Stanford cepa
Center for Education Policy Analysis

CEPA Working Paper No. 22-01

It is Surprisingly Difficult to Measure Income Segregation

AUTHORS

Josh Leung-Gagne
Stanford University

sean f. reardon
Stanford University

ABSTRACT

Recent studies have shown that Census- and ACS-based estimates of income segregation are subject to upward finite sampling bias (Logan et al. 2018, 2020; Reardon et al. 2018). We identify two additional sources of bias that are larger and opposite in sign to finite sampling bias: measurement error-induced attenuation bias and temporal pooling bias. The combination of these three sources of bias render estimates of the trend in income segregation unclear. We formalize the three types of bias, providing a method to correct them simultaneously using public Census and ACS data from 1990 to 2015-2019. We use these methods to produce bias-corrected estimates of income segregation in the U.S. from 1990 to 2019. We find that (1) segregation is on the order of 50 percent greater than previously believed; (2) the increase from 2000 to the 2005-09 period was much greater than indicated by previous estimates; and (3) segregation has declined since 2005-09. Correcting these biases requires good estimates of the reliability of self-reported income and of the year-to-year volatility in neighborhood mean incomes.

VERSION

May 2022

Suggested citation: Leung-Gagne, J., & Reardon, S.F. (2022). It is Surprisingly Difficult to Measure Income Segregation.. (CEPA Working Paper No. 22-01). Retrieved from Stanford Center for Education Policy Analysis: https://cepa.stanford.edu/wp22-01

Title Page

Title: It is Surprisingly Difficult to Measure Income Segregation

Authors: Josh Leung-Gagné (Stanford University, jgagne@stanford.edu) & sean f. reardon (sean.reardon@stanford.edu)

Suggested Running Head: It is Surprisingly Difficult to Measure Income Segregation

Word Count: 7,483

Acknowledgement: We thank Ann Owens for providing the necessary code to match our sample to past studies. We received helpful comments from discussants and audience members at Population Association of America and American Education Research Association conferences. Leung-Gagné acknowledges support from the Institute for Education Sciences (R305B140009).

It is Surprisingly Difficult to Measure Income Segregation

Abstract

Recent studies have shown that Census- and ACS-based estimates of income segregation are subject to upward finite sampling bias (Logan et al. 2018, 2020; Reardon et al. 2018). We identify two additional sources of bias that are larger and opposite in sign to finite sampling bias: measurement error-induced attenuation bias and temporal pooling bias. The combination of these three sources of bias render estimates of the trend in income segregation unclear. We formalize the three types of bias, providing a method to correct them simultaneously using public Census and ACS data from 1990 to 2015-2019. We use these methods to produce bias-corrected estimates of income segregation in the U.S. from 1990 to 2019. We find that (1) segregation is on the order of 50 percent greater than previously believed; (2) the increase from 2000 to the 2005-09 period was much greater than indicated by previous estimates; and (3) segregation has declined since 2005-09. Correcting these biases requires good estimates of the reliability of self-reported income and of the year-to-year volatility in neighborhood mean incomes.

KEYWORDS: Income Segregation, Residential Segregation, Measurement, Pooled Sampling, Attenuation, Finite Sampling Bias

Neighborhood Income Segregation in the US

A growing scholarship seeks to offer a nuanced description of patterns of income segregation across income level, over time, and with respect to other types of segregation (Jargowsky 1996; Massey, Rothwell, and Domina 2009; Owens 2016; Owens, Reardon, and Jencks 2016; Reardon and Bischoff 2011). These descriptions of income segregation – the uneven distribution of people with different incomes across space, place, and institutions – are important given how much people's lives and futures are shaped by the socioeconomic conditions of their surroundings (Chetty et al. 2014; Chetty, Hendren, and Katz 2016; Chetty and Hendren 2018a, 2018b; Sampson 2012).

Sociological theories suggest income segregation exacerbates inequality and hampers class mobility across generations. Since Park (1915), sociologists have theorized that distinct social norms and environments emerge from clustering people into neighborhoods. More recently, collective efficacy theory explains how poor neighborhoods' organizational characteristics and reputations – promoted by higher mobility and biased perceptions, respectively – increase residents' exposure to crime (Sampson 2012; Sampson, Raudenbush, and Earls 1997). That is, the different lived experiences of the rich and the poor are not confined to the household; rather, they are exacerbated by contextual disparities that emerge from residential segregation.

A similar conclusion is reached by a different line of scholarship that emphasizes the political economy of place (Lichter, Parisi, and Taquino 2012; Logan and Molotch 1987; Massey et al. 2009). These scholars describe how elites capture municipalities by maneuvering both politically and residentially among cities, suburbs, and towns. They argue that the growth imperative of municipalities binds them to elite interests as they compete for growth and affluent tax bases. Income segregation within places enables municipalities to boost their elite clients by investing more in richer neighborhoods, while income segregation between places concentrates resources and needs into distinct political units. This is potentially self-reinforcing as the higher tax bases, social capital, and human capital of rich neighborhoods and municipalities enables those places to attract high-income residents through investments in community resources and amenities like schools, social services, and parks.

One theorized consequence of income segregation, then, is differences in educational opportunity that enable the rich to transfer their class and status across generations. Empirical studies have validated this concern, showing that neighborhood socioeconomic conditions strongly influence child development, economic mobility, and educational attainment (Chetty and Hendren 2018a, 2018b; Chetty et al. 2016; Wodtke, Elwert, and Harding 2016). A complementary scholarship focuses on the role of income segregation between schools and school districts, with recent studies highlighting the strong relationship between income segregation and both income and race/ethnic achievement gaps (Fahle et al. 2020; Fahle and Reardon 2018; Owens 2016, 2017; Owens et al. 2016; Reardon, Weathers, et al. 2019; Reardon, Kalogrides, and Shores 2019).

Scholars have attended to these concerns by describing the income segregation of residences and schools in the United States. This literature shows that residential income segregation has increased over the last four decades, particularly in the 1980s and since 2000 (Bischoff and Reardon 2014; Jargowsky 1996; Massey et al. 2009; Owens 2016; Reardon et al. 2018; Reardon and Bischoff 2011). This increase was driven by increasing segregation of poverty in the 1970s and 2000s as well as increasing segregation at all income levels in the 1980s (Bischoff and Reardon 2014; Reardon and Bischoff 2011). The increase in segregation is also concentrated among households with school-aged children (Owens 2016). Partly as a result, income segregation between schools and school districts has increased markedly since 1990 (Owens et al. 2016), with affluent white families increasingly living in affluent school districts (Owens 2017).

Finite Sampling Bias

This story has recently been questioned. Finite sampling bias partly accounts for the apparent increase in residential income segregation during the 2000s, when the Census Bureau switched from collecting income data in the Census to collecting it in the American Community Survey (ACS) (Logan et al. 2018, 2020; Reardon et al. 2018). One of the more recent, prominent studies showing increasing income segregation was by Bischoff and Reardon (2014). It compared residential income segregation in 2000 measured using the Census to segregation measured using the 2005-2009 through 2007-2011 ACS 5-year pools. Because the effective sample of the latter is much smaller (roughly 8% vs. roughly 17% in the Census), it is subject to more sampling variation, which upwardly biases income segregation estimates (Logan et al. 2018).

Both Logan, Foster, Ke, and Li (2018) and Reardon, Bischoff, Owens, and Townsend (2018) demonstrated that the apparent increase in income segregation between 2000 and 2005-2009 is exaggerated by this bias. Reardon, Bischoff, Owens, and Townsend (2018) proposed a correction for finite sampling bias in segregation measures using publicly available data while Logan, Foster, Xu, and Zhang (2020) proposed an alternative approach using restricted microdata and new methods that account for weighted sampling. Replicating Bischoff and Reardon (2014), the former finds that income segregation increased by about half of what was initially reported while the latter finds that income segregation was stable over the period in question.

The recent attention to finite sampling bias is important, but misses—as we show below—two additional sources of downward bias that are much larger than the upward bias that results from finite samples. We formalize the three types of bias – finite sampling bias, attenuation bias, and pooling bias – and provide a method for correcting them using public Census and ACS data. Finally, we report new estimates of the national trend in residential income segregation, finding that it is substantially different than previously believed.

Measurement Error-Induced Attenuation Bias

Estimates of income segregation typically use error-prone self-reported measures of income. In addition to oft-cited sources of error like nonresponse and motivated misreporting, error can also come from the cognitive demands of income reporting in the form of misunderstanding income concepts and terms, information retrieval errors, and motivated misremembering (Moore, Stinson, and Welniak 2000).

Measurement error in self-reported income presents an issue for estimating segregation because it increases the apparent variance of income within neighborhoods, exaggerating the extent to which income distributions of different neighborhoods overlap. In other words, it makes neighborhoods appear more similar than they are, downwardly biasing estimates of income segregation.

Attenuation in segregation estimates has received only passing attention in the literature. In a comment, Dickens and Levy (2003) describe downward bias from response error and income volatility in measures of segregation by permanent income, reporting that estimates using dissimilarity indices should be inflated by 15 to 30 percent. In a note, Owens, Reardon, and Jencks (2016) disattenuate their income segregation estimates by dividing H by the reliability of self-reported annual income. Though attenuation bias appears to be severe, it remains unformalized, the previously applied corrections have not been evaluated, and reported estimates rarely consider it.

Temporal Pooling Bias

A third source of bias emerges when samples of respondents reporting annual income are pooled across multiple years (as is done in the ACS). In this case, temporal variation in neighborhood mean income is included in calculations of within-neighborhood income variation, artificially inflating the variance of incomes within neighborhoods. This biases measures of income segregation downward. This source of bias has not been discussed in the literature.

The ACS uses pooled samples (pooled over 5 years, in the case of tract-level data), but the decennial Census does not. As a result, pooling bias affects the trend during the 2000 to 2005-2009 period that has been the focus of recent work. If neighborhood means vary over five-year sampling periods, income segregation increased more than previously believed during the 2000s. Furthermore, the bias may affect trends even when using the ACS only; if the extent to which neighborhood means vary has been changing since the switch to the ACS, the estimated trend from 2005-2009 onward is biased in an unclear direction.

Formalizing the Three Biases

Simplifying Assumptions

For simplicity, and to build intuition, we assume the distribution of observed (self-reported) income in log-dollars each neighborhood j and year t is described by the data generating model:

$$\hat{Y}_{ijt} = Y_{ijt} + \epsilon_{ijt} = \Theta + \delta_t + \mu_j + v_{jt} + e_{ijt} + \epsilon_{ijt}, \qquad \begin{bmatrix} \delta_t \\ \mu_j \\ v_{jt} \\ e_{ijt} \\ \epsilon_{ijt} \end{bmatrix} \sim \left[\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \Delta & 0 & 0 & 0 & 0 \\ 0 & \tau & 0 & 0 & 0 \\ 0 & 0 & \sigma & 0 & 0 \\ 0 & 0 & 0 & \omega & 0 \\ 0 & 0 & 0 & 0 & \eta \end{pmatrix} \right] \tag{1}$$

Here Y_{ijt} indicates the true income of household i in neighborhood j in year t; $\hat{Y}_{ijt}$ indicates its self-reported income; Θ is the average value of Y in the population over time; δ_t and μ_j are year- and neighborhood-specific deviations from this average; v_{jt} is a neighborhood-by-year specific deviation; e_{ijt} is the difference between a household's true income and the average income in their neighborhood and year; and ϵ_{ijt} is measurement error in observed household income.

Additionally, we assume that all neighborhoods are the same size (so that they each are weighted equally in computing segregation). We consider the case where all neighborhoods in the relevant region are included in the data collection, but a fixed sample of size n is drawn from each neighborhood. As a result, we have the full population of neighborhoods, but a finite sample of households within each.

Because all neighborhoods are included in estimation, we treat both the number of neighborhoods and the total population over all neighborhoods as infinite in the derivations that follow.

Given the data generating model, the true within-neighborhood variance of income (denoted W) in a given year is

$$W = \omega; \tag{2}$$

the true between-neighborhood variance of income (denoted B) in a given year is

$$B = \tau + \sigma; \tag{3}$$

and the total population variance of income (denoted V) in a given year is

$$\begin{aligned} V &= B + W \\ &= \tau + \sigma + \omega. \end{aligned} \tag{4}$$

Segregation is generally defined as the proportion of variation in income that is due to between-neighborhood differences in income. A simple way to operationalize this is to define segregation in year t as the ratio of between-neighborhood variance to total variance of income:

$$S = \frac{B}{V} = 1 - \frac{W}{V}. \tag{5}$$

The latter formula is often used when it is simpler to estimate W than B.

Bias in Segregation Estimates

We start by considering the most general case, when all three features of data collection described above are present: individual income is measured with error; neighborhood income distributions are estimated based on finite samples; and neighborhood income distributions are estimated by pooling data across several years. Below we adopt the following notation: to indicate that a measure of segregation is based on data that include measurement error, we add a superscript accent symbol (S'); to indicate that a measure is based on a sample, we add a superscript asterisk (S^*); and to indicate that a measure is based on data pooled over years, we add a tilde above ($\tilde{S}$). In addition, let T be the number of years over which neighborhood samples are pooled; let n be the sample size in each neighborhood. Note that in our stylized data generating model, the reliability of self-reported income would be equal to $r = V/(V + \eta)$.

