$. Even if this is not strictly accurate, the bias due to non-zero C will be trivial unless the sampling rate is low, the average unit size is small, and I is small.

A.9. Bias in Other Segregation Measures

We focus in this paper on *H* and *R*. Here, however, we briefly demonstrate that other sample-based segregation estimates will also be biased. Both the Dissimilarity and Gini indices can be written as the weighted sums of differences in the absolute values of estimated proportions (James and Taeuber 1985), and the expected value of the absolute value of a random variable will be biased upwards.

The dissimilarity index can be written as a weighted average of $\left|\hat{p}_j - P\right|$ over all tracts j, where P is the proportion of a group in the total population (and where P is assumed known here for simplicity). When $E\left[\hat{p}_j\right] = p_j$ the expected value of this difference is biased upwards relative to its true value. To see this, consider the case where $p \geq P$ (here $\rho(x)$ is the density function describing the sampling distribution of $\hat{p}$):

$$\begin{aligned}
E[|\hat{p} - P|] &= \int_x \rho(x)|x - P|dx \\
&= \int_{x \geq P} \rho(x)(x - P)dx + \int_{x < P} \rho(x)(P - x)dx \\
&= \int_x \rho(x)(x - P)dx + 2\int_{x < P} \rho(x)(P - x)dx \\
&= E[x] - P + 2\int_{x < P} \rho(x)(P - x)dx \\
&= (p - P) + 2\int_{x < P} \rho(x)(P - x)dx \\
&= |p - P| + 2\int_{x < P} \rho(x)(P - x)dx \\
&\geq |p - P|.
\end{aligned}$$

Likewise, in the case where $p < P$, we get

$$\begin{aligned}
E[|\hat{p} - P|] &= \int_x \rho(x)(P - x)dx + 2\int_{x \geq P} \rho(x)(x - P)dx \\
&= |p - P| + 2\int_{x \geq P} \rho(x)(x - P)dx \\
&\geq |p - P|.
\end{aligned}$$

Similarly, the Gini index of segregation can be written a weighted average of $\left|\hat{p}_j - \hat{p}_k\right|$ over all pairs of tracts j and k. The expected value of the absolute difference in two estimated proportions is biased upwards relative to its true value. Consider, for example, the case where $p > q$:

$$E[|\hat{p} - \hat{q}|] = \int_x \int_y \rho(x, y)|x - y|dydx$$

$$= |p - q| + 2\int_x \int_{y>x} \rho(x,y)|y - x| dy dx$$

$$\geq |p - q|.$$

Finally, note that exposure measures are also subject to sampling bias. The exposure of black households to poor households, for example is defined as

$$_bP_p = \sum_j \frac{b_j}{B} p_j,$$

where b_j and B are the number of black households in tract j and the population, respectively, and p_j is the poverty rate in tract j. If b_j and p_j must be estimated from samples, then the expected value of $_bP_p$ will be

$$E\left[\sum_j \frac{\hat{b}_j}{B}\hat{p}_j\right] = {}_bP_p + \sum_j cov(\hat{b}_j, \hat{p}_j).$$

So the estimated exposure measure will be biased to the extent that race and poverty are correlated within tracts. If, on average, the black households in a tract are poorer than the white households in a tract, the covariance term will be positive, and the exposure measure will be upwardly biased.

Appendix B: Additional Results from Simulations

Appendix B presents additional simulation results. Figures B1 and B2 present the uncorrected and bias-corrected binary segregation measures $\hat{R}$ and $\widehat{H}$ as a function of the proportion of families in the metropolitan area below the income threshold used to define the binary measure; these are analogous to Figures 1 and 2. Unlike Figures 1 and 2, however, here we have artificially constrained all tracts in each metropolitan area to have the same population size, in this case, 1000. By holding tract size constant, C (see equation 5) is set to zero. Recall that in the bias correction derivation, we assume that $C = 0$, despite the fact that C is generally positive for high-income threshold and negative for low-income thresholds in the data generating models we use in our simulations. Figures B1 and B2 show that when $C = 0$, the bias-corrected $\hat{R}^*$ estimator performs equally well at all income thresholds. The bias-corrected $\widehat{H}^*$ estimator, however, still performs poorly at income thresholds near the top or bottom of the income distribution, a result of the failure of the second-order Taylor expansion of $\hat{E}$ used in Appendix section A.3.

