Differential Privacy and Racial Residential Segregation

David Van Riper, Tracy Kugler PhD, Jonathan Schroeder PhD, Steven Ruggles PhD

10/28/2020

DRAFT - please do not cite without permission. Contact David Van Riper (vanriper@umn.edu (mailto:vanriper@umn.edu)) for permission.

Abstract

For the 2020 Decennial Census, the US Census Bureau is adopting a new disclosure avoidance technique to provide stronger confidentiality protections to respondents. This technique, based on differential privacy, injects noise into nearly all published statistics, including race/ethnicity counts. These race/ethnicity counts are frequently used to measure racial residential segregation in the United States. This paper examines whether the new disclosure avoidance technique enhances or mutes common measures of racial residential segregation, including both aspatial and spatial measures. We compute multiple segregation indices from four versions of the 2010 decennial census data—the Summary File 1 data produce via traditional disclosure avoidance techniques and three differentially private versions. The index of dissimilarity, probably the most commonly used segregation measure, performs poorly when based on differentially private data. Other segregation indices, including multigroup entropy and its spatial counterpart, change less when based on differentially private data.

Introduction

Residential segregation is a form of structural racism, imposed and enforced by racist policies and systems that structure our world. Racial residential segregation has been a topic of study for scientists from a broad range of disciplines, including demography, sociology, economics, geography, history, law, and public health. In addition to work devoted to measuring segregation (Duncan & Duncan, 1955; Grigoryeva & Ruef, 2015; Logan & Parman, 2017; Massey & Denton, 1988; Morrill, 1991; Reardon et al., 2008; Reardon & O'Sullivan, 2004; Yao et al., 2019), scientists have examined the relationship between segregation and economic (Cutler & Glaeser, 1997; Iceland & Wilkes, 2006), educational (Quillian, 2014), and health outcomes (Logan & Parman, 2018; Mehra et al., 2017; Mendez et al., 2014; Williams & Collins, 2001).

These studies rely on segregation measures usually computed from decennial census data. The measures use counts of persons who identify with particular races or ethnicities to quantify segregation for cities or metropolitan areas. Shifting to a disclosure avoidance technique that injects noise into counts may adversely impact segregation metrics, given a false sense of whether segregation is increasing or decreasing.

To better understand the potential impacts the new disclosure avoidance system may have on the measurement racial residential segregation, we will compare a variety of segregation measures computed from four datasets based on the 2010 Decennial Census—Summary File 1 (SF1), the 2010 Demonstration Data Product (DDP) (Bureau, 2019), the v20200527 dataset (Van Riper et al., 2020a), and the v20200917 (Van Riper et al., 2020b) dataset. The latter three datasets were produced by different versions of the Bureau's new disclosure avoidance system, and SF1 was produced using household swapping for disclosure avoidance. These comparisons will help us understand the potential issues scientists may face when using differentially private data to study residential segregation.

Differential Privacy and the Decennial Census

In the summer of 2018, the US Census Bureau announced a transition from traditional disclosure avoidance techniques such as cell suppression or household swapping[1] to a technique based on differential privacy. Data from the 2020 Decennial Census of Population and Housing will be produced using the new technique (Abowd, 2018), marking “a sea change for the way that official statistics are produced and published” (Garfinkel et al., 2018).

The change to a new disclosure avoidance technique was driven by a reconstruction and re-identification attack executed by Bureau scientists. Using publicly available census tract and block tabulations from the 2010 Decennial Census, the Bureau reconstructed (Dinur & Nissim, 2003) microdata and linked the microdata to a commercial database that contained names. Approximately 45% (139 million) of the microdata records matched a record in the commercial database. The Bureau then merged the matched reconstructed records to the confidential microdata on all attributes—census block identifier, age, sex, name, race, Hispanic ethnicity. This merge yielded 52 million matches, representing 17% of the US resident population in 2010 (Leclerc, 2019).

That they could successfully re-identify 17% of the US population using published tabulations and commercial database pushed the Bureau to adopt a new disclosure avoidance technique that provides stronger confidentiality protections than prior methods. The new technique is grounded in differential privacy, which provides a mathematically provable lower bound on the amount of private information leaked when statistics are published (Dwork et al., 2006). Algorithms typically satisfy differential privacy’s requirements by injecting noise drawn at random from a statistical distribution into the cells of a cross-tabulation. The type and shape of the statistical distribution is determined by the data producer before data production begins.[2]

Policy Decisions

All disclosure avoidance techniques or algorithms include parameters that control the amount of noise, suppression or swapping applied to the data, and these parameters influence the accuracy and utility of the output data. It is critically important that data users understand the significance of the parameters. This section discusses the Top Down Algorithm’s parameters and the values used to generate the DDP, v20200527, and v20200917 datasets.

Global Privacy Loss Budget

The global privacy loss budget (PLB) (ϵ) controls the trade-off between the privacy afforded to Census respondents and the accuracy of the published data. Values for ϵ range from 0 to infinity, with 0 representing perfect privacy/no accuracy and infinity representing no privacy/perfect accuracy. After establishing the global PLB, it is then allocated to geographic levels and queries. Geographic levels or queries that receive larger fractions of the PLB will be more accurate than those that receive smaller fractions of the PLB.

The DDP, v20200527, and v20200917 all had the same PLB of 4.0 for person-based tables.[3] Combination of geographic levels and queries receive allocations of the PLB, and those allocations ultimately control the magnitude of the noise injected into counts.

Geographic Levels

Geographic levels define the rows of cross-tabulations, and each row represents a geographic unit within a geographic level (e.g., Minnesota is a geographic unit in the State geographic level). For the DDP/v20200527 and v20200917 datasets, seven and six geographic levels, respectively, received direct allocations of the PLB (Table 1). The census tract group[4] received no allocation in the v20200917 and was dropped from the hierarchy. In the DDP and v20200527 datasets, nation and state received a larger allocation, meaning less noise will be injected into their counts compared to counties, tracts…blocks. In the v20200917, nation received a larger allocation. Compared to the DDP and v20200527 datasets, however, the county, tract, block group, and block levels received larger allocation (0.16 vs. 0.12) in v20200917.

Table 1. Fractional Privacy Loss Budget Allocations to Geographic Levels for DDP/v20200527 and v20200917[1]

Geographic level	Allocation	
	DDP/v20200527	v20200917
Nation	0.20	0.20
State	0.20	0.16
County	0.12	0.16
Tract Group	0.12	NA[2]
Tract	0.12	0.16
Block Group	0.12	0.16
Block	0.12	0.16

[1] DDP, v20200527, and v20200917 are 2010 demonstration products produced by the US Census Bureau using three versions of their new disclosure avoidance system.

[2] The tract group geographic level was provided with a direct PLB allocation in the v20200917 dataset.

Queries

Queries define the rows of a cross-tabulation, and they are essentially combinations of demographic variables (i.e., age by sex). As with geographic levels, the PLB is allocated to these queries. The disclosure avoidance system uses two types of queries. The “Detailed” queries consist of all unique combinations of variables (e.g., single year of age by sex by race by Hispanic ethnicity by household/group quarters status), and “DP” queries are specific combinations of variables (e.g., voting age * race by Hispanic ethnicity). Both query types play a key role in the disclosure avoidance system. “Detailed” queries allow the Bureau to reconstruct the underlying microdata, and the “DP” queries allow policymakers to increase the accuracy of specific statistics in the published data.

Queries defined in the disclosure avoidance system do not have a one-to-one relationship with tables published by the Bureau. The queries are used in the noise injection and optimization processes, and the published tables are created from microdata created by those processes. Categories in the published data differ from those used in the queries.

The Census Bureau used three sets of queries for the person tables in the DDP, v20200527 and v20200917 datasets. Query descriptions and PLB allocations are shown in Table 2. The *voting age * Hispanic * race * citizenship* query, which is used for legislative redistricting, received the largest allocation (50%) in the DDP, and the *total population* query received the largest allocation (30%) in the v20200527 and v20200917 datasets.

Invariants and Constraints

Invariants and constraints play key roles in the disclosure avoidance system, particularly in the post-processing routines applied to the noisy counts. Invariants are counts computed directly from the confidential data and are not subject to noise injection. Constraints control the types and ranges of values in the published statistics.