We first consider the within-neighborhood variance we expect to observe in this case. The within-year, within-neighborhood variance of observed income will be $\omega + \eta$ (true income variance plus measurement error). But we do not observe a sample drawn from one year; nor do we observe income of all households. Instead, we estimate the within-neighborhood variance from a finite sample of size n pooled over T years. If we demean the sample distributions in each year, pool them over years, and then compute the sample variance of the pooled distribution, the expected variance of the observed demeaned within-neighborhood distribution will be $\left(\frac{n-1}{n}\right)(\omega + \eta)$. However, because the sample is pooled over years, the observed within-neighborhood sample variance will be larger, because neighborhoods do not have stable mean income over time. Over a period of T years, the neighborhood means have an expected sample variance of $\Delta^T + \left(\frac{T-1}{T}\right)\sigma$ where Δ^T represents the expected sample variance of the population mean income over T years and $\left(\frac{T-1}{T}\right)\sigma$ represents the expected variance of the neighborhood means over T years, net of the overall changes in income over the period.[1]

Thus, the expected variance of observed income within a neighborhood, when the distribution of income is estimated from a sample of size n surveyed over a period of T years will be

[1] We think of the T years as a random draw of years from some long time period such that the expected value of Δ^T is $\left(\frac{T-1}{T}\right)\Delta$. Note that while any specific period of T consecutive years might have substantially different variance than the long-run average variance of a set of random years, in practice we estimate Δ^T from computations of Δ over periods only slightly longer or shorter than T, depending on data availability.

$$E\left[\widehat{W}'^*\right] = \left(\frac{n-1}{n}\right)(\omega+\eta) + \left(\frac{T-1}{T}\right)\sigma + \Delta^T. \tag{6}$$

Next, we consider the total variance of observed income in a sample pooled over T years. This will be equal to V, plus two additional components: additional variance η due to measurement error in observed income; and additional variance Δ^T resulting from changes in mean population income over time:

$$\begin{aligned}\hat{V}'^* &= \Delta^T + \tau + \sigma + \omega + \eta \\ &= \Delta^T + \frac{1}{r}V\end{aligned} \tag{7}$$

Subtracting the expected within-neighborhood sample variance from the total variance yields an expression for the expected value of the between-neighborhood variance in observed income:

$$E\left[\hat{B}'^*\right] = \tau + \frac{1}{T}\sigma + \left(\frac{1}{n}\right)(\omega+\eta). \tag{8}$$

The first term represents between-neighborhood variance due to stable differences in neighborhood incomes; the second term represents variance among neighborhoods in their average v_{jt}; and the third term represents variance in the estimated neighborhood means due to the fact that the neighborhood means are estimated from finite samples of size n.

After some algebraic rearranging of terms, the expected value of S can be written as

$$E\left[\hat{S}'^*\right] = r\left[S\left(\frac{n-1}{n}\right) + \frac{1}{rn} - \left(\frac{T-1}{T}\right)\frac{\sigma}{V}\right]\left(\frac{V}{V+r\Delta^T}\right). \tag{9}$$

Note that, because the sample sizes used for computing segregation are generally quite large, we will drop the expected value notation going forward.

Given Equations (6)-(9), it is straightforward to derive expressions for the bias in W, B, V, and S under any combination of the three data collection processes. If there is no measurement error, we set $r = 1$ in the above expressions. If income data are collected from the full population, rather than samples, we set $n = \infty$. If data are collected in a single year, rather than pooled over years, $T = 1$ and $\Delta^T = 0$ (because a sample of one has no variance). Table 1 displays the full set of variance and segregation expressions under all combinations of the three data collection conditions. For example, estimates of segregation based on decennial Census data have $T = 1$ and $\Delta^T = 0$, which yields

$$B'^{*} = \tau + \sigma + \frac{1}{n}(\omega + \eta),$$

$$V'^{*} = \frac{1}{r}V,$$

and

$$S'^{*} = rS\left(\frac{n-1}{n}\right) + \frac{1}{n}. \tag{10}$$

Both the between-neighborhood and total variance components are biased upwards in this case, when $r < 1$ (which implies $\eta > 0$) and n is finite. The direction of bias in segregation, however, is ambiguous: measurement error biases it downwards, while sampling biases it upwards; the net bias will depend on the specific values of r, n, and S. For example, if true $S = 0.2$ and $r = 0.80$ and $n = 100$, the estimated value of S will be $S'^{*} \approx 0.17$; however, if $r = 0.90$ and $n = 25$, $S'^{*} \approx 0.22$.

[Insert Table 1 about here]

In summary, measurement error and pooling bias segregation estimates downward whereas sampling biases segregation estimates upward. ACS and Census estimates are known to differ in two ways: (1) the ACS has downward bias from pooling unlike the Census and the ACS and (2) the ACS has greater upward bias from sampling at a lower rate than the Census. Due to these countervailing factors, the direction of the bias when comparing ACS and Census estimates is unclear.

Consider Fig. 1, which depicts how segregation is biased under ACS- and Census-type sampling when $S = .3$. Tracking the orange line, we see that an annual sample where true income is observed results in upward bias which is greater at the ACS's lower sampling rate than at the Census's sampling rate. Tracking the blue line, we see that an annual sample of observed income ($r = .75$) leads to severely underestimating income, but this is slightly mitigated when sampling at a lower rate. Tracking the maroon line, we see that a pooled sample of observed income (using our average metro-area estimates of Δ^T and σ as benchmarks) leads to further underestimation. Pooled sampling is a substantially more potent source of downward bias than finite sampling is a source of upward bias, hence ACS-type data which uses a pooled sample of observed income will be more underestimated than Census-type data which uses a larger, annual sample of observed income. That is, there is reason to suspect that Bischoff and Reardon's (2014) estimated segregation increase from 2000 to 2005-2009 was underestimated, not overestimated.

[Insert Fig. 1 about here]

Correcting for Bias in Segregation Estimates

We now turn to the question of how to estimate segregation given these three sources of bias. After defining our generalized estimator for segregation, we consider how to estimate Δ^T and σ using ACS data. This subsection focuses on estimation with continuous income data while the following subsection concerns estimation with coarsened income data. We validate these approaches below and discuss practical considerations in Census and ACS data in the next section in which we apply our method to major metro areas.

Equation (9) can be rearranged to get an estimator for S when any combination of the three biases is present. We observe naïve estimates of total within-neighborhood variance and between-neighborhood variance, which we denote $\widehat{W}$ and $\widehat{B}$ respectively; these variance estimates may have any combination of the three biases depending on the data condition. True segregation is

$$S = \left(\frac{1}{r}\right)\left(\frac{n}{n-1}\right)\left(\frac{\hat{B} + \frac{T-1}{T}\sigma}{\hat{B} + \widehat{W} - \Delta^T} - \frac{1}{n}\right). \tag{11}$$

As before, we set $r = 1$ if there is no measurement error, $n = \infty$ if the data are collected from the full population, and $T = 1$, and $\Delta^T = 0$ if the data are annual rather than pooled.

While we focus on S for simplicity, alternative segregation indices can also be recovered; in Appendix A, we describe how we estimate the rank-order segregation indices used in the Bischoff and Reardon (2014) and Reardon et al. (2018) analyses. Additionally, we can relax the assumptions that J and the total population are large; Appendix B provides equations for when there are finite tracts and/or a finite population.

In any case, we can use Equation (11) to estimate segregation from Census data given tract sample size n and an estimate of r from the literature, but to estimate segregation from ACS data we will need to first estimate Δ^T and σ. To do so, we first consider the case in which neighborhood segregation is computed within metropolitan areas and income data are continuous. We get our estimators for Δ^T and σ from a system of three equations, which we turn to now.

The first equation we consider is crucial to estimating Δ^T. Here, we leverage that the ACS provides publicly available annual income data for metro areas, which allows us to compute the unadjusted sample variance of estimated annual metro means over time, $svar\left(\hat{\bar{Y}}_t\right)$. The observed estimated metro mean in year t is:

$$\hat{\bar{Y}}_t = \bar{Y}_t + u_t, \tag{12}$$

where $\bar{Y}_t$ is the true year t metro mean in constant dollars and u_t is the error in $\hat{\bar{Y}}_t$. Let $\boldsymbol{u}_t \sim MVN(0, \boldsymbol{U})$. The expected sample variance of $\hat{\bar{Y}}_t$ over T years is

$$svar\left(\widehat{\widehat{Y}_t}\right) = \Delta^T + \frac{T-1}{T}U.$$

(13)

We observe $svar\left(\widehat{\widehat{Y}_t}\right)$ and T but need more information to estimate the sampling error variance, U, and be able solve for Δ^T.

The second equation we consider is crucial to estimating σ. Here, we make use of $Var(\widehat{d_J})$, the variance of the observed difference in pooled neighborhood mean income (in constant dollars) between adjacent pooled samples. Note that because this equation requires adjacent pools, it uses $T+1$ years of data. Suppose $T = 5$. The difference, $\widehat{d_J}$, between neighborhood j's observed pooled mean in the years 1 through 5 pool, $\widehat{\overline{Y_{j15}}}$, and its observed pooled mean in the years 0 through 4 pool, $\widehat{\overline{Y_{J04}}}$, is

$$\begin{aligned}
\widehat{d_J} &= \widehat{\overline{Y_{j15}}} - \widehat{\overline{Y_{j04}}} \\
&= \overline{Y_{j15}} - \overline{Y_{j04}} + (q_{j15} - q_{j04}), \\
&= \frac{1}{T}\left[(\delta_5 - \delta_0) + (v_{j5} - v_{j0})\right] + (q_{j15} - q_{j04}),
\end{aligned}$$

(14)

where q_{j15} and q_{j04} are the errors in neighborhood pooled mean income in the two pooled samples. Let $\boldsymbol{q}_{jab} \sim MVN(0, \boldsymbol{Q})$. The variance of $\widehat{d_J}$ is

$$Var(\widehat{d_J}) = \frac{2}{T^2}(\sigma + TQ).$$

(15)

We observe $Var(\widehat{d_J})$ and T, but need more information to estimate the error variance in the pooled neighborhood means, Q, and be able to solve for σ.

Correcting for Bias Using Continuous Income Data

When data are continuous, U and Q are not observed because they rely on the within-unit variance of observed income in the population, W'. Specifically, $U = \frac{TW'}{Jn} + \frac{\sigma}{J}$ and $Q = \frac{W'}{n}$.[2] We can use the equation for $\widetilde{W'}^*$ to define $\widehat{W}$ in terms of W':

$$\widehat{W} = \Delta^T + \frac{T-1}{T}\sigma + \frac{n-1}{n}W', \tag{16}$$

where $n = \infty$ when observing the full population.

This gives us a system of three equations (for $svar(\widehat{Y}_t)$, $Var(\widehat{d}_J)$, and $\widehat{W}$) with three unknowns (σ, W', and Δ^T) which we can solve for Δ^T and σ. Solving for Δ^T:

$$\Delta^T = svar(\widehat{Y}_t) - \frac{T(T-1)}{2J}Var(\widehat{d}_J). \tag{17}$$

and for σ:

$$\sigma = \frac{T\left(n - \frac{J+T-1}{J}\right)}{2\left(\frac{n}{T} - 1\right)} Var(\widehat{d}_J) - \frac{\widehat{W} - svar(\widehat{Y}_t)}{\frac{n}{T} - 1}, \tag{18}$$

where $n = \infty$ when observing the full population.

Correcting for Bias Using Coarsened Income Data

We typically have income data that have been coarsened to preserve privacy. In this case, we can use constrained heteroskedastic ordered probit models (HETOP) to estimate tract and metro area means and variances as needed under the assumption that income is log-normal within neighborhoods (Reardon et

[2] Note in the equation for U that $\frac{\sigma}{J}$ is the variance of the annual metro mean of v_{jt} over years.

al. 2017). Correcting for bias in coarsened data requires a different approach because HETOP estimates include error variance from the HETOP modeling. This modeling error variance inflates U, Q, and $\hat{B}$.

Estimating Δ^T and σ is easier in coarsened data because HETOP's reported standard errors can be used to estimate U and Q. In coarsened data, U and Q include error variance from HETOP modeling in addition to sampling error variance and HETOP's reported standard errors for each estimated tract mean include error variance from both sampling and modeling. Thus, we can estimate U and Q as the average squared standard error of the estimated metro-year means and tract-pool means, respectively, then compute Δ^T and σ by a simple rearranging of terms in Eqs. 13 and 15.

However, there is an added step before plugging into Equation (11) once we have estimated Δ^T and σ, because $\hat{B}$ is inflated by the HETOP modeling error in the neighborhood means. Let $\hat{B}^{coarse}$ refer to the estimated between-neighborhood variance in coarsened data while $\hat{B}$ is the between-neighborhood variance one would estimate if the data were continuous. We estimate $\hat{B}$ by first removing the error variance in the estimated neighborhood means, E, which includes both modeling error variance and sampling variance, then adding back in the sampling variance:

$$\hat{B} = \hat{B}^{coarse} - E + \frac{\widehat{W}}{n}, \tag{19}$$

where E is the error variance in the estimates of the neighborhood means and $\frac{\widehat{W}}{n}$ is the sampling variance of neighborhood sample means. We estimate E from as the average of the squared standard errors that HETOP reports for the estimted neighborhood means.

Assessing the Segregation Estimator's Accuracy in Four Data Conditions

We assess the validity of our approach by simulating $M = 100$ metros over $K = 12$ years according to the data-generating model in Equation (1). In each metro, we set $J = 500$, $n = 120$, $S = .3$, $\Theta = 0$, $r =$

$.75$, $\omega = .7$, $\tau = .27$, $\sigma = .03$, and $\Delta = .001$, values similar to what we observe for major metro areas in the analysis in the following section. We first generate tract-year means $Y_{jt} = \delta_t + \mu_j + v_{jt}$ such that the sample variance of δ_t over the K years is exactly $\frac{K-1}{K}$, the sample variance of μ_j over the J tracts is exactly τ[3], and the sample variance of v_{jt} over the JK tract-years is exactly $\frac{JK}{JK-1}\sigma$. We then generate samples of n households in each tract-year, drawing e_{ijt} and ϵ_{ijt} from normal distributions with variances of ω and η, respectively.

We use the resulting annual continuous data to model three additional conditions: annual coarsened data, pooled continuous data, and pooled coarsened data. When coarsening, we use 10 bins with cutpoints set at the deciles of a normal distribution with variance $\frac{V}{r} + \frac{4}{5}\Delta$ where $\Delta = 0$ if data is annual. When pooling, we draw annual samples of $\frac{n}{T}$ then pool over $T = 5$ years. We estimate uncorrected segregation (S) and corrected estimates, computing $KM = 1200$ estimates in the annual conditions and $(K - T)M = 700$ estimates in the pooled conditions.