Table B1 presents the proportion of bias explained in corrected rank-order H^R and R^R by sampling rate, mean tract size, and race. This table provides additional guidance on when the bias-corrected rank-order segregation measures fail to eliminate bias. For "All Families", both $\widehat{H}^{R*}$ and $\hat{R}^{R*}$ perform very well, virtually eliminating all bias, even at low sampling rates and small mean tract sizes. This is not the case for black and Hispanic families. Here it is clear that $\widehat{H}^{R*}$, in particular, does not effectively eliminate bias when the sampling rate is small and the mean tract size is small. This is due to the fact that the black family population is both small and unevenly distributed among geographic units (tracts) with metropolitan areas. For "Black Families" at a 4% sampling rate, it is only at a mean tract size of 300 that more than 70% of the bias is removed. Below this mean tract size, H^{R*} is not an improvement over H^R. A similar pattern is evident for Hispanic families, although the bias correction is an improvement when mean tract size reaches roughly 150. This mirrors a general rule of thumb: H^{R*} typically represents an

improvement over H^R when mean tract size exceeds 200. In nearly all cases, R^{R*} outperforms H^{R*}, and is an improvement over R^R when mean tract size is greater than 100.

The poor performance of the bias-corrected estimators of black and Hispanic income segregation is due to the confluence of three factors: variable population sizes across tracts (meaning that the number of black or Hispanic families varies substantially across tracts, due to racial residential segregation patterns); small within-tract samples in many tracts; and an underlying correlation between the tract population size and tract median income. In the absence of any one of these conditions, the bias-corrected estimators perform very well. In many metropolitan areas, however, all three conditions hold for black and Hispanic populations. As a result, the bias-corrected estimators are of little use in such places. More generally, researchers are cautioned that when the population of interest is unevenly distributed among geographic units and samples are small in many units, the bias-corrected rank-order estimators may fail to provide accurate estimates.

Table B1. Proportion of bias explained in rank-order H and R, by sampling rate, mean tract size, and race, 380 metropolitan areas, 100 replications.

Tract Mean Size	4% Sampling Rate		8% Sampling Rate		16% Sampling Rate	
	H^R	R^R	H^R	R^R	H^R	R^R
			All Families			
50	0.95	0.98	0.97	1.02	0.98	1.04
100	0.96	1.01	0.97	1.03	0.98	1.04
250	0.98	1.04	0.99	1.04	1.01	1.04
500	1.00	1.04	1.01	1.04	1.03	1.05
			Black Families			
25	3.43	1.10	5.55	0.84	12.33	0.01
50	2.64	1.56	3.48	1.12	3.88	1.66
100	1.76	1.46	1.90	1.42	1.51	1.57
150	1.51	1.33	1.64	1.45	1.67	1.34
200	1.39	1.28	1.47	1.30	1.49	1.28
250	1.34	1.25	1.38	1.26	1.38	1.23
300	1.26	1.20	1.30	1.20	1.30	1.18
400	1.16	1.13	1.19	1.13	1.22	1.12
500	1.18	1.10	1.14	1.10	1.17	1.11
			Hispanic Families			
25	3.70	0.46	3.88	1.27	3.68	1.31
50	2.33	1.41	2.60	1.51	2.88	1.43
100	1.82	1.35	2.01	1.39	2.14	1.34
150	1.53	1.29	1.64	1.31	1.71	1.28
200	1.38	1.23	1.45	1.24	1.51	1.22
250	1.26	1.18	1.31	1.18	1.36	1.18
300	1.18	1.13	1.21	1.14	1.26	1.14
400	1.07	1.08	1.08	1.07	1.11	1.09
500	1.01	1.04	1.00	1.04	1.03	1.06

Notes: Proportion of bias explained in rank-order H and R is calculated as: [(bias in uncorrected H/R - bias in corrected H/R) / bias in uncorrected H/R]. Values greater than 1 represent an overcorrection in rank-order H/R. The proportion of bias explained values presented under "All Families" are derived from simulations in which we rescale each tract size (by applying a constant multiplier to each income category) within a metropolitan area so that the mean of all tract sizes within a metropolitan area statisfies a target value, where the targets are 50, 100, 250, and 500. This rescaling is done to all 380 metropolitan areas. The simulation is then run on this set of adjusted metropolitan areas for 100 replications. The bias in both the uncorrected and corrected rank-order measures is calculated by first averaging the bias within metropolitan areas across replications, and then averaging across all metropolitan areas for a given sampling rate. Data for these simulations come from the reported 2005-2009 ACS for all families. The values presented in the "Black" and "Hispanic Families" sections are derived by interpolating the remaining bias in corrected and uncorrected rank-order H/R from smoothed lines capturing the relationship between bias and mean tract size; see figure 4 for an example of one of these smooth lines. The data for these simulations come from the reported 2005-2009 ACS for all black and Hispanic families, respectively.