The Bureau has not yet selected invariants for the 2020 Decennial Census, but it did set four invariants for the DDP and v20200527 datasets and five invariants for the v20200917 dataset. Total population is invariant at the state level, and total housing units, total group quarters facilities, and total group quarters facilities by type are invariant at the census block level. The v20200917 included the same four invariants as the DDP, and it include a state-level total population living on American Indian/Alaska Native area invariant. The 2010 Summary File 1 dataset had 6 census block invariants—total population, voting age population, total group quarters facilities, total group quarters facilities by type, total housing units, and occupied housing units.

Constraints are the set of rules that the data produced by the disclosure avoidance system must follow. For all three datasets, constraints included non-negativity, integer, and hierarchical consistency. The non-negativity and integer constraints require that all published statistics are positive integer values. The hierarchical consistency constraint imposes consistency among geographic and category hierarchies. For geographic hierarchies, counts for child units must sum to the counts of their parent unit. For category hierarchies, counts of child categories must sum to the counts of their parent category.

Invariants and constraints are not required by the disclosure avoidance system; instead, they are imposed to satisfy the expectations of policymakers and data users. The state-level total population invariant was established to avoid debates and litigation over Congressional reapportionment. The non-negativity and integer constraints are imposed to avoid negative counts or non-integer values for geographic units, since we only count whole persons in the census. These constraints avoid such illogical values. Consistency constraints guarantee that if you sum the male and female totals from any published data table, you will obtain the same sum.

Disclosure avoidance system

The disclosure avoidance system that generated the DDP, v20200527, and v20200917 consists of three steps: generating counts from the CEF, injecting noise, and post-processing to satisfy invariants and constraints. The first two steps were the same for both datasets, but the post-processing step differed substantially between the DDP and the v20200527/v20200917 and is described below.

Generating Counts

The first step produces counts from the CEF. The disclosure avoidance system consumes the CEF, the queries, and the geographic levels and creates a set of cross-tabulations (or histograms) - one for each combination of query and geographic level. The cells in each cross-tabulation contain the counts of a particular set of categories (e.g., 0-4 year old males) for a given geographic unit. The number of cells in these cross-tabulations may be massive, particularly at the census block level, and the counts in the cells may be small or even zero.

Noise Injection

The second step injects noise into the counts generated in the prior step. These “noisy counts” are, by definition, differentially private, but they may not satisfy the specified invariants or constraints. The noise injection step is implemented with three sub-steps.

Compute ϵ for Each Geographic Level * Query Combination

The value of ϵ is computed for each combination by multiplying the PLB, geographic level fraction, and query fraction:

$$\epsilon = PLB * GeogLevel_{fraction} * Query_{fraction}$$

For the three datasets, the PLB was 4.0, the geographic level fractions are listed in Table 1, and the query fractions are shown in the *Allocation* column of Table 2.

Compute the Scale Parameter for the Statistical Distribution

Noise-injection values are generated by randomly drawing a value from a statistical distribution. The DDP and v20200527 used a double geometric distribution, which is essentially the discrete version of the Laplace distribution. The v20200917 dataset used a discrete Gaussian distribution[5]

The shape of the distribution is controlled by the scale parameter calculated using the following formula:

$$s = \frac{2}{\epsilon}$$

ϵ is the geographic level * query value computed in the previous sub-step. The numerator is the sensitivity of the query, which is always 2 for histograms.[6]

Scale parameters for the three dataset are shown in Table 2. For the DDP and v20200527 datasets, the nation column represent the nation and state levels. The block column represents the county, census tract group, census tract, census block group, and census block levels. For the v20200917 dataset, the nation column represents the nation scale parameter, and the block column represents the state, county, tract, block group, and block geographic levels.

The standard deviation for the DDP and v20200527 scale factors can be computed as follows:

$$\sigma = \sqrt{2 * s^2}$$

For the Gaussian-based v20200917 dataset, the standard deviation is same as the scale factor.

Table 2. Fractional Privacy Loss Budget Allocations, Scale Factors, and Standard Deviations for Queries

		Nation		Block	
Query	Allocation	Scale	SD	Scale	SD
DDP[1]					
Voting age * Hispanic * Race * Citizenship	0.500	5.0	7.1	8.3	11.8
Household/Group quarters	0.200	12.5	17.7	20.8	29.5
Detailed	0.100	25.0	35.4	41.7	58.9

[1] DDP: 2010 Demonstration Data Product produced by the Census Bureau using one version of the new disclosure avoidance system.

[2] v20200527: Second differentially private demonstration dataset produced by the Census Bureau using a different version of the new disclosure avoidance system. This dataset

[3] v20200917: Third differentially private demonstration dataset produced by the Census Bureau using a different version of the new disclosure avoidance system. This dataset only supports tables included in the PL94-171 redistricting file.

Table 2. Fractional Privacy Loss Budget Allocations, Scale Factors, and Standard Deviations for Queries

Sex * Age (single year of age)	0.050	50.0	70.7	83.3	117.9
Sex * Age (4-year age bins)	0.050	50.0	70.7	83.3	117.9
Sex * Age (16-year age bins)	0.050	50.0	70.7	83.3	117.9
Sex * Age (64-year age bins)	0.050	50.0	70.7	83.3	117.9
v20200527[2]					
Total population	0.300	8.3	11.8	13.9	19.6
Voting age * Hispanic * Race	0.290	8.6	12.2	14.4	20.3
Age * Sex * Hispanic * Race	0.250	10.0	14.1	16.7	23.6
Household/Group quarters	0.150	16.7	23.6	27.8	39.3
Detailed	0.010	250.0	353.6	416.7	589.3
v20200917[3]					
Total Population	0.300	8.3	8.3	10.4	10.4
Number of races	0.100	25.0	25.0	31.2	31.2
Race	0.100	25.0	25.0	31.2	31.2
Hispanic * Number of races	0.100	25.0	25.0	31.2	31.2
Hispanic * Race	0.025	100.0	100.0	125.0	125.0
Voting age * Number of races	0.100	25.0	25.0	31.2	31.2
Voting age * Race	0.025	100.0	100.0	125.0	125.0

[1] DDP: 2010 Demonstration Data Product produced by the Census Bureau using one version of the new disclosure avoidance system.

[2] v20200527: Second differentially private demonstration dataset produced by the Census Bureau using a different version of the new disclosure avoidance system. This dataset

[3] v20200917: Third differentially private demonstration dataset produced by the Census Bureau using a different version of the new disclosure avoidance system. This dataset only supports tables included in the PL94-171 redistricting file.

Table 2. Fractional Privacy Loss Budget Allocations, Scale Factors, and Standard Deviations for Queries

Voting age * Number of races * Race	0.025	100.0	100.0	125.0	125.0
Voting age * Race * Hispanic	0.025	100.0	100.0	125.0	125.0
Institution type	0.100	25.0	25.0	31.2	31.2
Group quarters type	0.075	33.3	33.3	41.7	41.7
Detailed	0.025	100.0	100.0	125.0	125.0

[1] DDP: 2010 Demonstration Data Product produced by the Census Bureau using one version of the new disclosure avoidance system.

[2] v20200527: Second differentially private demonstration dataset produced by the Census Bureau using a different version of the new disclosure avoidance system. This dataset

[3] v20200917: Third differentially private demonstration dataset produced by the Census Bureau using a different version of the new disclosure avoidance system. This dataset only supports tables included in the PL94-171 redistricting file.

Generate and Inject Random Noise in Each Cell

For each cell in a particular combination of a *geographic level * query*, a random value is drawn from the appropriate distribution and added to the confidential cell value. This sum is the noisy count. If we consider the *county* by *voting age * Hispanic * race* for the v20200527 dataset, we will take 811,692 random values (3,221 county * 252 unique categories [2 voting age * 2 Hispanic * 63 race] in the query) from a distribution with a scale parameter of 14.4 and an standard deviation of 20.3. Thus, 75% of those random values will fall within the standard deviation and 25% will fall outside.

Post-processing

The prior step yields a set of noisy, differentially private, histograms—one for each *geographic level * query* combination. These histograms are integer-valued but do not satisfy invariants, constraints, and the required tabulation software format. Additionally, the set of queries and geographic levels used for noise injection does not match the set of cross-tabulations and geographic levels desired for publication. In order to produce the final dataset that satisfy all requirements, the Census Bureau conducts a series of optimization steps that ultimately produce microdata (each record represents a person), which can then be tabulated for publication.