[Insert Fig. 2 about here]

Fig. 2 reports our corrected estimates of S compared to the uncorrected estimates in each condition (Appendix C discusses the results for Δ^T and σ). We report the 95 percent confidence interval around each mean estimate to display the bias in the estimates and we report the 90th percentile to 10th percentile range of our estimates to display their imprecision.

Across conditions, our correction substantially reduces bias; across all simulations and conditions, the magnitude of bias is reduced by 99 percent on average. In both continuous data conditions, our estimates are unbiased; in the annual continuous condition, the 95 percent confidence interval of the mean is (.2993,.3002) while in the pooled continuous condition it is (.2996,.3001). There is small downward bias in the annual coarsened and pooled coarsened conditions, which have confidence

[3] In accordance with the assumption that J is sufficiently large to ignore sampling over finite J.

intervals of (.2982,.2991) and (.2982,.2988), respectively. This appears to be due to HETOP models tending to slightly overestimate within-tract variance, as has been noted elsewhere (Reardon et al. 2017).

The corrected estimates are not substantially less efficient than uncorrected estimates, most of the imprecision owing to sampling variability, but the imprecision is nonetheless non-negligible. Across conditions, the standard deviation of the corrected estimates is 32.7 to 36.4 percent greater than that of the uncorrected estimates. This variation in the estimates is some cause for caution when comparing single observations of individual metro areas.

New National Trend Estimates

Data and Measures

In the simulations above, applying our bias-correcting estimator to annual data sharply increased estimated segregation and applying it to pooled data raised estimated segregation further still. This suggests that the various estimates of national income segregation at the center of recent debates are all severe underestimates. Moreover, given the change from estimates based on annual Census data in 2000 to estimates based on pooled ACS data starting in 2005-2009, the national trend may differ substantially from the various estimates in prior research.

We produce new national trend estimates using our bias-correcting estimator. Our data relies primarily on 1990 and 2000 Census and 2005-2009 through 2015-2019 ACS tabulations of tract-level household income in constant dollars. We focus on the Bischoff and Reardon (2014) sample of the 116 largest metro areas. To facilitate HETOP modeling, we further coarsen income from 16 to 8 bins by combining adjacent bins.[4] We also exclude the small number of tracts that have small populations (fewer

[4] Note that the ACS reports estimated population counts, which can be scaled to approximate sample bin counts by multiplying by the sampling rate. Alternatively, the standard errors on estimates using population bin counts can be scaled up by dividing by the square root of the sampling rate.

than 50 households) or sparse income distributions (observations in fewer than 5 bins). In 2015-2019, for example, these exclusions remove 11 of the 45,882 tracts with non-zero populations.

Applying our estimator requires estimates of r, which we derive from the literature. In Appendix D, we review the literature on the reliability of self-reported income. The literature provides a wide range of estimates; studies use a variety of methods, compare observed data to a variety of approximations of true values, and estimate the reliability of various types of income reports (e.g., household vs. personal income, total income vs. wage income, etc.). As a result, it is not clear exactly how reliable income reports are, though most evidence suggests that the reliability of self-reported household income in the Census and ACS is roughly between $.7$ and $.8$. We use the midpoint, $r = .75$, to produce our primary estimates. We treat the reliability of income as stable over time and place and assume it is the same for the Census and ACS income items, which have the same design.

To apply our estimator to the pooled ACS data, we need to estimate Δ^T and σ. We estimate a constant value of Δ^T in each metro area using ACS 1-year data from 2010 (when the data first became available) to 2019.[5] We estimate $svar(Y_t')$ when $T = 5$ by estimating the population variance of metro mean income over all 10 years with available data, then multiply this by $\frac{4}{5}$ to estimate the 5-year sample variance.[6] This improves precision at the cost of forcing Δ^T to be constant over time within metros. We estimate a 6-year running average of σ in each metro area using adjacent 5-year pools from the ACS in 2005-2009 (the start of the ACS) through 2015-2019. In practice, the σ estimates are imprecise, which

[5] We estimate a constant value of Δ^T due to data constraints. Moving averages of Δ^T would be exceedingly noisy (see Fig. C2) and require substantial imputation to account for missing observations, most critically during 2005-2009 when we do not observe annual metro means.

[6] Of the 116 metro areas in our analytic sample, 103 appear in all 10 years; 2 appear in the first 9 years and 8 appear in the first 3 years so we estimate Δ^{T} from those years in these areas; 2 are missing, so we impute Δ^{T} in these areas as the mean over the observed metros; and the Philadelphia metro area appears in all 10 years but it has an unbelievable income decline from 2012 to 2013, so we estimate Δ^{T} separately using the 2010-2012 and 2013-2019 time spans and use the observation-weighted average of these estimates as Δ^{T} in our analyses.

can inflate how much σ appears to vary over time and across metros. In Appendix E, we describe a procedure we use to reduce imprecision and better estimate variation in σ.

We estimate household income segregation between tracts within metro areas and we estimate it in three metrics: S, H, and R. While we use S above for illustrative purposes, it is not a standard income segregation measure. It is more common to estimate income segregation using the rank-order information theory index, H^R, and the rank-order variance ratio index, R^R, measures for computing segregation of a coarsened continuous variable (Reardon and Bischoff 2011). More crucially for our purposes, these are the metrics in which Bischoff and Reardon (2014) report their unadjusted estimates and Reardon et al. (2018) and Logan et al. (2018, 2020) report their finite sampling bias-adjusted estimates. Appendix A describes how we extend our method to estimating rank-order segregation and includes a simulation assessment of the bias-reduction using this strategy.

Our estimation procedure is unbiased under the simplifying assumptions that income is log-normal within tracts and observed with measurement error that is classical in the logged income metric. Tract income distributions are of course messier than our simple data generating model, so we assessed whether this assumption biases our estimates and applied a post-hoc adjustment to remove the bias. We estimate unadjusted rank-order income segregation in H^R and R^R in our analytic sample in two ways: first, using the income tabulations provided by the Census or ACS and, second, using the income tabulations implied by the HETOP-estimated means and variances (see Appendix A Eqs. A1, A2, and A5 for more details). The latter estimates, which assume income is well-behaved, tend to provide overestimates of unadjusted segregation compared to the tabulation-based estimates. The bias is smaller in later years (no more than 2.4 percent from 2006-2010 on) and is greatest in 1990 where the estimates are 11.1 percent greater. Our post-hoc adjustment removes this bias in each metro-year estimate by multiplying our adjusted segregation estimates by the ratio of the tabulation-based unadjusted estimates to the HETOP-based unadjusted estimates.[7]

Findings

Fig. 3 presents the estimated national trend in each segregation metric – S, H, and R – when $r = .75$. The top panel is the trend in each metric while the bottom panel is the trend as a percentage change since 1990 to ease comparisons across metrics. The trend is similar across metrics. The rank-order segregation trends, H and R, run nearly parallel to one another with estimates in R consistently above those in H such that they almost perfectly align when presented as a percentage change. The trend in S, which is roughly double the rank-order measures in scale, has a similar pattern but with a noticeably less steep percentage increase from 2000 to 2005-2009 (plotted at the midpoint, 2007). Across the three metrics, we find that income segregation declined by roughly 4 percent from 1990 to 2000 then increased substantially through 2006-2010 such that the net increase since 1990 was 12.1 percent in S, 19.0 percent in R, and 20.3 percent in H. Over the following 8 years, segregation declined such that, by 2014-2018, it was only about 5 percent above the 1990 value in the rank-order metrics and had returned to the 1990 value in the S metric. Segregation ticked back up between then and our last observation, the 2015-2019 period.

[Insert Fig. 3 about here]

[Insert Fig. 4 about here]

Fig. 4 compares our estimates of the national trend to what one would estimate using extant methods, with the detailed estimates reported in Tables 2 and 3. We provide 3 types of estimates: an unadjusted estimate following the procedure of Bischoff and Reardon (2014), an estimate adjusted for

[7] We can only compute this ratio in the H and R metrics and it is similar across the two metrics, so the ratio we use for our adjusted estimates in the S metric is the average of the H and R ratios.

finite sampling following the procedure of Reardon et al. (2018), denoted RBOT estimate; and new estimates that apply our correction to adjust for finite sampling, attenuation, and pooled sampling. We report three new estimates at each time point depending on the reliability of income, which is estimated with a wide range in the literature. The top panel of Fig. 4 and Table 2 provide estimates in the H metric while the bottom panel of Fig. 4 and Table 3 provide estimates in the R metric.

We see 3 main takeaways in these findings. First, segregation is far greater than previously believed. In the 2015-2019 ACS sample, we estimate that if the reliability of self-reported household income is $r = .75$, the average segregation across major metro areas in the H metric is 52.1 percent greater than unadjusted estimates suggest and 64.4 percent greater than would be estimated using the Reardon et al. (2018) correction for finite sampling bias. The findings are similar when using the R metric; segregation is 49.5 percent greater than unadjusted estimates suggest and 59.2 percent greater than estimates following the Reardon et al. (2018) procedure. Most of the downward bias in the unadjusted and Reardon et al. (2018) estimates is due to attenuation bias; attenuation bias accounts for the entirety of the difference between our estimates and the Reardon et al. (2018) estimates in the Census years, where our estimates show that segregation is over 40 percent greater than previously believed.

Second, the increase from 2000 to 2005-2009 (plotted at 2007) was much greater – not lower – than what we find when we use unadjusted estimates like Bischoff and Reardon (2014). This marks the switch from Census to ACS data, in which the additional downward bias from pooled sampling far outweighs the increase upward bias from sampling at a lower rate. Using unadjusted estimates yields an estimated increase of .0049 in the H metric, or 5.0 percent. Following the Reardon et al. (2018) procedure, we estimate a substantially smaller increase of .0015 (1.5 percent). At $r = .75$, we estimate that segregation actually increased by .0318 (23.4 percent), a change that was masked by unaccounted for pooling bias in the other estimates. The findings are similar when using the R metric; we estimate that segregation increased by 21.4 percent compared to changes of 3.8 percent using unadjusted estimates and 0.8 percent using finite sampling bias-adjusted estimates.

Third, the decline in segregation after 2005-09 is underestimated in prior estimation methods because those estimates include time-varying downward bias from pooled sampling bias. The difference in trends is starkest when comparing 2011-2015 to 2005-2009. Over this time, both unadjusted estimates and estimates following the Reardon et al. (2018) procedure find a small increase in segregation regardless of metric whereas we find a decrease of .0109 (6.5 percent) in the H metric and a decrease of .0113 (6.2 percent) in the R metric when $r = .75$. Our estimates of σ over this period show that pooled sampling bias was decreasing in magnitude as tract-level volatility settled down following the Great Recession, tilting the trend upward and masking the decline in estimates that do not correct for it (Appendix F Fig. F2).

This speaks to the sensitivity of segregation trend estimates to the trend in neighborhood-level income volatility when using pooled sample data, which will be standard practice for the foreseeable future. For example, we might suspect that both segregation and tract income volatility have recently been increasing during the COVID-19 pandemic, meaning both segregation and the downward pooled sampling bias are greater. If we estimate current segregation with future ACS data without adjusting for pooled sampling, it is plausible income segregation will appear steady despite having increased. Correcting for pooled sampling bias will be crucial for accurately estimating segregation trends for as long as researchers are reliant on pooled ACS data.

[Insert Table 2 about here]

[Insert Table 3 about here]

Discussion

We consider three sources of potential bias in recent income segregation estimates, yielding several important findings. We extend the recent discussion of upward bias in sample-based segregation measures by considering two sources of downward bias that have received less attention. We confirm that segregation measures using noisy characteristic data, like self-reported income, are severely biased downward. We also demonstrate that segregation measures drawing from pooled samples, like we find in ACS data, are biased downward. This bias can be substantial when the characteristic of interest varies over time; our estimates of tract- and metro-level income variation in the ACS indicate that pooling bias is typically greater in magnitude than finite sampling bias in the case of income segregation. Ignoring any of these three sources of bias may lead to incorrect inferences.

Additionally, formalizing the relationship between the three sources of bias indicates that the direction of bias is often unclear, including in both Census- and ACS-type data. We also show how one can use a single procedure to compute segregation estimates that are largely free of the three sources of bias. We demonstrate that this is possible using only publicly available Census and ACS tabulations and the extant literature on the reliability of income.

Our bias-corrected estimates indicate that income segregation is on the order of 50 percent greater than previously believed due to attenuation and pooled sampling bias. The increase in segregation from 2000 to the 2005-09 period estimated without adjustment is an underestimate by more than a factor of four, rather than an overestimate as has been argued. The decline in segregation in the years following 2005-09 has been underestimated due to pooled sampling bias; from 2005-2009 to 2011-2015, a period of declining tract-level volatility and consequently declining pooled sampling bias, segregation declined by over 6 percent, but estimates ignoring pooled sampling bias instead find a small increase. The latter finding is particularly noteworthy because it indicates that pooled sampling bias substantially distorts unadjusted segregation trends when using only ACS data such that correcting for pooled sampling bias will be crucial for the foreseeable future.

Though the focus of this paper is the measurement of income segregation, the new estimates of the national income segregation trend also have implications for how we understand income segregation. Reardon and Bischoff (2011) estimated that a 1 SD increase in income inequality increases income segregation by .25 SD, providing one account for what drives the trends we observe here. However, it remains unexplained why we find segregation decline in the 1990s when income inequality substantially increased and why find such a large increase in segregation from 2000 to 2005-2009, a period of only modestly increasingly inequality (World Bank 2021). Our estimates call for further examination of what drives changes in income segregation and, potentially, identification of additional mechanisms.

The three types of segregation measurement biases we consider are not limited to income segregation estimates, the segregation indices discussed here, or the datasets we focus on. As others have discussed in more detail, finite sampling bias is relevant to all sample-based measures of segregation (Logan et al. 2018; Reardon et al. 2018).