Figure B1. Bias in Uncorrected and Corrected Binary H and R at 8% Sampling Rate, 380 Metropolitan Areas, by Income Percentile, All Tracts Adjusted to Have a Total Population of 1000

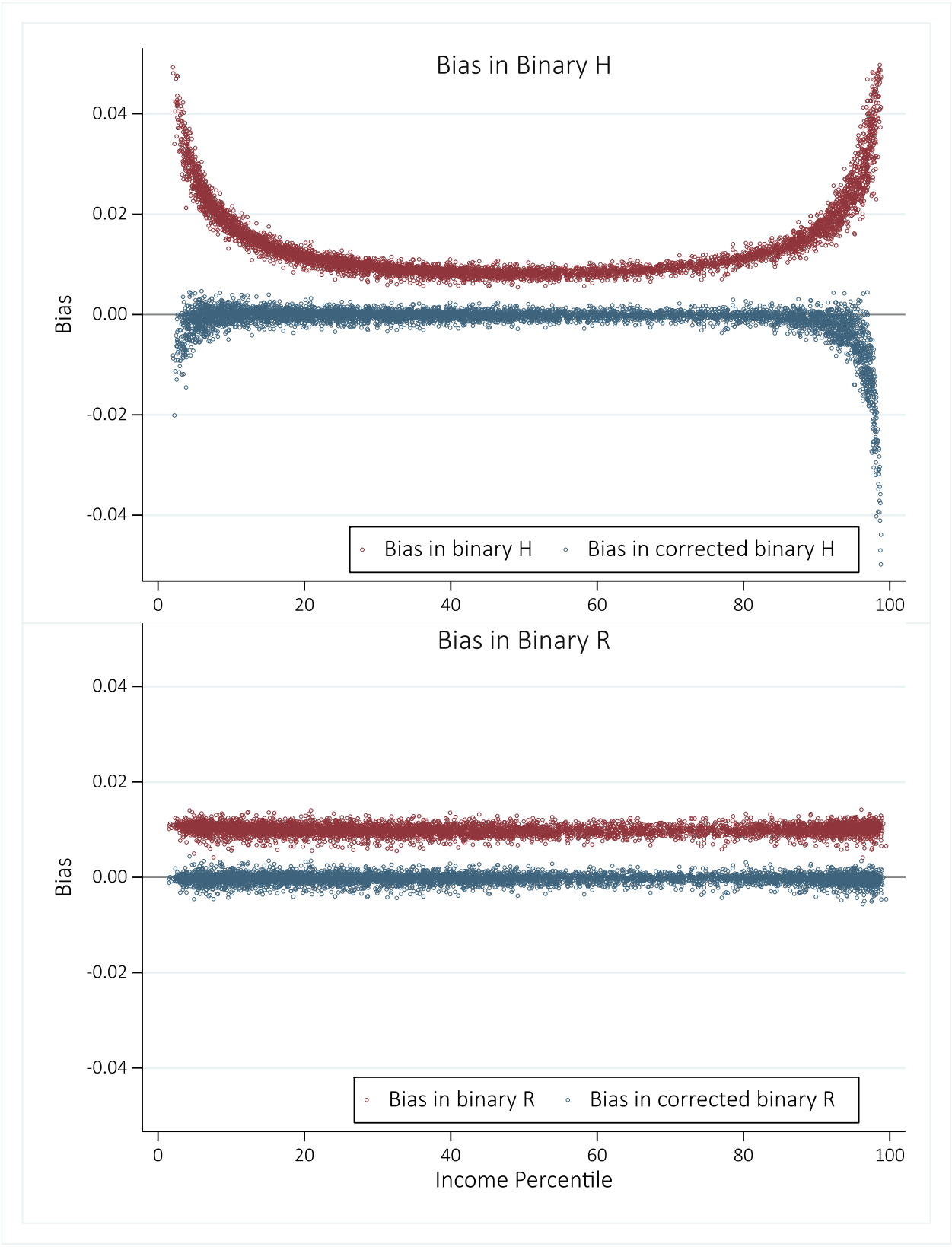
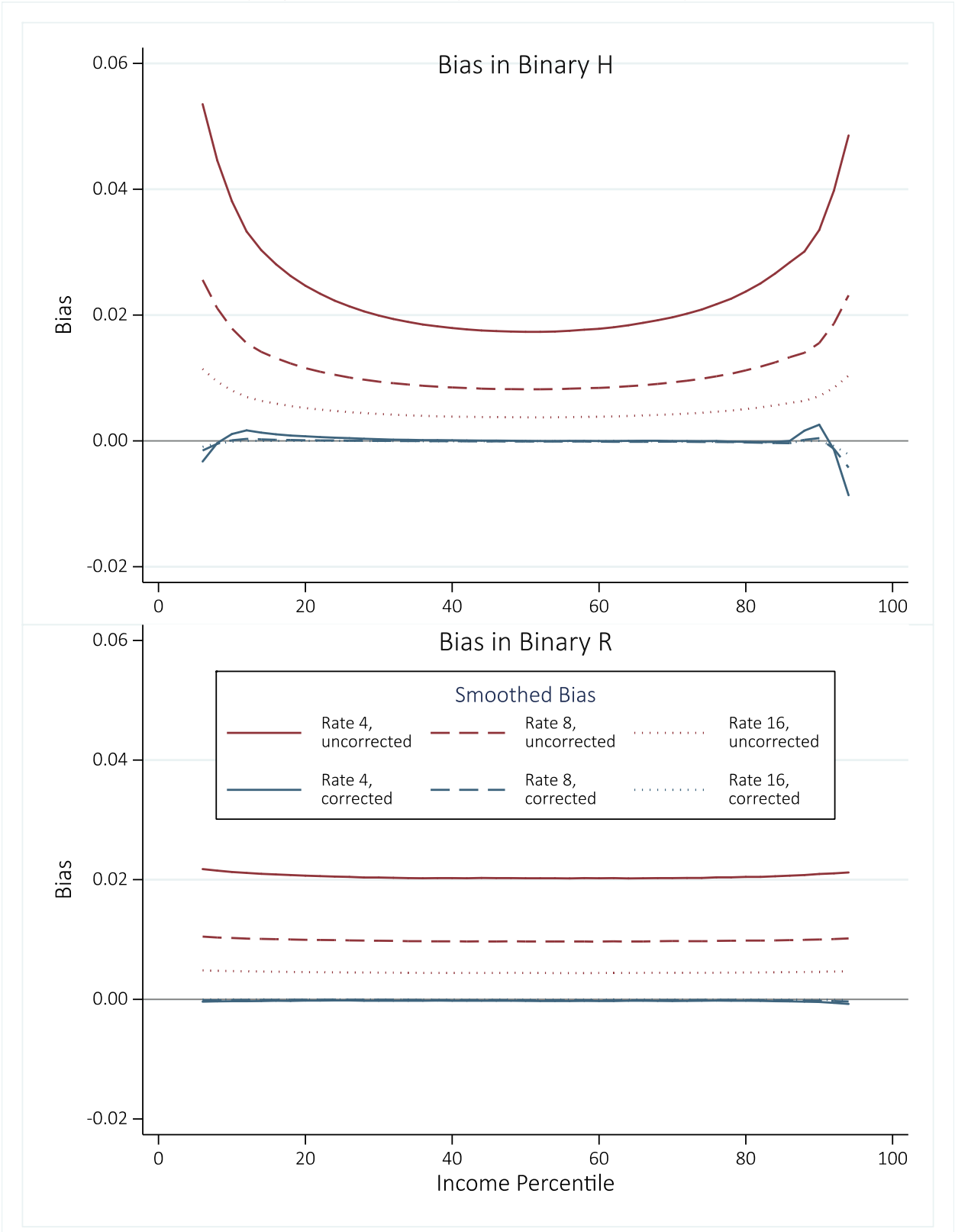