Different post-processing routines were used to generate the DDP and the v20200527/v20200719 datasets. Both routines follow the same general structure in that they start at the national level and work down the geographic hierarchy, successively generating data for finer geographic levels. However, the optimization step for each geographic level differs.

For the DDP, the algorithm minimized the differences between the noisy detailed histogram and the noisy DP queries, along with satisfying invariants, non-negativity, and consistency, in a single optimization step (Figure 1, Panel B). The algorithm first optimized the nation-level queries and then moved down the geographic hierarchy until it produced the optimized block-level queries (Figure 1, Panel A). Those queries were then converted to microdata that were tabulated for publication. Optimizing all queries at one time and enforcing all required invariants and constraints yielded inaccurate counts, particularly for areas with small population counts and sub-populations with small counts (Hotz & Salvo, 2020).

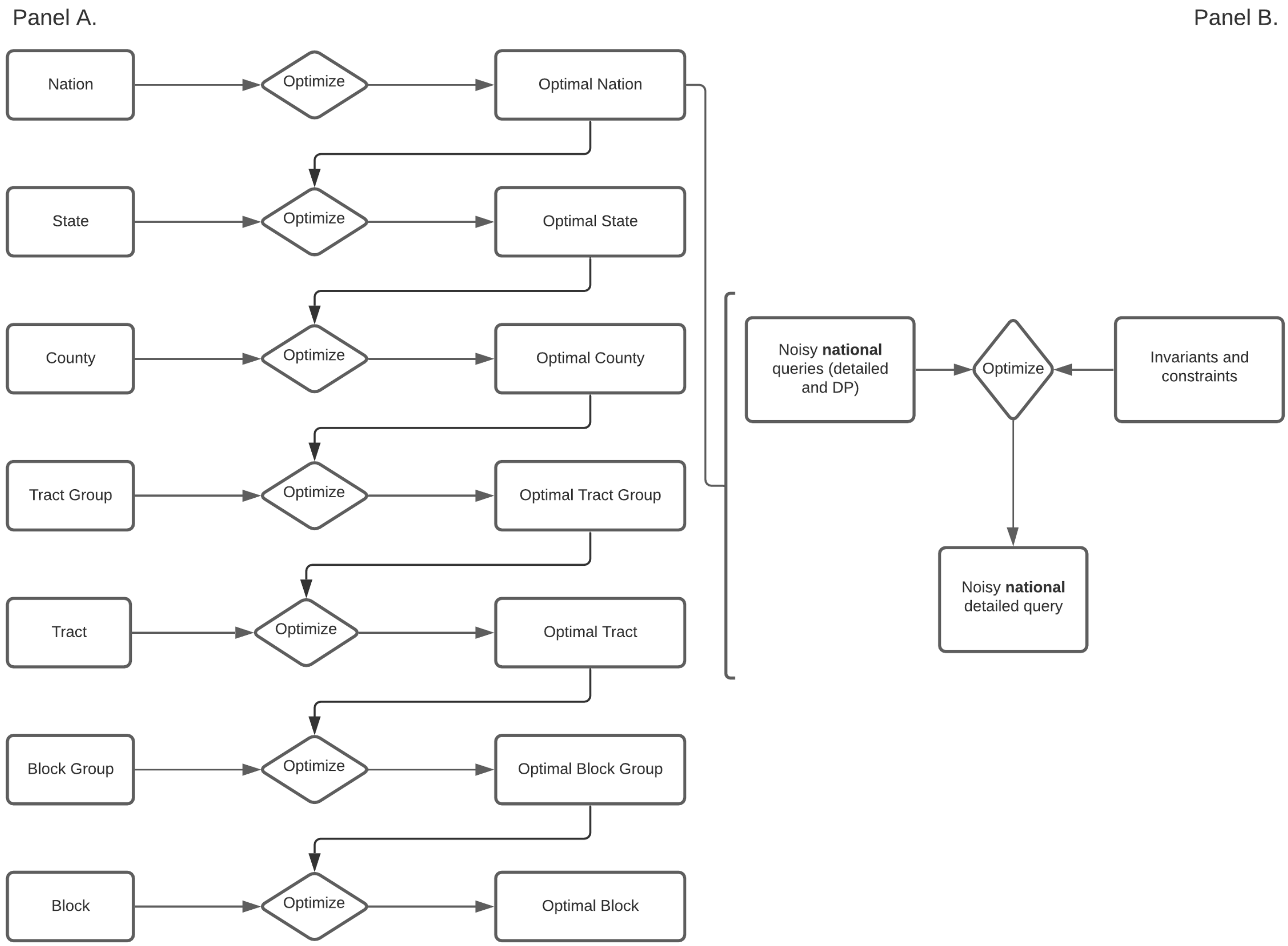

Figure 1. Top Down Algorithm used for the DDP dataset. Panel A depicts the flow of data from the nation down to the census block. Panel B depicts the optimization step run for each geographic level.

To ameliorate inaccuracies in the DDP, the Census Bureau redesigned the post-processing routine, creating what is now called the *multipass* routine. This new routine was used on the v20200527 and v20200917 datasets. Instead of optimizing all cells in the detailed and DP queries at one time for a given geographic level, the routine runs multiple optimization steps for each geographic level. In the first pass, the routine optimizes the differences between the noisy detailed histogram and the total population and household/group quarters counts. In the second pass, the routine optimizes differences between the noisy detailed histogram and the *voting age * Hispanic * race* query, controlling the counts to the results of the first pass (Figure 2, Panel B). In other words, if we sum the optimized counts from the *voting age * Hispanic * race* categories, the result will be close to the optimized total population count generated in the first pass.

The third pass optimizes differences between the noisy detailed histogram and the *age * sex * Hispanic * race* query, controlling the counts to the results of the second pass.