Similarly, all segregation indices are biased downwards by noise in the characteristic of interest, whether due to reporting error (e.g., self-reported educational attainment), the use of data reduction techniques (e.g., SES), or using a noisy proxy (e.g., using annual income to proxy for permanent income). Attenuation bias can also lead to erroneous inferences when comparing across times, groups, or datasets in which the reliability of the characteristic of interest differs. For example, income segregation comparisons across groups that differ by educational attainment, race/ethnicity, or occupations may be biased by group differences in response error (Kim and Tamborini 2012).

All segregation indices are potentially biased downwards by pooled sampling when data are collected as a pooled sample, as in the ACS, or when researchers combine samples to increase statistical power. This is a relevant concern regardless of the characteristic of interest, but it depends on whether there is unit-level (e.g., neighborhood-level) variation over time within the pooling period. This may occur primarily because people's characteristics change over time, as it does in the case of income, or because of mobility across units. In addition to pooled sampling in the ACS biasing the comparison of Census and ACS data, we find that changes over time in tract-level income volatility masks the recent downward trend in income segregation estimates that ignore pooled sampling bias. Comparing neighborhood segregation across groups in pooled samples will also be biased if the groups have different levels of neighborhood mobility and/or volatility in the characteristic of interest.

Moreover, outside of the case of annual population data, the aggregate bias in segregation measures is both multiplicative and additive such that the bias depends on the true segregation level whether comparing raw or proportional differences.

The bias-corrected estimator we describe here will yield approximately unbiased estimates in each of these cases, however. Researchers can use it to make valid comparisons when true segregation, sample sizes, unit-level volatility, and reliability differ. In the case of national income segregation trends, this approach allows us to compare income segregation across metro areas and years with different sample sizes, levels of true segregation, neighborhood-level income volatility, and metro-level income volatility. It would also allow comparisons across metro areas and years with different reliabilities of income if such information was available. Though our estimator assumes a large number of units and a large population, researchers can use the adjustments in Appendix B to estimate segregation in smaller populations.

Bias-corrected estimates will not be accurate in all cases. As we are using publicly available data, we do not consider bias due to weighted sampling, which Logan et al. (2020) find poses a larger problem for race-specific analyses for which sample weights differ more. Our estimates of the national trend are likely overestimated in light of bias from weighted sampling, though the size of this bias is small relative to the three biases considered and the uncertainty in segregation due to uncertainty in income reliability. Logan et al. (2020) provide formulas correcting for weighted sampling that, like our estimator, estimate segregation by estimating within-unit variance and total variance, so extending our estimator to accommodate weighted sampling is straightforward when the necessary data are available.

Our estimator cannot, however, be readily extended to account for violations of the simplified data-generating model's log-normality assumptions, the foremost threat being non-classical reliability. Yet we can speculate how our estimates may be affected. If lower-income households report their incomes with more error, as Kim and Tamborini (2012) report, measurement error widens income distributions more than we have assumed such that we may have under-corrected our estimates. Conversely, we may have over-corrected our estimates if self-report errors at the tails are skewed toward the population mean (Kim and Tamborini 2012; Pedace and Bates 2000; c.f. Bingley and Martinello 2017). Non-classical response error also raises the possibility that response errors are related to tract- and metro-level volatility; for example, if self-reports err toward period means, our estimates may have over-corrected for attenuation bias.

Income segregation is surprisingly difficult to measure from Census or ACS data. In this paper we partially address the challenges of measuring it. We formalize three important threats to segregation measures. Note that these sources of bias pertain equally to aggregated, publicly reported Census and ACS data and to household-level Census and ACS microdata. Given these sources of bias, we provide a strategy for estimating segregation when income and bias are well-behaved and income reliability is known. Our new estimates of the national income segregation trend demonstrate that previous estimates of both the levels and trends in income segregation were severely biased. That said, there remains some uncertainty about the actual trend, because our estimates rely on simplifying assumptions about the shape of income distributions and estimates of the reliability of self-reported income. Violations of these assumptions or variation (over time or place) in the reliability of self-reported income might bias the trends further in ways we have not been able to correct for. Nonetheless, we hope the groundwork laid here improves our understanding of the measurement and trends in income segregation.

References

Abowd, John M., and Martha H. Stinson. 2013. "Estimating Measurement Error in Annual Job Earnings: A Comparison of Survey and Administrative Data." *The Review of Economics and Statistics* 95(5):1451–67. doi: 10.1162/REST_a_00352.

Alwin, Duane F. 2007. *Margins of Error: A Study of Reliability in Survey Measurement*. John Wiley & Sons.

Bielby, William T., Robert M. Hauser, and David L. Featherman. 1977. "Response Errors of Black and Nonblack Males in Models of the Intergenerational Transmission of Socioeconomic Status." *American Journal of Sociology* 82(6):1242–88. doi: 10.1086/226465.

Bingley, Paul, and Alessandro Martinello. 2017. "Measurement Error in Income and Schooling and the Bias of Linear Estimators." *Journal of Labor Economics* 35(4):1117–48. doi: 10.1086/692539.

Bischoff, Kendra, and Sean F. Reardon. 2014. "Residential Segregation by Income, 1970–2009." *Diversity and Disparities: America Enters a New Century* 43.

Brady, David, Marco Giesselmann, Ulrich Kohler, and Anke Radenacker. 2018. "How to Measure and Proxy Permanent Income: Evidence from Germany and the U.S." *The Journal of Economic Inequality* 16(3):321–45. doi: 10.1007/s10888-017-9363-9.

Chetty, Raj, and Nathaniel Hendren. 2018a. "The Impacts of Neighborhoods on Intergenerational Mobility I: Childhood Exposure Effects*." *The Quarterly Journal of Economics* 133(3):1107–62. doi: 10.1093/qje/qjy007.

Chetty, Raj, and Nathaniel Hendren. 2018b. "The Impacts of Neighborhoods on Intergenerational Mobility II: County-Level Estimates*." *The Quarterly Journal of Economics* 133(3):1163–1228. doi: 10.1093/qje/qjy006.

Chetty, Raj, Nathaniel Hendren, and Lawrence F. Katz. 2016. "The Effects of Exposure to Better Neighborhoods on Children: New Evidence from the Moving to Opportunity Experiment." *American Economic Review* 106(4):855–902. doi: 10.1257/aer.20150572.

Chetty, Raj, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez. 2014. "Where Is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States *." *The Quarterly Journal of Economics* 129(4):1553–1623. doi: 10.1093/qje/qju022.

Dickens, William T., and Frank Levy. 2003. "Comments." *Brookings-Wharton Papers on Urban Affairs* 2003(1):30–38. doi: 10.1353/urb.2003.0003.

Fahle, Erin M., and Sean F. Reardon. 2018. "How Much Do Test Scores Vary Among School Districts? New Estimates Using Population Data, 2009–2015." *Educational Researcher* 47(4):221–34. doi: 10.3102/0013189X18759524.

Fahle, Erin M., Sean F. Reardon, Demetra Kalogrides, Ericka S. Weathers, and Heewon Jang. 2020. "Racial Segregation and School Poverty in the United States, 1999–2016." *Race and Social Problems* 12(1):42–56. doi: 10.1007/s12552-019-09277-w.

Gottschalk, Peter, and Minh Huynh. 2010. "Are Earnings Inequality and Mobility Overstated? The Impact of Nonclassical Measurement Error." *The Review of Economics and Statistics* 92(2):302–15. doi: 10.1162/rest.2010.11232.

Haider, Steven, and Gary Solon. 2006. "Life-Cycle Variation in the Association between Current and Lifetime Earnings." *American Economic Review* 96(4):1308–20. doi: 10.1257/aer.96.4.1308.

Hyslop, Dean R. 2001. "Rising U.S. Earnings Inequality and Family Labor Supply: The Covariance Structure of Intrafamily Earnings." *American Economic Review* 91(4):755–77. doi: 10.1257/aer.91.4.755.

Jargowsky, Paul A. 1996. “Take the Money and Run: Economic Segregation in U.S. Metropolitan Areas.” *American Sociological Review* 61(6):984–98. doi: 10.2307/2096304.

Kim, ChangHwan, and Christopher R. Tamborini. 2012. “Do Survey Data Estimate Earnings Inequality Correctly? Measurement Errors Among Black and White Male Workers*.” *Social Forces* 90(4):1157–81. doi: 10.1093/sf/sor042.

Lichter, Daniel T., Domenico Parisi, and Michael C. Taquino. 2012. “The Geography of Exclusion: Race, Segregation, and Concentrated Poverty.” *Social Problems* 59(3):364–88. doi: 10.1525/sp.2012.59.3.364.

Logan, John R., Andrew Foster, Jun Ke, and Fan Li. 2018. “The Uptick in Income Segregation: Real Trend or Random Sampling Variation?” *American Journal of Sociology* 124(1):185–222. doi: 10.1086/697528.

Logan, John R., Andrew Foster, Hongwei Xu, and Wenquan Zhang. 2020. “Income Segregation: Up or Down, and for Whom?” *Demography* 57(5):1951–74. doi: 10.1007/s13524-020-00917-0.

Logan, John R., and Harvey Luskin Molotch. 1987. *Urban Fortunes: The Political Economy of Place*. University of California Press.

Massey, Douglas S., Jonathan Rothwell, and Thurston Domina. 2009. “The Changing Bases of Segregation in the United States.” *The ANNALS of the American Academy of Political and Social Science* 626(1):74–90. doi: 10.1177/0002716209343558.

Mazumder, Bhashkar. 2001. *The Mis-Measurement of Permanent Earnings: New Evidence from Social Security Earnings Data*. *SSRN Scholarly Paper*. ID 296131. Rochester, NY: Social Science Research Network. doi: 10.2139/ssrn.296131.

Moore, Jeffrey C., Linda L. Stinson, and Edward J. Welniak. 2000. “Income Measurement Error in Surveys: A Review.” *Journal of Of®cial Statistics* 16(4):331–61.

Owens, Ann. 2016. “Inequality in Children’s Contexts: Income Segregation of Households with and without Children.” *American Sociological Review* 81(3):549–74. doi: 10.1177/0003122416642430.

Owens, Ann. 2017. “Racial Residential Segregation of School-Age Children and Adults: The Role of Schooling as a Segregating Force.” *RSF: The Russell Sage Foundation Journal of the Social Sciences* 3(2):63–80. doi: 10.7758/RSF.2017.3.2.03.

Owens, Ann, Sean F. Reardon, and Christopher Jencks. 2016. “Income Segregation Between Schools and School Districts.” *American Educational Research Journal* 53(4):1159–97. doi: 10.3102/0002831216652722.

Park, Robert E. 1915. “The City: Suggestions for the Investigation of Human Behavior in the City Environment.” *American Journal of Sociology* 20(5):577–612. doi: 10.1086/212433.

Pedace, Roberto, and Nancy Bates. 2000. “Using Administrative Records to Assess Earnings Reporting Error in the Survey of Income and Program Participation.” *Journal of Economic and Social Measurement* 26(3–4):173–92. doi: 10.3233/JEM-2000-0180.

Reardon, Sean F., and Kendra Bischoff. 2011. “Income Inequality and Income Segregation.” *American Journal of Sociology* 116(4):1092–1153. doi: 10.1086/657114.

Reardon, Sean F., Kendra Bischoff, Ann Owens, and Joseph B. Townsend. 2018. “Has Income Segregation Really Increased? Bias and Bias Correction in Sample-Based Segregation Estimates.” *Demography* 55(6):2129–60. doi: 10.1007/s13524-018-0721-4.

Reardon, Sean F., Demetra Kalogrides, and Andrew Ho. 2017. "Linking U.S. School District Test Score Distributions to a Common Scale (CEPA Working Paper No.16-09)." Retrieved from Stanford Center for Education Policy Analysis: http://cepa.stanford.edu/wp16-0.

Reardon, Sean F., Demetra Kalogrides, and Kenneth Shores. 2019. "The Geography of Racial/Ethnic Test Score Gaps." *American Journal of Sociology* 124(4):1164–1221. doi: 10.1086/700678.

Reardon, Sean F., Ericka S. Weathers, Erin M. Fahle, Heewon Jang, and Kalogrides, Demetra. 2019. "Is Separate Still Unequal? New Evidence on School Segregation and Racial Academic Achievement Gaps." *CEPA Working Papers* 19–06.

Reardon, Sean, Benjamin R. Shear, Katherine E. Castellano, and Andrew D. Ho. 2017. "Using Heteroskedastic Ordered Probit Models to Recover Moments of Continuous Test Score Distributions From Coarsened Data." *Journal of Educational and Behavioral Statistics* 42(1):3–45. doi: 10.3102/1076998616666279.

Rothstein, Jesse, and Nathan Wozny. 2013. "Permanent Income and the Black-White Test Score Gap." *Journal of Human Resources* 48(3):510–44. doi: 10.3368/jhr.48.3.510.

Sampson, Robert J. 2012. "Moving and the Neighborhood Glass Ceiling." *Science* 337(6101):1464–65. doi: 10.1126/science.1227881.

Sampson, Robert J., Stephen W. Raudenbush, and Felton Earls. 1997. "Neighborhoods and Violent Crime: A Multilevel Study of Collective Efficacy." *Science* 277(5328):918–24. doi: 10.1126/science.277.5328.918.

Siegel, Paul M., and Robert W. Hodge. 1968.in *Methodology in social research*, edited by H. M. Blalock and A. B. Blalock. New York, McGraw-Hill.

Stinson, Martha H. 2002. “Estimating Measurement Error in SIPP Annual Job Earnings: A Comparison of Census Survey and SSA Administrative Data.” *Center for Economic Studies, US Census Bureau* No. 2002-24.

US Bureau of the Census. 1968. *The Current Population Survey Reinterview Program, January 1961 Through December 1966*. U.S. Government Printing Office.

Withey, Stephen B. 1954. “Reliability of Recall of Income.” *Public Opinion Quarterly* 18(2):197–204. doi: 10.1086/266505.