Figure B2. Average Bias in Uncorrected and Corrected Binary H and R, 380 Metropolitan Areas, by Income Percentile and Sampling Rate, All Tracts Adjusted to Have a Total Population of 1000



Appendix C: Estimates of Income Segregation and Replication of Previously Published Results

Appendix C provides additional income segregation estimates and replication of published results using the bias-corrected estimators. Appendix Table C1 provides uncorrected and bias-corrected estimates of binary family income segregation (*H* and *R*) at the 10th, 50th, and 90th percentiles of the income distribution from 1970 to 2014. (Figure 7 presents bias-corrected *H10, H50,* and *H90*.)

Appendix Table C2 presents estimates of rank-order income segregation among families by racial/ethnic group. Bias-corrected estimates of R^R from the top and bottom three panels are presented in Appendix Figure C1. The second and third panels provide additional results for white families, presenting average income segregation only in the metropolitan areas that meet the sample criteria (over 500,000 residents as of 2007, at least 10,000 families of the relevant racial/ethnic group from 1970-2000, and average tract size of at least 200 families of the relevant racial group) for black and Hispanic families, respectively.[22] Comparing the second and fifth panels, for example, confirms that income segregation among white families increased less than among black families in the same set of 22 metropolitan areas.

Finally, Tables C3-C5 replicate multivariate regression tables from previously published papers estimating the relationship between income inequality and rank-order income segregation. The key coefficients from these tables are presented in Table 3.

[22] The 22 metropolitan areas in the black family sample are: Atlanta-Sandy Springs-Marietta, Augusta-Richmond County, Baltimore-Towson, Baton Rouge, Birmingham-Hoover, Charleston-North Charleston, Charlotte-Gastonia-Concord, Columbia, Detroit-Livonia-Dearborn, Fort Lauderdale-Pompano Beach-Deerfield, Greensboro-High Point, Houston-Baytown-Sugar Land, Jackson, Jacksonville, Little Rock-North Little Rock, Memphis, New Orleans-Metairie-Kenner, Newark-Union, Raleigh-Cary, Richmond, Virginia Beach-Norfolk-Newport News, and Washington, DC - Arlington-Alexandria. The 20 metropolitan areas in the Hispanic family sample are: Albuquerque, Austin-Round Rock, Bakersfield, Dallas-Plano-Irving, El Paso, Fresno, Houston-Baytown-Sugar Land, Los Angeles-Long Beach-Glendale, McAllen-Edinburg-Pharr, Miami-Miami Beach-Kendall, New York-Wayne-White Plains, Oxnard-Thousand Oaks-Ventura, Phoenix-Mesa-Scottsdale, Riverside-San Bernardino-Ontario, San Antonio, San Diego-Carlsbad-San Marcos, San Jose-Sunnyvale-Santa Clara, Santa Ana-Anaheim-Irvine, Stockton, and Tucson.

Table C1. Mean Uncorrected and Bias-Corrected Estimates of Binary Family Income Segregation at the 10th, 50th, and 90th Income Percentiles, 116 Largest Metropolitan Areas, 1970 to 2014

	Uncorrected *H10*	Bias-Corrected *H10*	Uncorrected *H50*	Bias-Corrected *H50*	Uncorrected *H90*	Bias-Corrected *H90*
1970	0.077	0.071	0.092	0.089	0.140	0.132
	(0.024)	(0.024)	(0.028)	(0.028)	(0.038)	(0.039)
1980	0.090 ***	0.079 ***	0.089 ^	0.085 **	0.123 ***	0.113 ***
	(0.030)	(0.031)	(0.027)	(0.026)	(0.038)	(0.038)
1990	0.121 ***	0.112 ***	0.106 ***	0.102 ***	0.159 ***	0.149 ***
	(0.039)	(0.039)	(0.028)	(0.028)	(0.041)	(0.041)
2000	0.113 ***	0.104 ***	0.109	0.105	0.153 **	0.143 **
	(0.031)	(0.031)	(0.027)	(0.027)	(0.037)	(0.037)
2007	0.126 ***	0.107 ^	0.116 ***	0.108 ^	0.167 ***	0.148 **
	(0.032)	(0.030)	(0.028)	(0.028)	(0.039)	(0.038)
2014	0.124 ***	0.106	0.118 ***	0.110 **	0.164 ***	0.146
	(0.029)	(0.029)	(0.028)	(0.028)	(0.038)	(0.037)
	Uncorrected *R10*	Bias-Corrected *R10*	Uncorrected *R50*	Bias-Corrected *R50*	Uncorrected *R90*	Bias-Corrected *R90*
1970	0.057	0.054	0.120	0.116	0.105	0.101
	(0.020)	(0.020)	(0.035)	(0.035)	(0.034)	(0.034)
1980	0.069 ***	0.064 ***	0.115 *	0.110 **	0.093 ***	0.087 ***
	(0.026)	(0.027)	(0.034)	(0.033)	(0.032)	(0.032)
1990	0.095 ***	0.090 ***	0.137 ***	0.132 ***	0.117 ***	0.112 ***
	(0.036)	(0.036)	(0.035)	(0.035)	(0.036)	(0.036)
2000	0.086 ***	0.081 ***	0.141 *	0.136 ^	0.114	0.109
	(0.028)	(0.028)	(0.034)	(0.034)	(0.033)	(0.033)
2007	0.095 ***	0.085 *	0.149 ***	0.139	0.118 *	0.108
	(0.028)	(0.028)	(0.035)	(0.035)	(0.034)	(0.033)
2014	0.092 ***	0.082	0.152 ***	0.142 **	0.116	0.106 ^
	(0.025)	(0.025)	(0.035)	(0.035)	(0.033)	(0.033)