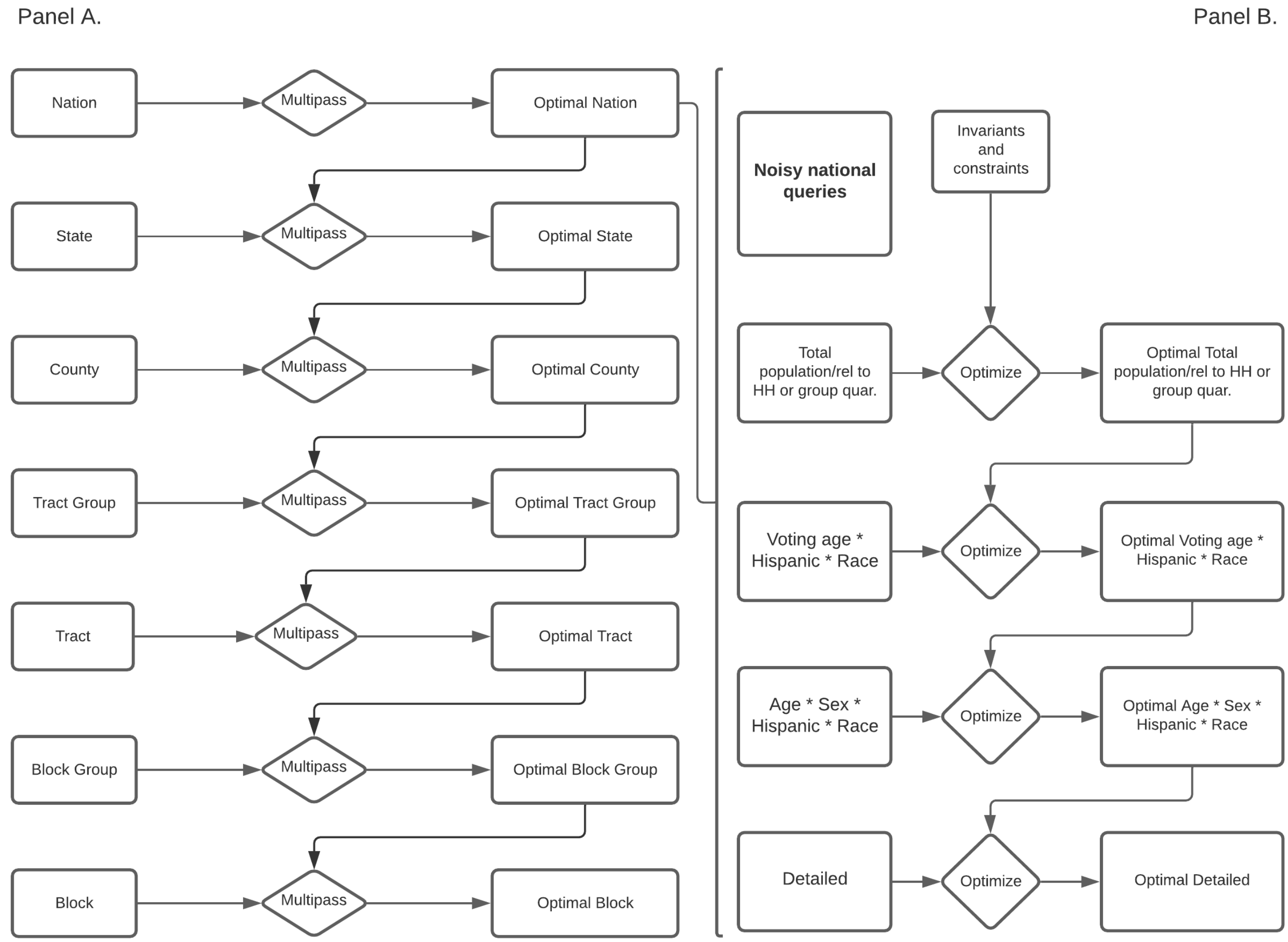

Figure 2. Top Down Algorithm used for the v20200527 dataset. Panel A depicts the flow of data from the nation down ot the census block. Panel B depicts the optimization steps run for each geographic level.

These passes are run for each geographic level in the geographic hierarchy (Figure 2, Panel A). All passes are run for the nation level first, then the state level, and down the hierarchy until they are run for the block level. Essentially, the multipass routine is top-down along two dimensions—the geographic hierarchy and the query hierarchy. By adding the query hierarchy and the extra optimization steps for queries, the Bureau may prioritize particular use cases. For the v20200527 dataset, the top query priority was total population counts and counts of persons in households or group quarters, and the second and third query priorities were counts that support redistricting and population estimates, respectively. We assume the query priority for the v20200917 dataset is these as the order listed in Table 2, but we do not have confirmation about priority from the Bureau.

Data

We use race/ethnicity counts from 2010 Summary File 1 and the three differentially private datasets described above. The Census Bureau provides data for many different race categories, and we generate counts for the following groups:

- White alone
- Black alone
- non-Hispanic White
- non-Hispanic Black

- non-Hispanic All other races
- Hispanic

The segregation measures we compute require different sets of counts, and the above list covers all of the required sets.

Following the example of most racial residential segregation studies, we use census tracts and core-based statistical areas (CBSA) as our main units of geography. For better or worse, census tracts are typically used as proxies for residential neighborhoods. Core-based statistical areas are composed of counties having a high degree of economic and social integration, often through commuting patterns.

Methods

Since Duncan and Duncan published their index of dissimilarity in 1955 (Duncan & Duncan, 1955), scientists have invented or refined dozens of segregation indices. These indices all attempt to measure "the extent to which individuals of different groups occupy of experience different social environments" (Reardon & O'Sullivan, 2004). Readers interested in learning more about measures of segregation are directed to Massey and Denton (Massey & Denton, 1988) and Reardon and O'Sullivan (Reardon & O'Sullivan, 2004). These two papers do an excellent job comparing measures and describing what aspects of segregation are being captured. For our paper, we select a number of indices, including the commonly used index of dissimilarity (*D*), the multigroup entropy (or information theory) index (*H*), the multigroup index of dissmilarity, spatial versions of the index of dissimilarity and multigroup entropy, and the spatial isolation index.

We use both aspatial and spatial measures of segregation to examine whether one family performs better using noise injected data. Since these measures are used across a broad range of disciplines, it is crucial that we provide guidance regarding their usage with 2020 decennial census data.

Index of dissimilarity

The index of dissimilarity, *D*, compares the distribution of mutually exclusive population subgroups among a set of census tracts nested within a CBSA (Iceland et al., 2002; Massey & Denton, 1988). Typically, *D* is computed for two mutually exclusive subgroups but has been extended to three or more sub-groups (Reardon & Firebaugh, 2002). *D* is calculated as:

$$D = \frac{1}{2} \sum | \frac{w_i}{W} - \frac{b_i}{B} |$$

where w_i is the White population of census tract *i*, W is the White population of the CBSA containing tract *i*, b_i is the Black population of tract *i*, and B is the black population of the CBSA containing tract *i*.

D varies from 0 to 1. Dissimilarity is minimized when the proportions of the subgroups within each census tract is the same as their proportion in the CBSA. Dissimilarity is maximized when census tracts contain persons from a single subgroup.

D is a widely used segregation measure (Frey & Myers, 2005) but is extremely sensitive to small changes. Measurement error, such as that introduced by the new disclosure avoidance technique, in low-population areas can cause upward bias (Napierala & Denton, 2017). If a CBSA contains few members of a subgroup and the noise injection adds new members, particularly if those new members are concentrated in a tract or two, *D* will increase.

Multigroup entropy

Multigroup entropy, H, is a measure of evenness that examines the simultaneous distribution of three of more mutually exclusive population subgroups among a set of census tracts nested within a CBSA (Iceland et al., 2002; Iceland, 2004). We divide our data into four subgroups—non-Hispanic Whites, non-Hispanic Blacks, non-Hispanic all other races[7], and Hispanic. H is effectively the weighted deviation of a census tract's entropy from the CBSA's entropy.

Entropy of a census tract is calculated as:

$$e_i = \sum_{j=1}^{k} p_{ij} ln(\frac{1}{p_{ij}})$$

where k is the number of subgroups and p_{ij} is the jth subgroup's proportion of census tract i's total population.

To compute the multigroup entropy, H, for a CBSA:

$$H = \sum_{i=1}^{n} [\frac{t_i(e - e_i)}{et}]$$

where n is the number of census tracts in the CBSA, t_i is the total population of census tract i, e is the entropy index for the CBSA, e_i is the entropy index for census tract i, and t is the total population for the CBSA.

H ranges between 0 to 1, with lower values indicating less segregation. CBSAs with relatively even distributions of subgroups across census tracts will have lower values of H.

Multigroup entropy is often preferred over other segregation measures because of its decomposability and its sensitivity to the reconfiguration of population among census tracts (Reardon, 2017; Reardon & Firebaugh, 2002; Reardon & O'Sullivan, 2004). Measurement error will affect the value of H but its impact is lessened because H uses the population-weighted sum.

Spatial versions of the index of dissimilarity and multigroup entropy

Both the index of dissimilarity and multigroup entropy are aspatial measures of segregation. These measures treat census tracts as independent units of analysis and do not take into account the spatial structure of the tracts (Morrill, 1991; Oka & Wong, 2015; Reardon & O'Sullivan, 2004; Yao et al., 2019). Reardon and O'Sullivan (Reardon & O'Sullivan, 2004) describe spatial versions of these two commonly used measures. The spatial versions convert polygon-based census tracts to continuous surfaces and then use a two-dimensional kernel when computing the indices. The kernel weights values near its center the highest, and the weights decrease as we move further outward. The kernel allows us to capture population in nearby census tracts and use these neighboring values to more properly capture segregation.

Using population counts from neighboring units may attenuate the impact of measurement error on index values. If the count for a subgroup in a reference tract increases after disclosure avoidance and counts for the same subgroup increase and decrease in various neighboring units, the net effect of the measurement error will be less than it would be if we failed to take neighbors into account.

Results

Index of Dissimilarity

We begin by comparing CBSA values of D computed from SF1 with values computed from the three demonstration datasets (Figure 1). The red dots represent metropolitan CBSAs, and the blue dots represent micropolitan CBSAs. The micropolitan units consist of counties with an urban cluster of 10,000 people or more and less than 50,000; thus, micropolitan units are smaller than metropolitan units. In all three demonstration datasets, we observe larger differences between SF1 and the demonstration datasets in the micropolitan areas but we do still see variation in metropolitan areas, particularly in the v20200917 scatterplot.