Wodtke, Geoffrey T., Felix Elwert, and David J. Harding. 2016. “Neighborhood Effect Heterogeneity by Family Income and Developmental Period.” *American Journal of Sociology* 121(4):1168–1222. doi: 10.1086/684137.

World Bank. 2021. “GINI Index for the United States.” *FRED, Federal Reserve Bank of St. Louis*. Retrieved March 30, 2022 (https://fred.stlouisfed.org/series/SIPOVGINIUSA).

Table 1 Bias in segregation and variance components under different data collection conditions.

	Within-Unit Variance	Bias	Between-Unit Variance	Bias	Total Variance	Bias	Segregation	Bias
Annual True Income in Population	$W = \omega$	0	$B = \tau + \sigma$	0	$V = \tau + \sigma + \omega$	0	$S = \frac{\tau+\sigma}{\tau+\sigma+\omega}$	0
Annual True Income in Sample	$W^* = \frac{n-1}{n}\omega$	-	$B^* = B + \frac{1}{n}\omega$	+	$V^* = V$	0	$S^* = \frac{n-1}{n}S + \frac{1}{n}$	+
Pooled True Income in Population	$\widetilde{W} = \Delta^T + \frac{T-1}{T}\sigma + \omega$	+	$\tilde{B} = \tau + \frac{1}{T}\sigma$	-	$\tilde{V} = V + \Delta^T$	+	$\tilde{S} = \left(S - \frac{T-1}{T}\left(\frac{\sigma}{V}\right)\right) \cdot \frac{V}{V+\Delta^T}$	-
Annual Observed Income in Population	$W' = \omega + \eta$	+	$B' = B$	0	$V' = \frac{V}{r}$	+	$S' = rS$	-
Annual Observed Income in Sample (e.g., Census)	$W'^* = \frac{n-1}{n}(\omega + \eta)$	?	$B'^* = B + \frac{1}{n}(\omega + \eta)$	+	$V'^* = \frac{V}{r}$	+	$S'^* = \frac{n-1}{n}rS + \frac{1}{n}$	?
Pooled True Income in Sample	$\widetilde{W}^* = \Delta^T + \frac{T-1}{T}\sigma + \frac{n-1}{n}\omega$	?	$\tilde{B}^* = \tau + \frac{1}{T}\sigma + \frac{1}{n}\omega$	?	$\tilde{V}^* = V + \Delta^T$	+	$\tilde{S}^* = \left(\frac{n-1}{n}S + \frac{1}{n} - \frac{T-1}{T}\left(\frac{\sigma}{V}\right)\right)$	?
Pooled Observed Income in Population	$\widetilde{W}' = \Delta^T + \frac{T-1}{T}\sigma + \omega + \eta$	+	$\tilde{B}' = \tau + \frac{1}{T}\sigma$	-	$\tilde{V}' = \frac{V}{r} + \Delta^T$	+	$\tilde{S}' = r\left(S - \frac{T-1}{T}\left(\frac{\sigma}{V}\right)\right) \cdot \frac{V}{V+\Delta^T}$	-
Pooled Observed Income in Sample (e.g., ACS)	$\widetilde{W}'^* = \Delta^T + \frac{T-1}{T}\sigma + \frac{n-1}{n}(\omega + \eta)$	?	$\tilde{B}'^* = \tau + \frac{1}{T}\sigma + \frac{1}{n}(\omega + \eta)$	?	$\tilde{V}'^* = \frac{1}{r}V + \Delta^T$	+	$\tilde{S}'^* = r\left(\frac{n-1}{n}S + \frac{1}{rn} - \frac{T-1}{T}\left(\frac{\sigma}{V}\right)\right) \cdot \frac{V}{V+\Delta^T}$	?

Table 2 Rank-H^R household income segregation estimates, by year and estimate type.

Year	Unadjusted Estimate	RBOT Estimate	New Estimates		
			$r = .7$	$r = .75$	$r = .8$
1990	.1009	.0973	.1547	.1415	.1299
	(.0232)	(.0231)	(.0393)	(.0355)	(.0322)
2000	.0977***	.0940***	.1488***	.1357***	.1247***
	(.0203)	(.0203)	(.0346)	(.0309)	(.0279)
2005-2009	.1026***	.0955*	.1848***	.1675***	.1532***
	(.0209)	(.0206)	(.0390)	(.0346)	(.0311)
2006-2010	.1068***	.0985***	.1875*	.1702*	.1557**
	(.0211)	(.0209)	(.0388)	(.0349)	(.0313)
2007-2011	.1061	.0979	.1811***	.1642***	.1502***
	(.0210)	(.0208)	(.0384)	(.0342)	(.0306)
2008-2012	.1052	.0975	.1784*	.1620*	.1483
	(.0208)	(.0206)	(.0379)	(.0337)	(.0303)
2009-2013	.1045	.0970	.1783	.1620	.1482
	(.0204)	(.0202)	(.0368)	(.0329)	(.0296)
2010-2014	.1038	.0965	.1755*	.1591**	.1457**
	(.0207)	(.0206)	(.0373)	(.0333)	(.0299)
2011-2015	.1036	.0966	.1727*	.1566*	.1433*
	(.0209)	(.0208)	(.0379)	(.0335)	(.0302)
2012-2016	.1025	.0956	.1684***	.1527***	.1399***
	(.0210)	(.0209)	(.0376)	(.0335)	(.0301)
2013-2017	.1007	.0936	.1644	.1491	.1365
	(.0208)	(.0207)	(.0370)	(.0330)	(.0297)
2014-2018	.0991	.0917	.1639	.1486	.1361
	(.0209)	(.0208)	(.0371)	(.0330)	(.0297)
2015-2019	.1009	.0934	.1693	.1535	.1403
	(.0207)	(.0207)	(.0373)	(.0330)	(.0296)

Notes: Cells report means; standard deviations are shown in parentheses. Statistical significance tests come from regression models with metro area fixed effects that compare the estimate with that in the prior observation. RBOT estimates apply the Reardon et al. (2018) finite sampling bias correction. Sample is the 116 metro areas analyzed in Reardon et al. (2018).

$*p \leq .05$; $**p \leq .01$; $***p \leq .001$

Table 3 Rank-R^R household income segregation estimates, by year and estimate type.

Year	Unadjusted Estimate	RBOT Estimate	New Estimates $r = .7$	$r = .75$	$r = .8$
1990	.1131 (.0261)	.1098 (.0261)	.1702 (.0432)	.1563 (.0392)	.1441 (.0358)
2000	.1103*** (.0230)	.1069*** (.0230)	.1650*** (.0381)	.1511*** (.0343)	.1394*** (.0311)
2005-2009	.1145*** (.0232)	.1078 (.0231)	.2014*** (.0420)	.1834*** (.0375)	.1685*** (.0339)
2006-2010	.1188*** (.0235)	.1110*** (.0235)	.2038 (.0418)	.1859* (.0378)	.1708* (.0341)
2007-2011	.1181 (.0233)	.1104 (.0233)	.1970*** (.0412)	.1795*** (.0370)	.1649*** (.0333)
2008-2012	.1172 (.0231)	.1099 (.0230)	.1943* (.0408)	.1773 (.0365)	.1630 (.0330)
2009-2013	.1165 (.0227)	.1095 (.0227)	.1944 (.0397)	.1775 (.0357)	.1631 (.0323)
2010-2014	.1158 (.0230)	.1090 (.0230)	.1917* (.0402)	.1746* (.0361)	.1605* (.0327)
2011-2015	.1158 (.0233)	.1092 (.0233)	.1889* (.0409)	.1721* (.0365)	.1582* (.0331)
2012-2016	.1146 (.0235)	.1081 (.0236)	.1844*** (.0409)	.1680*** (.0366)	.1546*** (.0332)
2013-2017	.1126 (.0233)	.1059 (.0234)	.1800 (.0404)	.1640 (.0362)	.1508 (.0328)
2014-2018	.1107 (.0236)	.1039 (.0236)	.1794 (.0407)	.1635 (.0364)	.1503 (.0330)
2015-2019	.1120 (.0237)	.1052 (.0237)	.1837 (.0411)	.1674 (.0368)	.1537 (.0333)

Notes: Cells report means; standard deviations shown in parentheses. Statistical significance tests come from regression models with metro area fixed effects that compare the estimate with that in the prior observation. RBOT estimates apply the Reardon et al. (2018) finite sampling bias correction. Sample is the 116 metro areas analyzed in Reardon et al. (2018).
$^*p \leq .05$; $^{**}p \leq .01$; $^{***}p \leq .001$

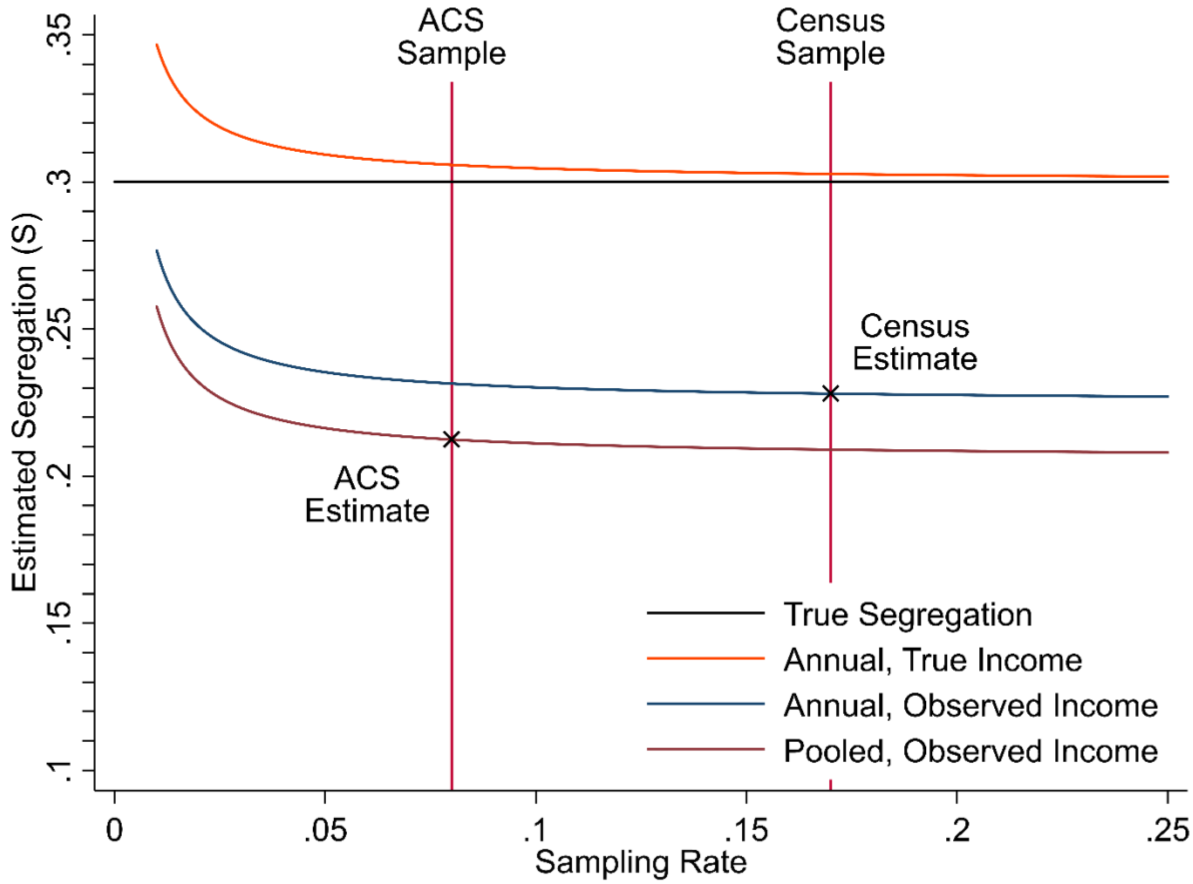

Fig. 1 Simulated estimates of segregation (S) by sampling rate and data collection condition.

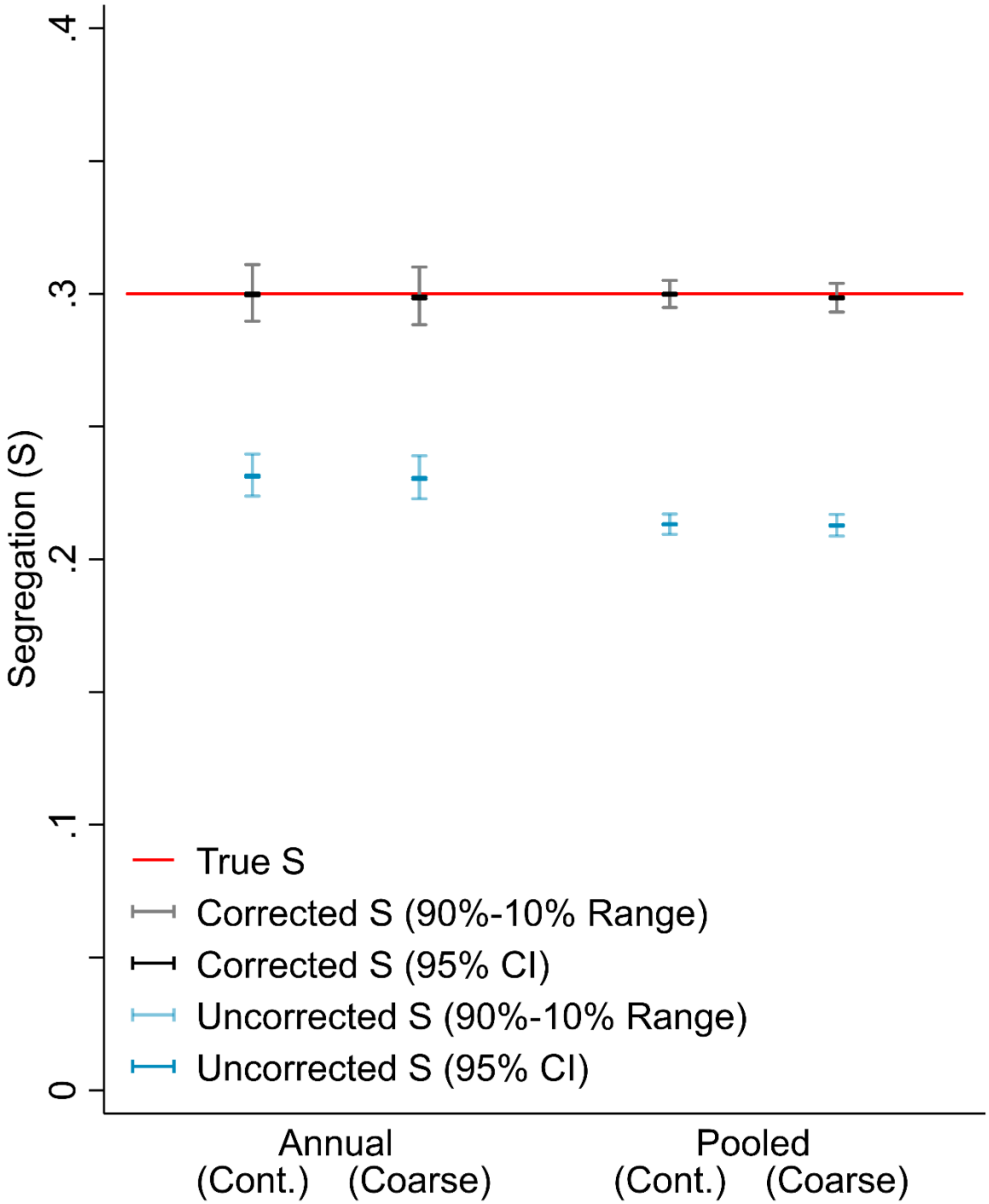

Fig. 2. Bias reduction in corrected segregation (S), by data collection condition, in simulations where the 90-10% range is a range of estimates (indicating imprecision) and the 95% CI is the confidence interval of the mean estimate (indicating bias).