Notes: Cells report means with standard deviations beneath. Sample is 116 metropolitan areas with over 500,000 residents as of 2007 (Cape Coral is excluded due to missing data in 1970). 2007 refers to 2005-09 ACS and 2014 refers to 2012-16 ACS. ^p≤0.10; ***p≤0.05; **p≤0.01; ***p≤0.001; statistical significance tests come from regression models with metropolitan area fixed effects that compare the estimate in each decade to the prior decade. Note that 2007 and 2014 are both compared to the 2000 estimate.

Table C2. Mean Uncorrected and Bias-Corrected Estimates of Rank-Order Family Income Segregation, by Race/Ethnicity, 2000-2014

Year	Uncorrected H^R	Bias-Corrected H^R	Uncorrected R^R	Bias-Corrected R^R
		White Families		
2000	0.102	0.095	0.114	0.107
	(0.025)	(0.024)	(0.028)	(0.028)
2007	0.112 ***	0.097 ***	0.123 ***	0.109 *
	(0.028)	(0.027)	(0.031)	(0.030)
2014	0.114 ***	0.099 ***	0.124 ***	0.111 ***
	(0.028)	(0.026)	(0.030)	(0.030)
		White Families (Black Metro Sample)		
2000	0.105	0.097	0.117	0.110
	(0.021)	(0.021)	(0.023)	(0.023)
2007	0.114 ***	0.097	0.125 ***	0.110
	(0.026)	(0.025)	(0.029)	(0.029)
2014	0.119 ***	0.101 *	0.130 ***	0.114 ^
	(0.025)	(0.025)	(0.028)	(0.028)
		White Families (Hispanic Metro Sample)		
2000	0.127	0.119	0.143	0.135
	(0.024)	(0.024)	(0.027)	(0.026)
2007	0.143 ***	0.127 **	0.157 ***	0.143 **
	(0.029)	(0.027)	(0.032)	(0.030)
2014	0.141 ***	0.126 *	0.155 ***	0.142 *
	(0.030)	(0.028)	(0.032)	(0.031)
		Non-Hispanic White Families		
2000	0.097	0.089	0.108	0.101
	(0.022)	(0.021)	(0.025)	(0.024)
2007	0.103 ***	0.086 ***	0.112 ***	0.097 ***
	(0.023)	(0.020)	(0.025)	(0.023)
2014	0.103 ***	0.085 ***	0.111 ***	0.095 ***
	(0.023)	(0.020)	(0.025)	(0.023)
		Black Families		
2000	0.113	0.093	0.123	0.105
	(0.025)	(0.023)	(0.028)	(0.027)
2007	0.152 ***	0.112 ***	0.160 ***	0.124 ***
	(0.028)	(0.025)	(0.030)	(0.029)
2014	0.159 ***	0.118 ***	0.167 ***	0.130 ***
	(0.027)	(0.024)	(0.029)	(0.027)
		Hispanic Families		
2000	0.104	0.084	0.113	0.095
	(0.019)	(0.014)	(0.020)	(0.015)
2007	0.135 ***	0.103 ***	0.143 ***	0.114 ***
	(0.029)	(0.019)	(0.029)	(0.021)
2014	0.137 ***	0.107 ***	0.146 ***	0.119 ***
	(0.026)	(0.018)	(0.026)	(0.020)

Notes: Cells report means with standard deviations beneath. Sample is metropolitan areas with over 500,000 residents as of 2007 (Cape Coral is excluded due to missing data in 1970), at least 10,000 families of the relevant racial/ethnic group from 1970-2000, and average tract size of at least 200 families of the relevant racial group. N=116 metropolitan areas for white families, 113 metropolitan areas for non-Hispanic white families, 22 for black families, and 20 for Hispanic families. "Black Metro Sample" and "Hispanic Metro Sample" are estimates for white families in the samples of metros where black or Hispanic populations are sufficient for estimation (N=22 and 20, respectively). 2007 refers to 2005-09 ACS and 2014 refers to 2012-16 ACS. ^p≤0.10; *p≤0.05; **p≤0.01; ***p≤0.001; statistical significance tests come from regression models with metropolitan area fixed effects and compare both 2007 and 2014 to 2000.