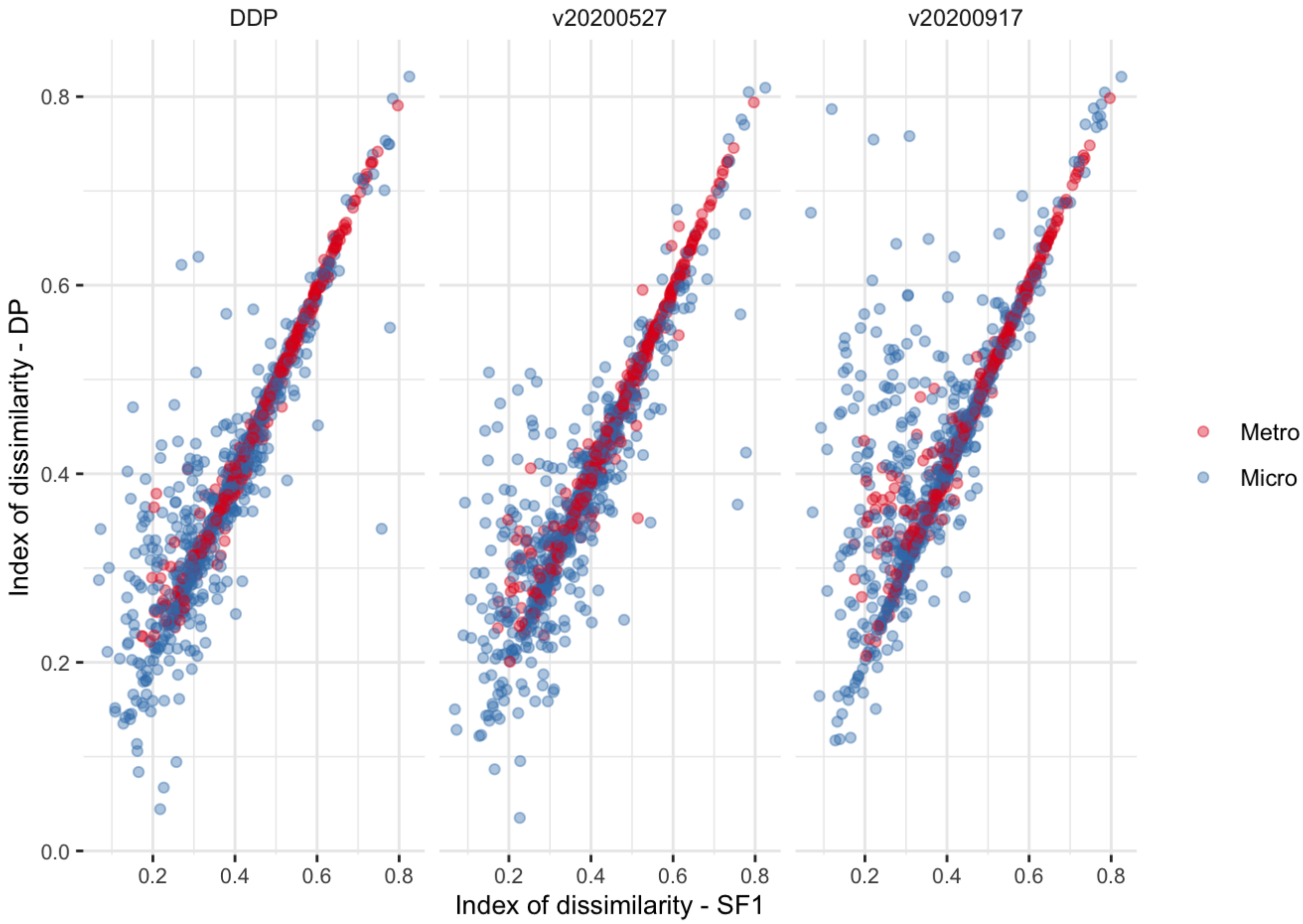

Figure 1. Index of dissimilarity (D) for White alone - Black alone. The x-axis is SF1-based value of D, and the y-axis is the value of D computed from the demonstration datasets.

Multigroup entropy

Comparing CBSA values of H computed from SF1 with values computed from the three demonstration datasets (Figure 2), we observe minimal differences between SF1 and the demonstration datasets. There are a few outliers in the v20200527 dataset but those are not seen on the v20200917 dataset.

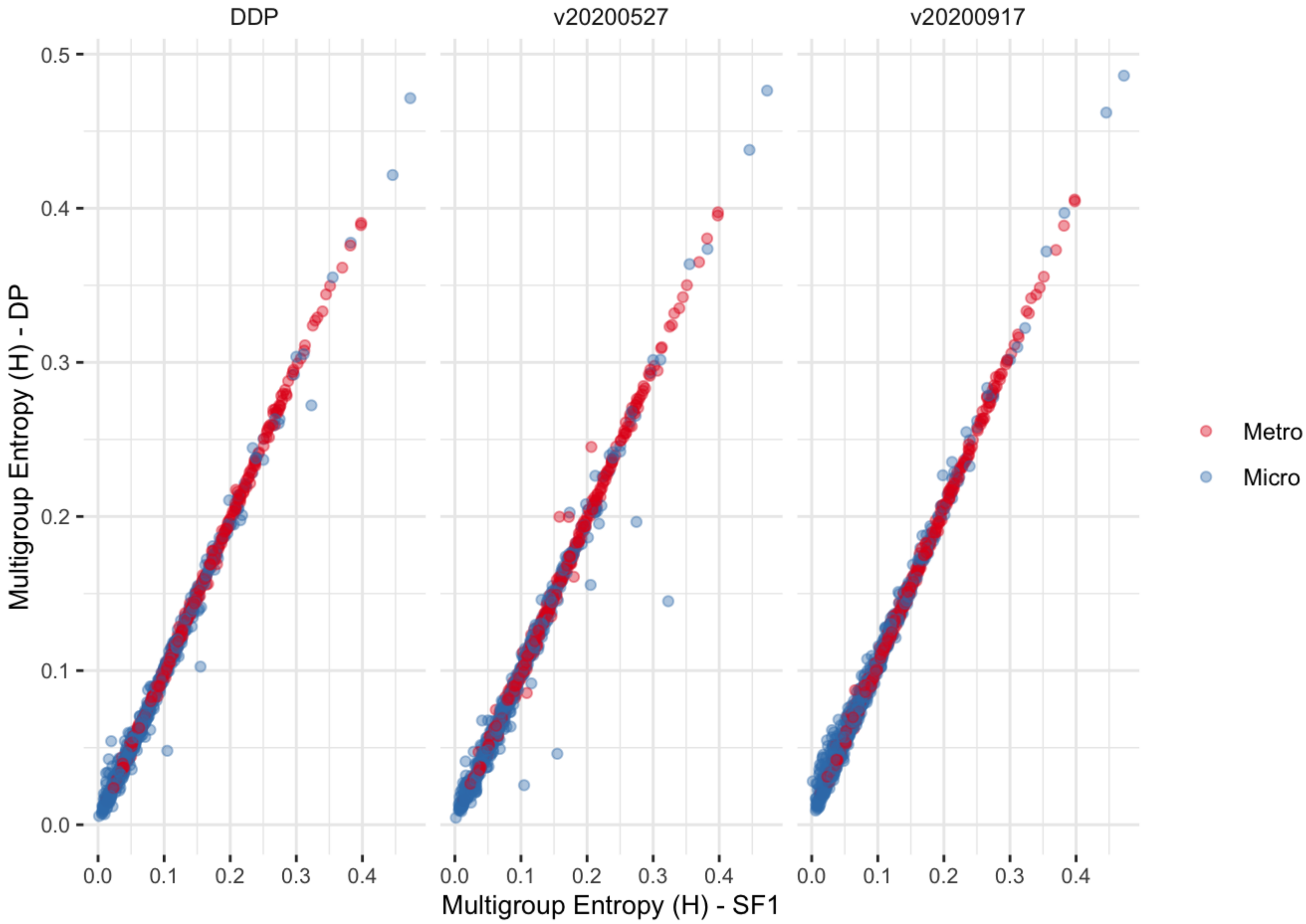

Figure 2. Scatterplot of CBSA values for multigroup entropy (H). The x-axis is SF1-based value of H, and the y-axis is the value of H computed from the demonstration datasets.