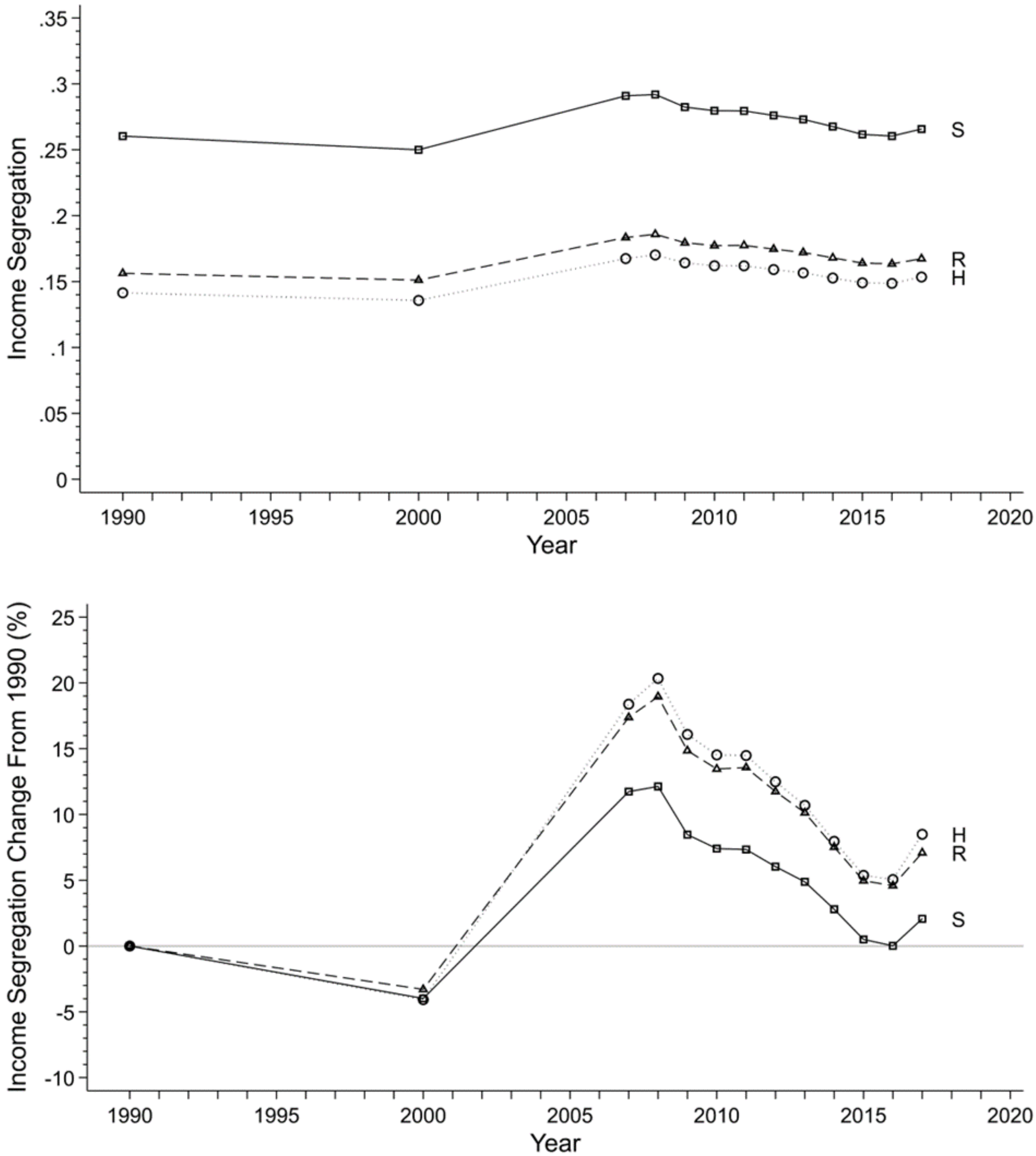

Fig. 3 Household income segregation trend estimates in S, Rank-R^R (R), and Rank-H^R (H) presented as a metric value (top) and as a percentage change relative to 1990 (bottom), with ACS 5-year pooled estimates plotted at middle year, for the 116 metro areas analyzed in Reardon et al. (2018).

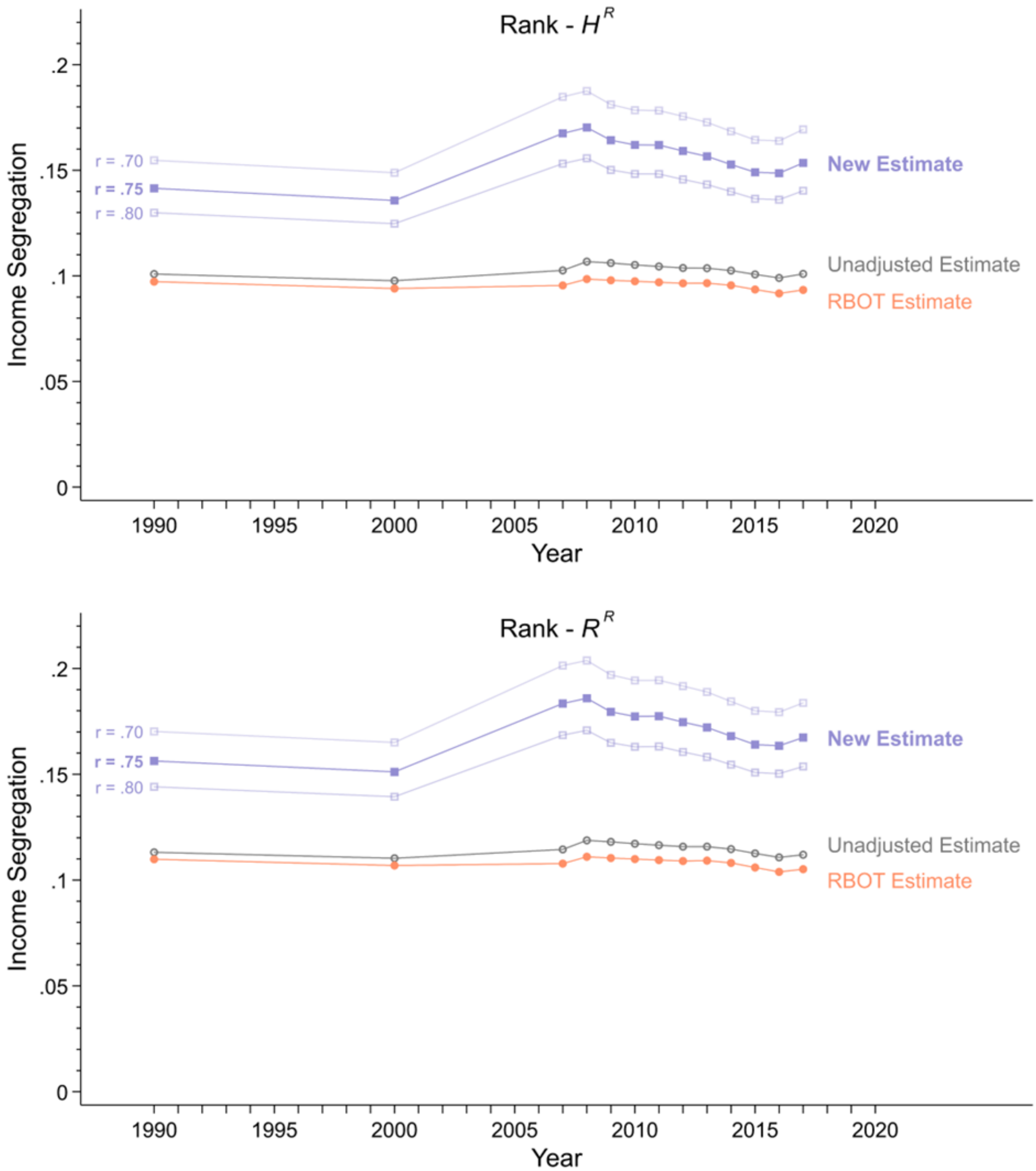

Fig. 4 Household income segregation trend estimates in Rank-H^R (top) and Rank-R^R (bottom) with ACS 5-year pooled estimates plotted at middle year, by estimate type, for the 116 metro areas analyzed in Reardon et al. (2018), where RBOT estimates apply the Reardon et al. (2018) finite sampling bias correction.

Appendix A. Estimating Rank-Order Segregation.

Given r, Δ^T, and σ, we can estimate S in both annual and pooled data as described in Equation (11). A few more steps are required to estimate the rank-order segregation measures more commonly used in income segregation analyses.

We use HETOP to estimate observed neighborhood means and variances in log-dollars. We then adjust the observed neighborhood variances and means to estimate the true neighborhood variances, W_j^s, and means, μ_j^s, in standardized units. Here, we draw from Reardon, Kalogrides, and Ho (2017), who provide equations for disattenuating the mean and variance of a normal variable measured with classical error.

Given a HETOP-estimated sample variance, $\widehat{W_j}$, and sample mean, $\widehat{\mu_j}$, for neighborhood j, the neighborhood's observed variance and mean in standardized units are

$$W_j^s = \frac{\widehat{W_j}}{\widehat{W} + \hat{B} - E}, \tag{A1}$$

and

$$\mu_j^s = \frac{\widehat{\mu_j} - \hat{\mu}}{\sqrt{\widehat{W} + \hat{B} - E}}, \tag{A2}$$

where $\hat{B}$ is the variance of HETOP-estimated neighborhood means, and E is the error variance in HETOP estimates of the neighborhood means.

The neighborhood's true variance and mean in standardized units are

$$W_j^s = \frac{1}{r}\left(\frac{n}{n-1}\right)\left(\widehat{W_j} - \frac{T-1}{T}\sigma - \Delta^T\right)\left(\frac{n+1}{n}\widehat{W} + \hat{B} - E - \Delta^T\right)^{-1} + 1 - \frac{1}{r}, \tag{A3}$$

and

$$\mu_j^s = \left(\widehat{\mu_J} - \hat{\mu}\right)\sqrt{\frac{1}{r}\left(\frac{n+1}{n}\widehat{W} + \hat{B} - E - \Delta^T\right)^{-1}}\sqrt{1 - \frac{\widehat{W} - \frac{T-1}{T}n\sigma - \Delta^T}{n-1}\left(\hat{B} - E + \frac{\widehat{W}}{n}\right)^{-1}}, \tag{A4}$$

where $\sigma = 0$ and $\Delta^T = 0$ when using annual data.

Given W_j^s and μ_j^s, the estimated number of households in neighborhood j with an income falling within population decile d is

$$N_j^d = N_j\left[\Phi\left(\frac{\Phi^{-1}\left(\frac{d}{10}\right) - \mu_j^s}{\sqrt{W_j^s}}\right) - \Phi\left(\frac{\Phi^{-1}\left(\frac{d-1}{10}\right) - \mu_j^s}{\sqrt{W_j^s}}\right)\right]. \tag{A5}$$

Rank-order income segregation between neighborhoods can be computed using Stata's -rankseg- command treating N_j^1 through N_j^{10} as neighborhood population counts.

Equations (A3) and (A4) were validated as part of the validation simulation by computing segregation as $S^s = \frac{Var\left(\mu_j^s\right)}{Var\left(\mu_j^s\right) + \frac{1}{J}\Sigma_j W_j^s}$. This approach produced results similar to our main approach, as shown in Fig. A1 below.