Table C3. Estimated Effects of Income Inequality on Rank-Order Income Segregation, 1970 to 2000 (Replication of Reardon and Bischoff, 2011)

	Published Estimates (using published H^R)	Estimates Using Bias-Corrected H^R	Estimates Using Bias-Corrected R^R
Gini	0.561 ***	0.480 ***	0.528 ***
	(0.085)	(0.124)	(0.142)
Year = 1980	0.027 ***	0.020 **	0.024 ***
	(0.007)	(0.008)	(0.009)
Year = 1990	0.025 *	0.031 ***	0.034 **
	(0.012)	(0.012)	(0.013)
Year = 2000	0.012	0.019	0.021
	(0.016)	(0.016)	(0.017)
Adjusted R^2	0.959	0.925	0.924
N	400	400	400

Notes: Models include metropolitan area fixed effects, year indicators, and metropolitan-year covariates (metro population, unemployment rate, proportion under age 18, proportion over age 65, proportion with high school diploma, proportion foreign born, proportion female headed families, per capita income, proportions employed in manufacturing, construction, financial and real estate, professional and managerial jobs, and proportions of housing built within ten, five, and one years). Bootstrapped standard errors in parentheses. Reported results published in Reardon & Bischoff 2011, Table 4, "All Families" model, predicting H^R. Sample includes 100 largest metropolitan areas in 2000; data from 1970, 1980, 1990, and 2000. ^p≤0.10; *p≤0.05; **p≤0.01; ***p≤0.001.

Table C4. Estimated Effects of Income Inequality on Rank-Order Income Segregation, 1970 to 2009 (Replication of Bischoff and Reardon, 2014)

	Published Estimates (using published H^R)	Estimates Using Bias-Corrected H^R	Estimates Using Bias-Corrected R^R
Gini	0.443 ***	0.448 ***	0.500 ***
	(0.090)	(0.083)	(0.098)
Average change, 1970s	0.000	0.001	0.002
	(0.008)	(0.007)	(0.006)
Average change, 1980s	0.006	0.008 *	0.008
	(0.005)	(0.004)	(0.004)
Average change, 1990s	-0.018 ***	-0.018 ***	-0.018 ***
	(0.004)	(0.003)	(0.014)
Average change, 2000s	0.006	0.000	-0.001
	(0.005)	(0.004)	(0.004)
Adjusted R^2	0.905	0.900	0.897
N	584	584	584

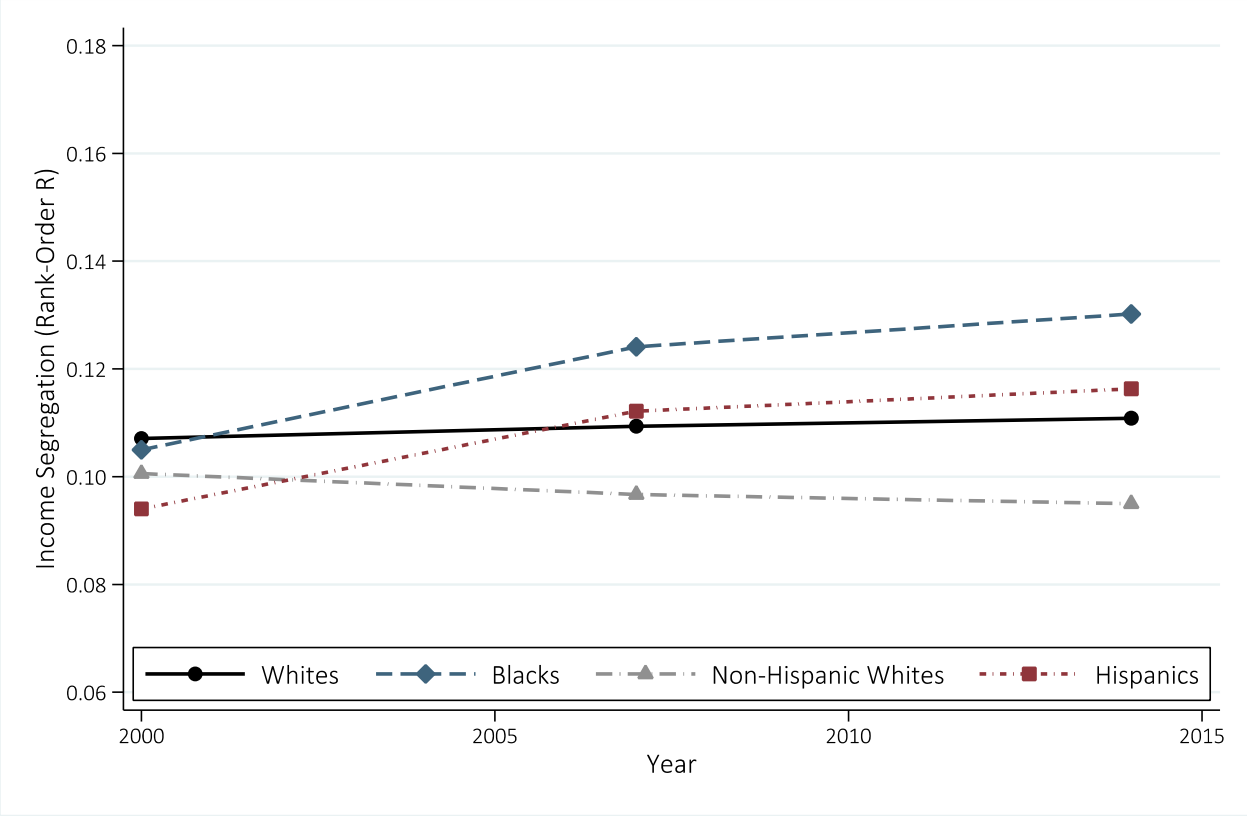
Notes: Models include metropolitan area fixed effects and metropolitan-year covariates (log population, age composition, residents' educational attainment, unemployment rate, proportion employed in manufacturing, per capita income, racial composition, foreign-born composition, and female-headed family rate). Reported results published in Bischoff and Reardon 2014, Table 5, predicting H^R. Sample includes 117 metropolitan areas with population >500,000 in 2007 (minus Cape Coral in 1970); data from 1970, 1980, 1990, 2000, and 2007-11. ^p≤0.10; *p≤0.05; **p≤0.01; ***p≤0.001.