Multigroup index of dissimilarity

When we consider the multigroup index of dissimilarity (Figure 3), we observe substantially less difference between the demonstration datasets and SF1 compared to the two-group D shown in Figure 1.

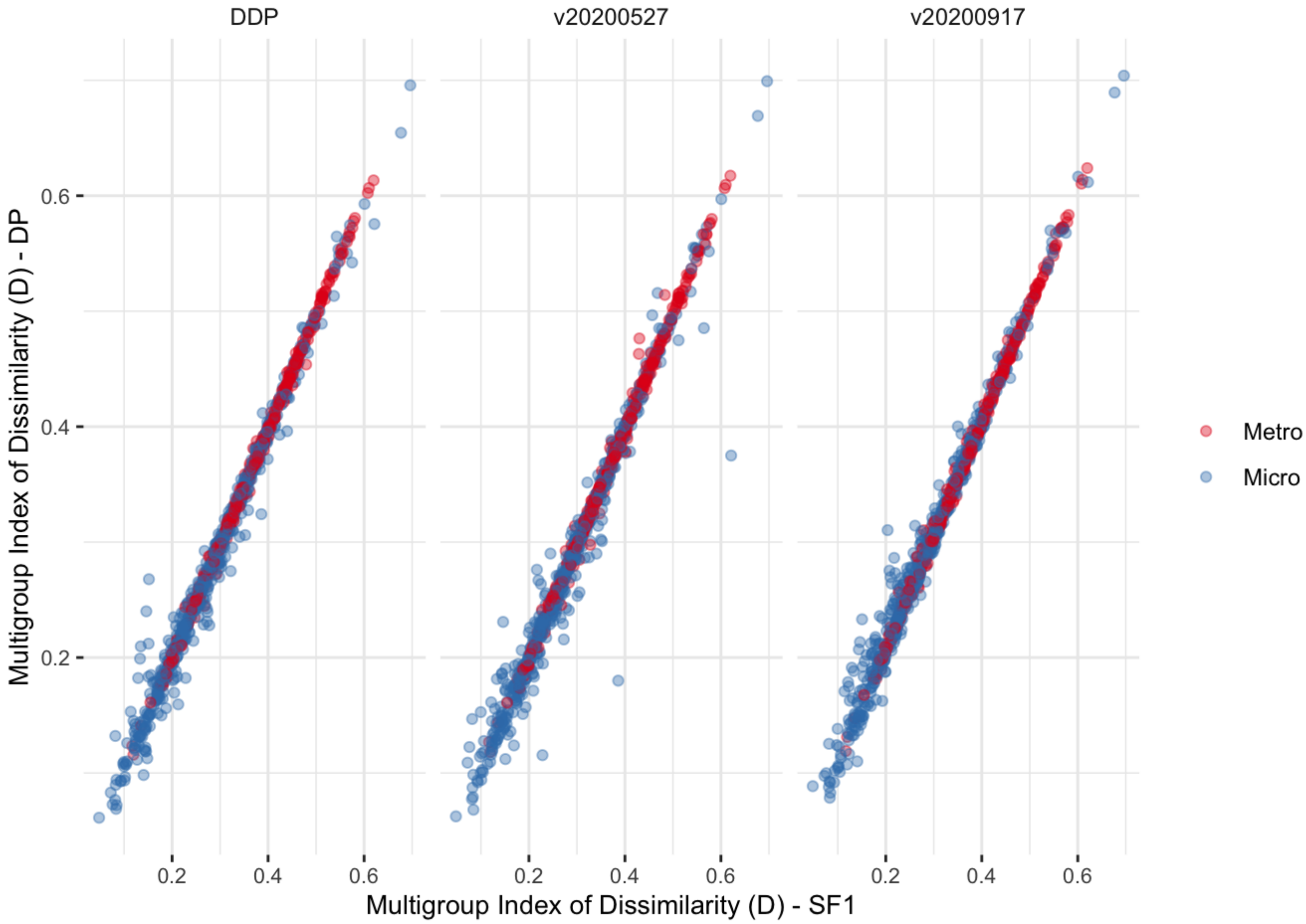

Figure 3. Scatterplot of CBSA values for multigroup index of dissimilarity (D). The x-axis is SF1-based value of D, and the y-axis is the value of D computed from the demonstration datasets.

Spatial Multigroup Index of Dissimilarity

When we consider the spatial version of the multigroup index of dissimilarity (Figure 4), we observe similar results to those seen in Figure 3. Whether using the aspatial or spatial version, the multigroup index of dissimilarity demonstrations less variation in the demonstration datasets compared to the two-group D shown in Figure 1.

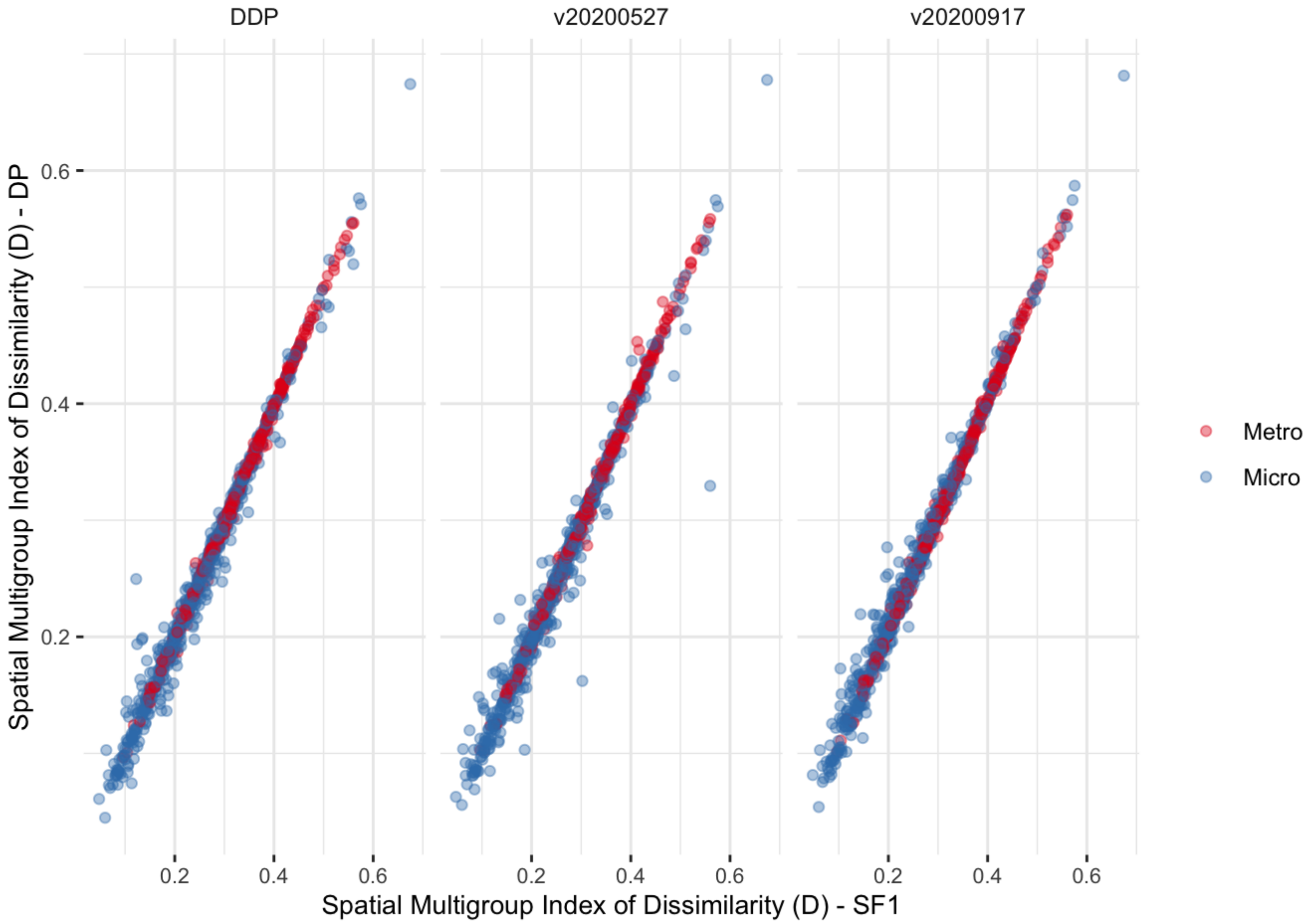

Figure 4. Scatterplot of CBSA values for the spatial version of the multigroup index of dissimilarity (H). The x-axis is SF1-based value of spatial multigroup D, and the y-axis is the value of spatial multigroup D computed from the demonstration datasets. The spatial multigroup D values incorporate information from neighboring census tracts via a smoothing function. We use the default kernel option.