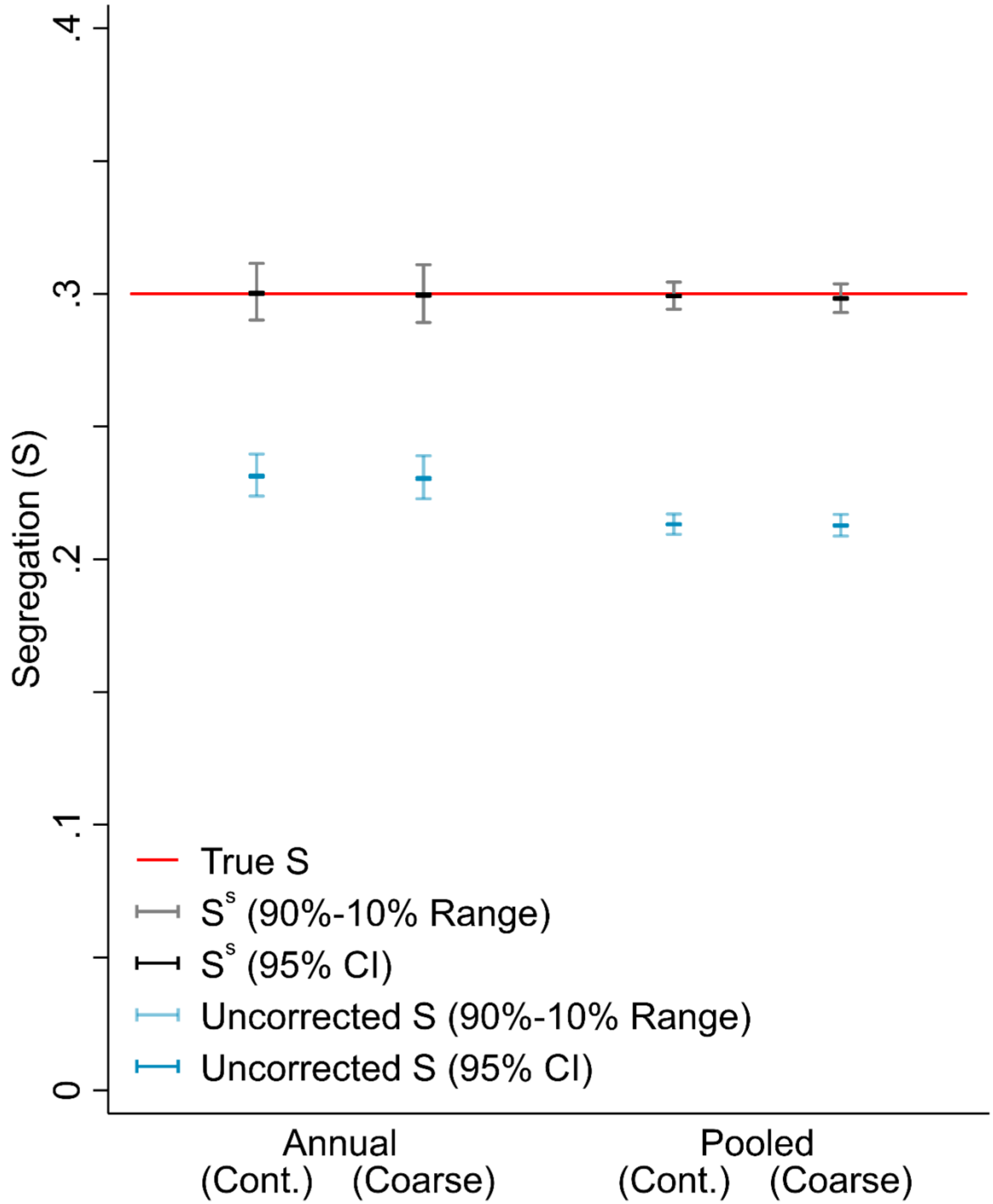

Fig. A1 Bias reduction using S^s, by data collection condition, in simulations where the 90-10% range is a range of estimates (indicating imprecision) and the 95% CI is the confidence interval of the mean estimate (indicating bias).

Appendix B. Relaxing the Large-J and Large Population Assumptions.

The equations in Table 1, Equation (11), and Appendix A estimate population variances as variances from a theoretical distribution, as though J and the full metro area population are infinitely large. This assumption makes the equations above easier to follow and are inconsequential in many applications. However, they will lead to bias when (1) J or the full metro area population are small and (2) segregation is estimated from a sample rather than the full population. When these conditions are met, the estimated true values of W, B, V, and S should be adjusted.

Let N be the neighborhood population and JN be the metro area population. In expectation, the adjusted estimates (denoted $\square^{\circ}$) are:

$$W^{\circ} = \frac{N-1}{N} W, \tag{B1}$$

$$B^{\circ} = \frac{JN-1}{JN} B + \frac{J-1}{JN} W, \tag{B2}$$

$$V^{\circ} = \frac{JN-1}{JN} V, \tag{B3}$$

and

$$S^{\circ} = S + \frac{J-1}{JN-1}(1-S). \tag{B4}$$

In expectation, the finite population values of the standardized within-tract variance and tract means discussed in Appendix A are:

$$W_j^{s\circ} = \frac{N-1}{N\sqrt{1-\frac{1}{JN}}} W_j^s$$

(B5)

and

$$\mu_j^{s^\circ} = \mu_j^s \left[\frac{N-1}{N} \left(1 - \frac{1}{JN}\right)^{-\frac{1}{2}} + \left(\left(1 - \frac{1}{JN}\right)^{\frac{1}{2}} - \frac{N-1}{N} \left(1 - \frac{1}{JN}\right)^{-\frac{1}{2}} \right) \frac{1}{S} \right].$$

(B6)

Appendix C. Assessing the Accuracy of $\boldsymbol{\Delta^T}$ and $\boldsymbol{\sigma}$ Estimates.

Fig. C1 and Fig. C2 present the estimates of σ and Δ^T, respectively, in the simulations validating the bias-corrected measure. To evaluate σ we use the estimates using the first 6 years and the last 6 years of the simulated data to avoid estimates for overlapping periods which would complicate estimating standard errors. We slightly underestimate σ. In continuous data, the average of the estimates is $.02991$ with a 95 percent confidence interval of $(.02985, .02997)$ while in coarsened data the average of the estimates is $.02869$ with a 95 percent confidence interval of $(.02865, .02873)$.

In Fig. C2, we have adjusted the benchmark for Δ^T to account for sampling variability due to simulating $K = 12$ instead of infinite years of data. Due to the finite number of years, the true value of Δ^T differs from its expected value of $\frac{T-1}{T}\Delta$. We compute the true value of $\frac{T-1}{T}\Delta$ in the simulated data as $Var(Y_t) - \frac{\sigma}{J}$ where Y_t is the true annual metro mean and $\frac{\sigma}{J}$ is the expected variance in the annual metro means due to v_{jt} having a non-zero annual mean in the metro. In our simulations, this value is $.000887$, which we use as our benchmark for Δ^T. We estimate Δ^T without bias in both conditions.

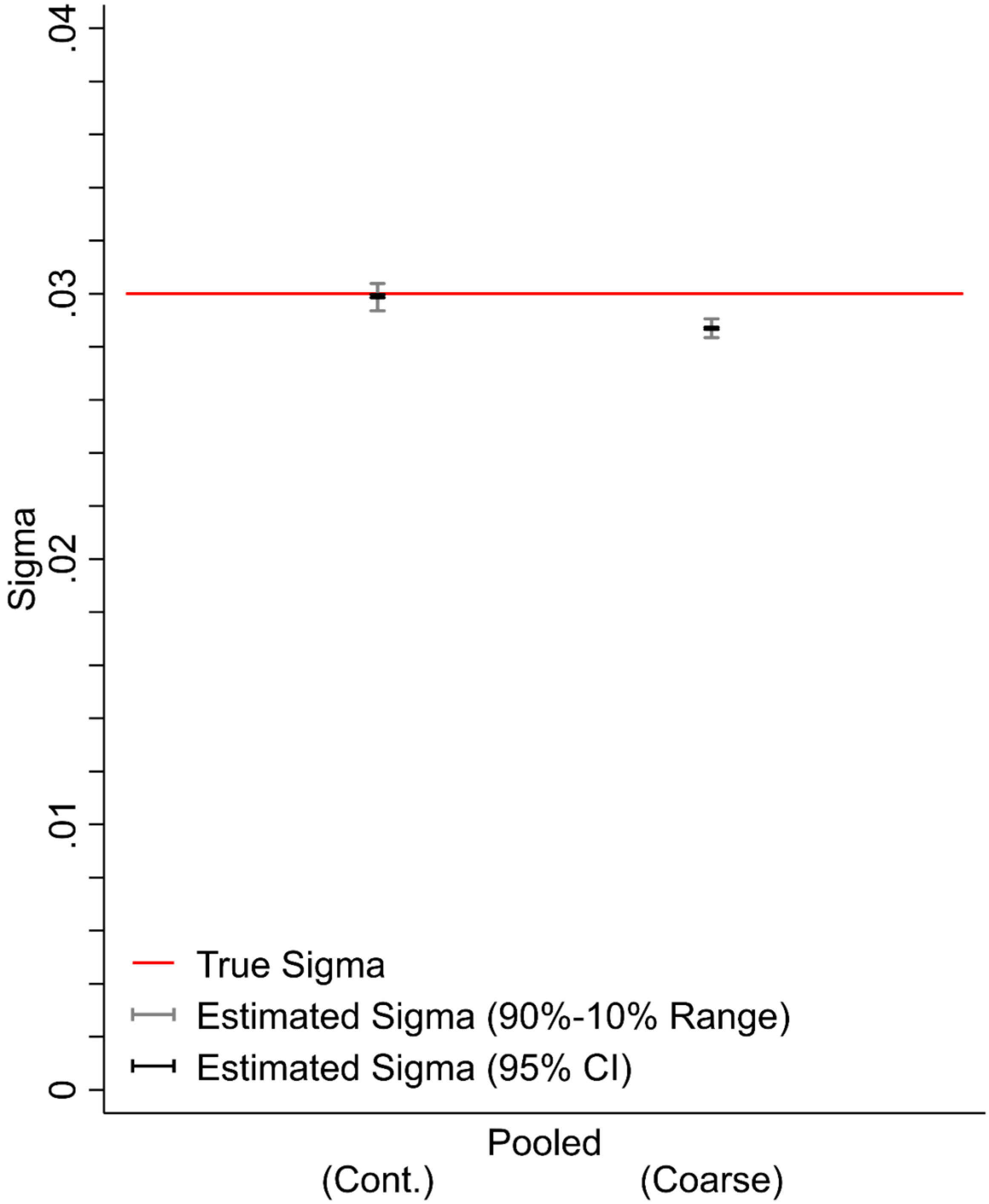

Fig. C1 Estimates of σ, by data collection condition, in simulations.

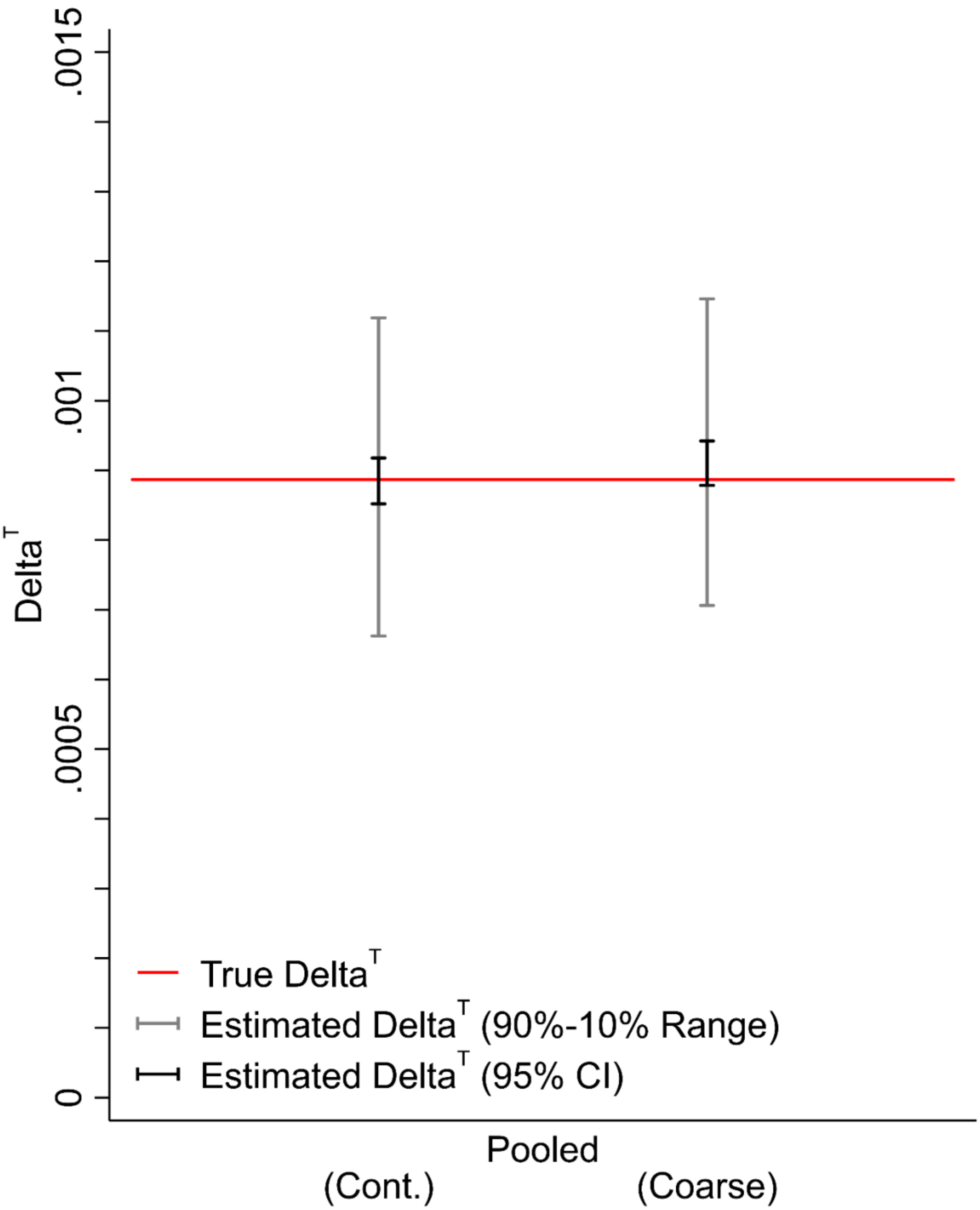

Fig. C2 Estimates of Δ^T, by data collection condition, in simulations where the True Δ^T benchmark accounts for sampling variability in the finite-K simulations.

Appendix D. Reliability of Self-Reported Household Income.