Table C5. Longitudinal Regression Predicting Rank-Order Income Segregation Among Households With and Without Children, 1990 to 2010 (Replication of Owens 2016)

	Published Estimates (using published H^R)	Estimates Using Bias-Corrected H^R	Estimates Using Bias-Corrected R^R
Income Inequality	0.232 ***	0.194 **	0.230 **
	(0.069)	(0.075)	(0.081)
Income Inequality x Families with Children	0.223 ***	0.260 ***	0.275 ***
	(0.057)	(0.061)	(0.066)
Fragmentation x Families with Children	0.028 **	0.031 **	0.034 **
	(0.010)	(0.011)	(0.012)
District Fragmentation x 2000	-0.001	-0.003	-0.003
	(0.009)	(0.010)	(0.011)
District Fragmentation x 2010	-0.012	-0.011	-0.012
	(0.009)	(0.010)	(0.011)
Families with Children x Fragmentation x 2000	0.014	0.016	0.019
	(0.013)	(0.014)	(0.015)
Families with Children x Fragmentation x 2010	0.023 ^	0.020	0.023
	(0.013)	(0.014)	(0.015)
Families with Children	-0.153 ***	-0.188 ***	-0.193 ***
	(0.029)	(0.031)	(0.034)
Year = 2000	-0.012 ***	-0.018 ***	-0.020 ***
	(0.002)	(0.002)	(0.003)
Year = 2010	-0.012 **	-0.014 ***	-0.018 ***
	(0.004)	(0.004)	(0.004)
Families with Children x 2000	0.003	0.003	0.005
	(0.003)	(0.003)	(0.003)
Families with Children x 2010	0.021 ***	0.023 ***	0.022 ***
	(0.005)	(0.005)	(0.005)
N	570	570	570

Notes: Models include metropolitan area fixed effects and group-metro-year covariates and their interaction with group (log population, age composition, female-headed household rate, racial composition, racial segregation, foreign born composition, residents' educational attainment, unemployment rate, proportion employed in manufacturing, and private school enrollment rate). Reported results published in Owens 2016, Table 4, Model 2, predicting H^R. Sample includes 95 largest metropolitan areas in 2010 with more than 1 school district; data from 1990, 2000, and 2008-2012. ^p≤0.10; *p≤0.05; **p≤0.01; ***p≤0.001.

Figure C1. Mean Bias-Corrected Estimates of Income Segregation (Rank-Order *R*) among Families by Race, 2000 to 2014



Note: Data are from Appendix Table C2 (panels 1,3, 4, 5), which reports statistical significance of changes over time. Sample includes metropolitan areas among the 116 largest in 2007 where there are at least 10,000 families of the relevant racial/ethnic group from 1970-2000, and the average tract population is at least 200 of the relevant group (N=116 for white families; 113 for non-Hispanic white families; 22 for black families; 20 for Hispanic families).