Spatial Multigroup Entropy

Comparing CBSA values of spatial H computed from SF1 with values computed from the three demonstration datasets (Figure 4), we observe similar results to those seen in Figure 2. There are a few outliers in the v20200527 dataset but those are not seen on the v20200917 dataset.

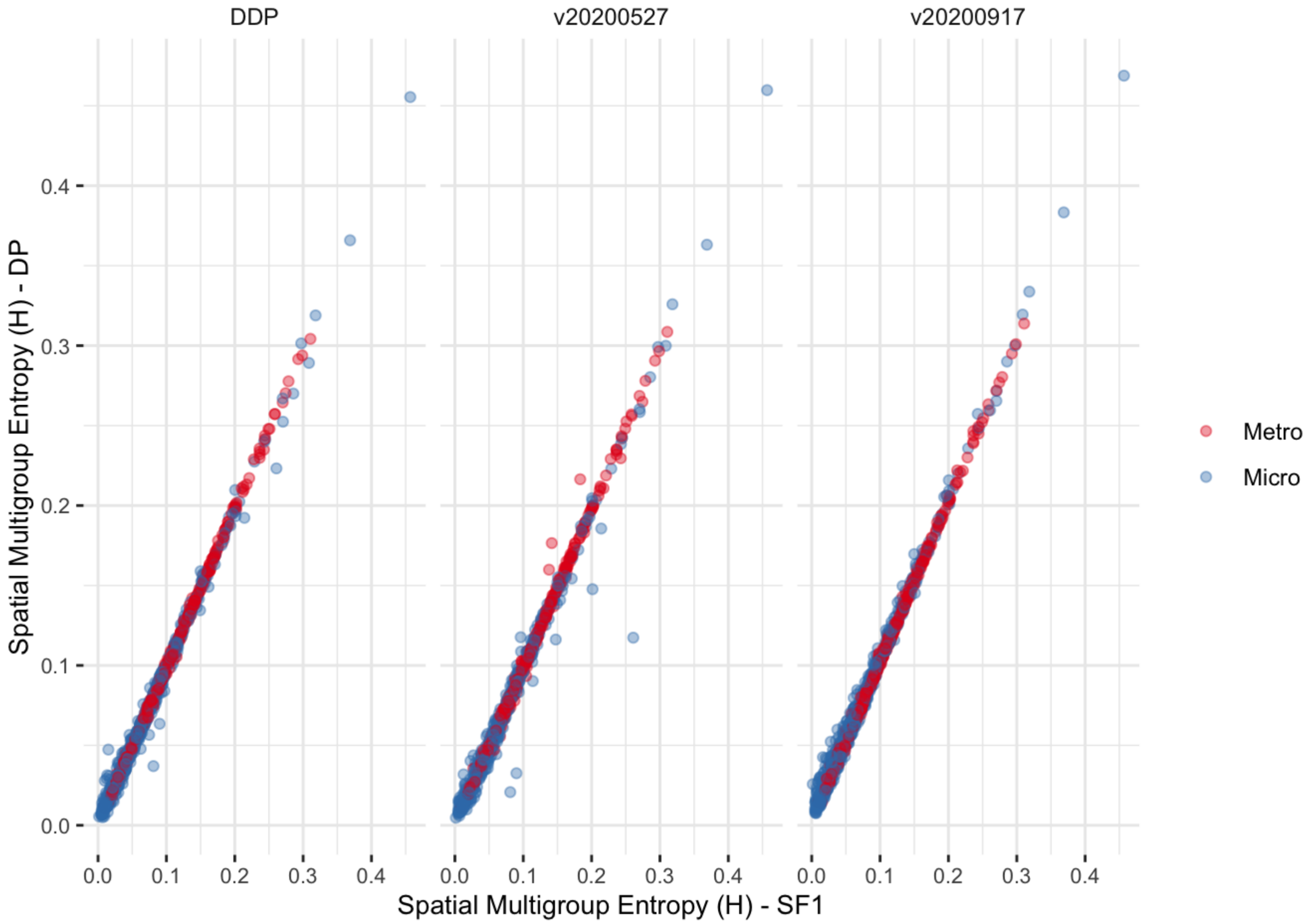

Figure 5. Scatterplot of CBSA values for the spatial version of multigroup entropy (H). The x-axis is SF1-based value of spatial H, and the y-axis is the value of spatial H computed from the demonstration datasets. The spatial H values incorporate information from neighboring census tracts via a smoothing function. We use the default kernel option.

Discussion

The Census Bureau's new disclosure avoidance system, based on differentially privacy, will inject noise into almost all published population counts for the 2020 decennial census. Our paper examines the impact that this noise injection may have on the measurement of racial residential segregation. We compute a number of commonly used segregation measures for core-based statistical areas using four different datasets—SF1, DDP, v20200527, and v20200917. We then plot the SF1-based metrics versus the metrics based on the differentially private datasets.

Scientists should be wary of calculating the index of dissimilarity D with the 2020 decennial data if the 2020 data are produced using similar parameters to those used for the DDP, v20200527 and v20200917 datasets. We observe substantial differences in the value of D based on differentially private datasets compared to SF1-based values. If scientists naively compute D from differentially private data and compare it with D based on prior decennial censuses, they may severely under or overestimate changes in segregation, particularly for smaller metropolitan areas and for micropolitan areas.

For the other segregation indices, it matters less whether the metrics are based on SF1 or differentially private data. We observe smaller differences between SF1-based and differentially private-based metrics for other indices. In particular, values for metropolitan CBSAs seem robust to the disclosure avoidance techniques used to generate published data. Additionally, it does not seem to matter whether we use the spatial or aspatial version of multigroup entropy with various 2010 datasets. The shapes of the scatterplots are similar between the spatial (Figure 5) and aspatial (Figure 2) indices.

One limitation of this study is our inability to quantify the uncertainty associated with segregation measures computed differentially private data. The counts in the DDP, v20200527, and v20200917 do not have associated error bounds (e.g., like margins of error in the American Community Survey), and we do not know of a method for computing error bounds that accounts for both noise injection and post-processing. The noise injection is unbiased but post-processing introduces bias into the final counts. If scientists can account for the uncertainty, they may be able to appropriately adjust their segregation indices.

References

Abowd, J. (2018). *Protecting the Confidentiality of America's Statistics: Adopting Modern Disclosure Avoidance Methods at the Census Bureau*. https://www.census.gov/newsroom/blogs/research-matters/2018/08/protecting_the_confi.html (https://www.census.gov/newsroom/blogs/research-matters/2018/08/protecting_the_confi.html)

boyd, danah. (2019). *Balancing Data Utility and Confidentiality in the 2020 US Census*. Data and Society. https://datasociety.net/wp-content/uploads/2019/12/DS_Differential_Privacy_L.pdf (https://datasociety.net/wp-content/uploads/2019/12/DS_Differential_Privacy_L.pdf)

Bureau, U. C. (2019). *2010 Demonstration Data Products*. US Census Bureau. https://www.census.gov/programs-surveys/decennial-census/2020-census/planning-management/2020-census-data-products/2010-demonstration-data-products.html (https://www.census.gov/programs-surveys/decennial-census/2020-census/planning-management/2020-census-data-products/2010-demonstration-data-products.html)

Cutler, D. M., & Glaeser, E. L. (1997). Are Ghettos Good or Bad? *The Quarterly Journal of Economics*, *112*(3), 827–872. https://doi.org/10.1162/003355397555361 (https://doi.org/10.1162/003355397555361)

Dinur, I., & Nissim, K. (2003). Revealing Information While Preserving Privacy. *Proceedings of the Twenty-Second ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems*, 202–210. https://doi.org/10.1145/773153.773173 (https://doi.org/10.1145/773153.773173)

Duncan, O. D., & Duncan, B. (1955). A Methodological Analysis of Segregation Indexes. *American Sociological Review*, *20*(2), 210–217. https://doi.org/10.2307/2088328 (https://doi.org/10.2307/2088328)

Dwork, C., McSherry, F., Nissim, K., & Smith, A. (2006). Calibrating Noise to Sensitivity in Private Data Analysis. In S. Halevi & T. Rabin (Eds.), *Theory of Cryptography* (pp. 265–284). Springer Berlin Heidelberg.