The highest-quality estimates of the reliability of self-reported income make use of administrative data, typically treating administrative data as truth. It may be more accurate to think of administrative data as highly reliable but not perfect, given the possibility of imperfect data linkage, filing mistakes, employer processing errors, and automated read-in malfunctions (Abowd and Stinson 2013; Kapteyn and Ypma 2007). Abowd and Stinson (2013) and Stinson (2002) take this approach, estimating true income using both self-reported and administrative data under various assumptions about how informative each data source is.

Income reliability studies that do not use administrative data compare self-reported income at one time to self-reported income at another or use self-reported income to estimate true income (e.g., the three-wave-simplex design). These approaches are more likely to overestimate reliability; if individuals err similarly when reporting their income, some of the noise will be mistaken for signal.

For the purposes of this study, we are interested in the reliability of household total income, the measure used in most income segregation studies. This presents a challenge; among the studies we found, all of the highest quality income reliability studies – those that make use of administrative data – focus on personal earnings, which excludes other income-earners in the household and other sources of income.

Moore, Stinson, and Welniak, Jr.'s (2000) review of sources of error in self-reported income provides some guidance for how to proceed. They highlight that measurement error varies by types of income. They report that the reliability of self-reported personal earnings is typically estimated at around .8, a finding that comports with the data in Table D1, which describes studies estimating income reliability. Meanwhile, income from transfers is in the .4 to .6 range and the reliability of income from assets is even worse (Moore, Stinson, and Welniak, Jr.'s 2000). Given that earnings make up most income in most households and under the assumption that errors in reports of different income sources are not negatively correlated, this suggests self-reported total household income is slightly less reliable than self-reported personal earnings. However, it is unclear by how much. We use a benchmark of .75 and provide estimates in which household income reliability is alternatively .8, similar to personal earnings, and .7, in the case that household income is substantially less reliable than personal earnings.

Note that this study is focused on annual income; studies considering the reliability of self-reported annual income as a proxy for long-run income (e.g., lifetime earnings) have produced far lower reliability estimates (Brady et al. 2018; Haider and Solon 2006; Hyslop 2001; Mazumder 2001; Rothstein and Wozny 2013).

Table B1 Literature review of estimates of the reliability of reported annual income.

Income Type	Source	Data	Measure Type	Target Type	Reliability	Design Details
Personal Total Income	Alwin 2007	NES 1990	Self-Report	Estimate	.95	Three-wave simplex design
	US Bureau of the Census 1968	CPS Reinterview 1961-66	Self-Report	Self-Report	.87	Test-Retest design
	Siegel and Hodge 1968	Census & March CPS 1960	Self-Report	Self-Report	.85	Test-Retest design
Personal Earnings	Moore et al. 2000	Literature Review	Self-Report	Administrative	About .8	Individual studies included in this review are excluded from this table
	Gottschalk and Huynh 2010	SIPP & SSA DER 1990-99	Self-Report	Administrative	.67	
	Abowd and Stinson 2013	SIPP & SSA DER 1990-99	Self-Report	Administrative	.7-.78	Estimate varies by data imputation strategy; Reliability appears to be increasing over time
		SIPP & SSA DER 1990-99	Self-Report	Estimate	.76-.82	Assume administrative data is much more informative than self-reported data; Estimate varies by data imputation strategy
		SIPP & SSA DER 1990-99	Self-Report	Estimate	.93-.94	Assume administrative and self-reported data are equally informative; Estimate varies by data imputation strategy

Table B1 (Continued) Literature review of estimates of the reliability of reported annual income.

Stinson 2002	SIPP & SSA DER 1996	Self-Report	Estimate	.82	Assumes administrative and self-reported data are equally informative
Alwin 2007	ACL 1986-1979-1994	Self-Report	Estimate	.96	Three-wave simplex design
US Bureau of the Census 1968	CPS Reinterview Program 1961-66	Self-Report	Self-Report	.7	Test-Retest design
Bielby, Hauser, and Featherman 1977	CPS & OCG March 1973	Self-Report	Self-Report	.9	Test-Retest design
Withey 1954	Repeated Interviews 1948-49	Self-Report	Self-Report	.84	Test-Retest design; Reporting 1947 earnings
Abowd and Stinson 2013	SIPP & SSA DER 1990-99	Administrative	Estimate	.94-.95	Assume administrative and self-reported data are equally informative; Estimate varies by data imputation strategy
	SIPP & SSA DER 1990-99		Estimate	1	Assume administrative data is much more informative than self-reported data
			Self-report	.74-.8	Estimate varies by data imputation strategy

Table B1 (Continued) Literature review of estimates of the reliability of reported annual income.

	Stinson 2002	SIPP & SSA DER 1996	Administrative	Estimate	.79	Assumes administrative and self-reported data are equally informative
Personal Transfers	Moore et al. 2000	Literature Review	Self-Report	Administrative	.4-.6	
Personal Asset Income	Moore et al. 2000	Literature Review	Self-Report	Administrative	< .4	
Household Total Income	Alwin 2007	NES 1990	Self-Report	Estimate	.88	Three-wave simplex design
		NES 1950			.9	Three-wave simplex design
		NES 1970			.87	Three-wave simplex design
		NES 1990			.92	Three-wave simplex design
		SAFMO 1976-79-84-92			.77	Detroit Women; Three-wave simplex design

Notes: "Measure Type" refers to the type of source producing the measure for which the reliability is estimated while "Target Type" refers to the type of source producing the data that proxies for true income.

Appendix E. Adjusting for Imprecision in $\boldsymbol{\sigma}$.

In practice, the σ estimates are imprecise, which can inflate how much σ appears to vary over time and across metros. To improve efficiency, we randomly split each sample into two halves in each pair of ACS pools used to estimate σ.[8] For our purposes, this means that a "half-sample" consists of half of the tracts in a metro area over a six-year period. We estimate σ in each half-sample.

We use a 3-level HLM model of half-samples s within metro-periods t within metros m:

$$\sigma_{stm} = v_{00m} + u_{0tm} + \sum_{t=1}^{T} \gamma_{0t0} * I(Period_{mt} = t) + e_{stm}$$

$$e_{stm} \sim N(0, \varphi), u_{0tm} \sim N(0, \tau_{00}), v_{00m} \sim N(0, \tau_{000}),$$

(E1)

where v_{00m} is a metro-specific intercept, u_{0tm} is a metro-period-specific average deviation from the metro intercept, and $\sum_{t=1}^{T} \gamma_{0t0} * I(Period_{mt} = t)$ is the non-parametric, secular time trend.

We estimate σ in each metro-period as $\sigma_{tm} = v_{00m} + u_{0tm}$ using Empirical-Bayes shrunken estimates. This model leverages variation across half-sample σ estimates to model imprecision in the full-sample σ estimates to create shrunken estimates that account for imprecision-inflated variation across metros and over time.

[8] Some tracts are only observed in one of two adjacent ACS datasets. We drop them assuming they are missing at random. In all but the 2005-2009 and 2006-2010 pair of ACS pools (in which 22 percent of tracts are dropped), no more than 0.11 percent of tracts are dropped.

Appendix F. $\boldsymbol{\Delta^T}$ and $\boldsymbol{\sigma}$ Estimates.

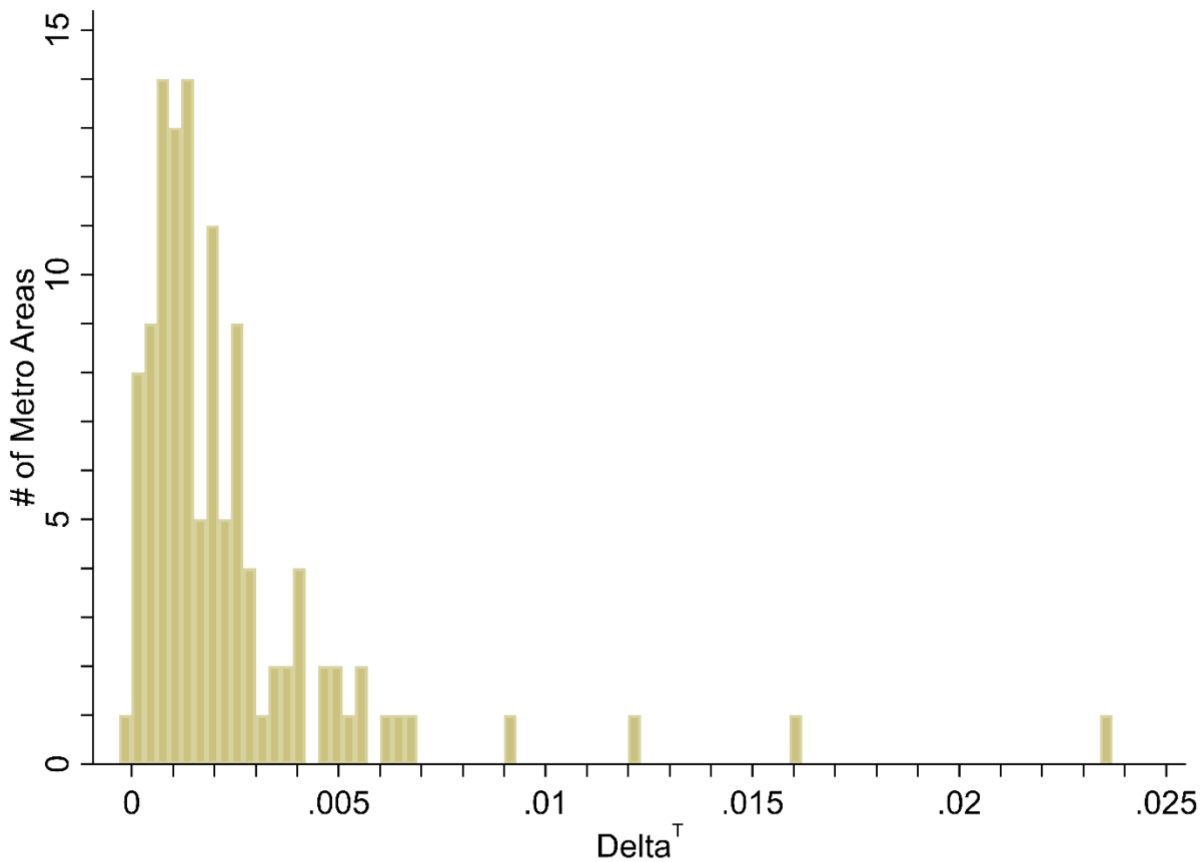

Fig. F1 Frequency distribution of Δ^T estimates.

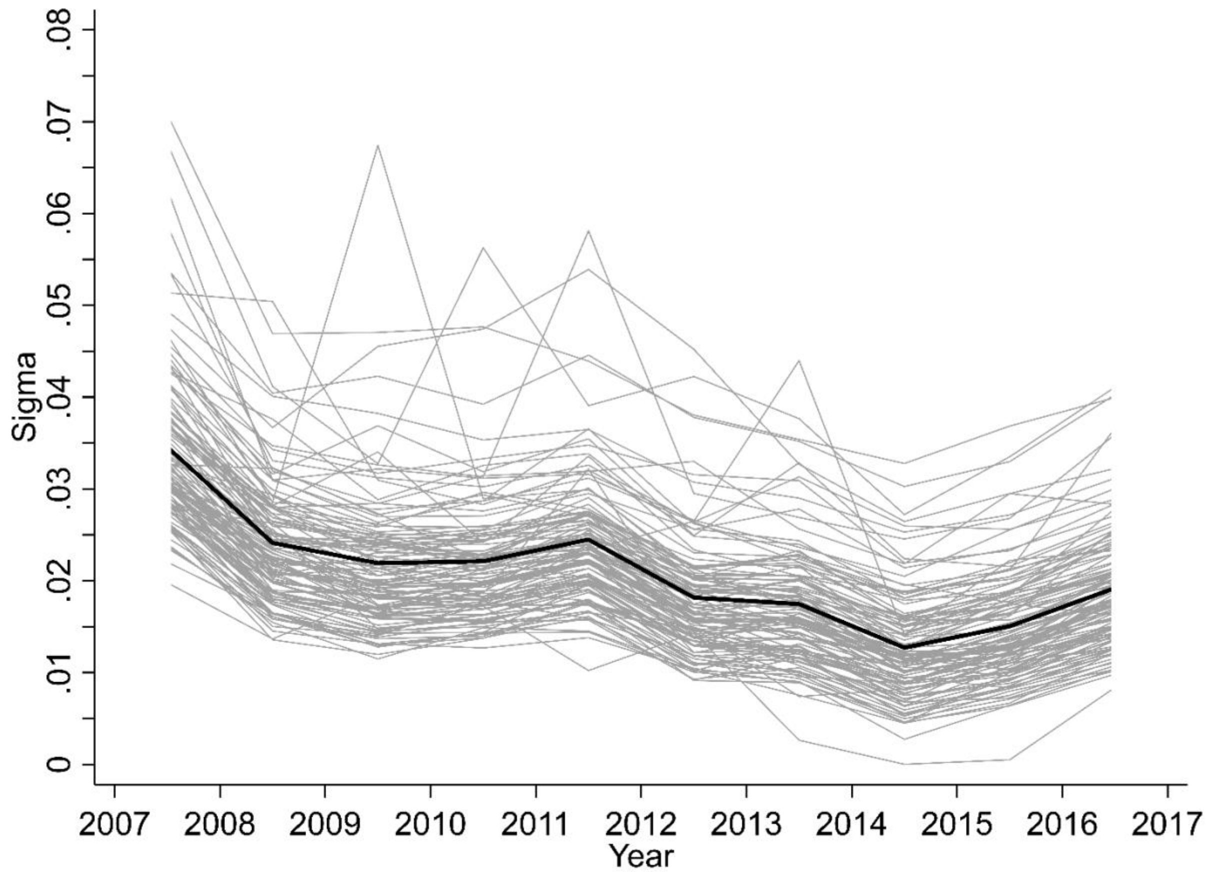

Fig. F2 Trend in σ estimates, by metro area and on average (bold).