Frey, W., & Myers, D. (2005). *Racial Segregation in US Metropolitan Areas and Cities, 1990-2000: Patterns, Trends, and Explanations*. University of Michigan Population Studies Center. https://www.psc.isr.umich.edu/pubs/abs/3199 (https://www.psc.isr.umich.edu/pubs/abs/3199)

Garfinkel, S. L., Abowd, J. M., & Powazek, S. (2018). Issues Encountered Deploying Differential Privacy. *Proceedings of the 2018 Workshop on Privacy in the Electronic Society*, 133–137. https://doi.org/10.1145/3267323.3268949 (https://doi.org/10.1145/3267323.3268949)

Grigoryeva, A., & Ruef, M. (2015). The Historical Demography of Racial Segregation. *American Sociological Review*, *80*(4), 814–842. https://doi.org/10.1177/0003122415589170 (https://doi.org/10.1177/0003122415589170)

Hansen, M. (2018, December 5). To Reduce Privacy Risks, the Census Plans to Report Less Accurate Data (Published 2018). *The New York Times: The Upshot*. https://www.nytimes.com/2018/12/05/upshot/to-reduce-privacy-risks-the-census-plans-to-report-less-accurate-data.html

(https://www.nytimes.com/2018/12/05/upshot/to-reduce-privacy-risks-the-census-plans-to-report-less-accurate-data.html)

Hotz, V. J., & Salvo, J. (2020). *Assessing the Use of Differential Privacy for the 2020 Census: Summary of What We Learned from the CNSTAT Workshop*. https://www.amstat.org/ASA/News/Co-Chairs-Share-Insights-Gleaned-from-2020-Census-Data-Products-Workshop.aspx (https://www.amstat.org/ASA/News/Co-Chairs-Share-Insights-Gleaned-from-2020-Census-Data-Products-Workshop.aspx)

Iceland, J. (2004). *The multigroup entropy index (also known as Theil's H or the information theory index)*. US Census Bureau. https://www2.census.gov/programs-surveys/demo/about/housing-patterns/multigroup_entropy.pdf (https://www2.census.gov/programs-surveys/demo/about/housing-patterns/multigroup_entropy.pdf)

Iceland, J., Weinberg, D. H., & Steinmetz, E. (2002). *Racial and Ethnic Residential Segregation in the United States: 1980-2000* (Nos. CENSR-3; Census 2000 Special Reports, p. 151). US Census Bureau. https://www.census.gov/prod/2002pubs/censr-3.pdf (https://www.census.gov/prod/2002pubs/censr-3.pdf)

Iceland, J., & Wilkes, R. (2006). Does Socioeconomic Status Matter? Race, Class, and Residential Segregation. *Social Problems*, *53*(2), 248–273. https://doi.org/10.1525/sp.2006.53.2.248 (https://doi.org/10.1525/sp.2006.53.2.248)

Leclerc, P. (2019, June 5). *Reconstruction of Person Level Data from Data Presented in Multiple Tables*. Challenges and New Approaches for Protecting Privacy in Federal Statistical Programs: A Workshop. https://vimeo.com/showcase/6076198/video/343287917 (https://vimeo.com/showcase/6076198/video/343287917)

Logan, T. D., & Parman, J. M. (2018). Segregation and mortality over time and space. *Social Science & Medicine*, *199*, 77–86. https://doi.org/10.1016/j.socscimed.2017.07.006 (https://doi.org/10.1016/j.socscimed.2017.07.006)

Logan, T. D., & Parman, J. M. (2017). The National Rise in Residential Segregation. *The Journal of Economic History*, *77*(1), 127–170. https://doi.org/10.1017/S0022050717000079 (https://doi.org/10.1017/S0022050717000079)

Massey, D. S., & Denton, N. A. (1988). The Dimensions of Residential Segregation. *Social Forces*, *67*(2), 281–315. https://doi.org/10.2307/2579183 (https://doi.org/10.2307/2579183)

McKenna, L. (2018). *Disclosure Avoidance Techniques Used for the 1970 through 2010 Decennial Censuses of Population and Housing* (Nos. 18-47; p. 39). US Census Bureau. https://www2.census.gov/ces/wp/2018/CES-WP-18-47.pdf (https://www2.census.gov/ces/wp/2018/CES-WP-18-47.pdf)

Mehra, R., Boyd, L. M., & Ickovics, J. R. (2017). Racial residential segregation and adverse birth outcomes: A systematic review and meta-analysis. *Social Science & Medicine*, *191*, 237–250. https://doi.org/10.1016/j.socscimed.2017.09.018 (https://doi.org/10.1016/j.socscimed.2017.09.018)

Mendez, D. D., Hogan, V. K., & Culhane, J. F. (2014). Institutional racism, neighborhood factors, stress, and preterm birth. *Ethnicity & Health*, *19*(5), 479–499. https://doi.org/10.1080/13557858.2013.846300 (https://doi.org/10.1080/13557858.2013.846300)

Morrill, R. (1991). On the Measure of Geographic Segregation. *Geography Research Forum*, *11*, 25–36. http://raphael.geography.ad.bgu.ac.il/ojs/index.php/GRF/article/view/91 (http://raphael.geography.ad.bgu.ac.il/ojs/index.php/GRF/article/view/91)

Napierala, J., & Denton, N. (2017). Measuring Residential Segregation With the ACS: How the Margin of Error Affects the Dissimilarity Index. *Demography*, *54*(1), 285–309. https://doi.org/10.1007/s13524-016-0545-z (https://doi.org/10.1007/s13524-016-0545-z)

Oka, M., & Wong, D. W. S. (2015). Spatializing Segregation Measures: An Approach To Better Depict Social Relationships. *Cityscape*, *17*(1), 97–114. http://www.jstor.org/stable/26326925 (http://www.jstor.org/stable/26326925)

Quillian, L. (2014). Does Segregation Create Winners and Losers? Residential Segregation and Inequality in Educational Attainment. *Social Problems*, *61*(3), 402–426. https://doi.org/10.1525/sp.2014.12193 (https://doi.org/10.1525/sp.2014.12193)

Reardon, S. (2017). A Conceptual Framework for Measuring Segregation and its Association with Population Outcomes. In *Methods in Social Epidemiology* (2nd ed., pp. 132–157). Jossey-Bass.

Reardon, S. F., & Firebaugh, G. (2002). Measures of Multigroup Segregation. *Sociological Methodology*, *32*(1), 33–67. https://doi.org/10.1111/1467-9531.00110 (https://doi.org/10.1111/1467-9531.00110)

Reardon, S. F., Matthews, S. A., O'Sullivan, D., Lee, B. A., Firebaugh, G., Farrell, C. R., & Bischoff, K. (2008). The geographic scale of metropolitan racial segregation. *Demography*, *45*(3), 489–514. https://doi.org/10.1353/dem.0.0019 (https://doi.org/10.1353/dem.0.0019)

Reardon, S. F., & O'Sullivan, D. (2004). Measures of Spatial Segregation. *Sociological Methodology*, *34*(1), 121–162. https://doi.org/10.1111/j.0081-1750.2004.00150.x (https://doi.org/10.1111/j.0081-1750.2004.00150.x)

Van Riper, D., Kugler, T., & Schroeder, J. (2020a). *IPUMS NHGIS Privacy-Protected 2010 Census Demonstration Data, version 20200527 [Database]*. IPUMS. https://nhgis.org/privacy-protected-demonstration-data (https://nhgis.org/privacy-protected-demonstration-data)

Van Riper, D., Kugler, T., & Schroeder, J. (2020b). *IPUMS NHGIS Privacy-Protected 2010 Census Demonstration Data, version 20200917 [Database]*. IPUMS. https://nhgis.org/privacy-protected-demonstration-data (https://nhgis.org/privacy-protected-demonstration-data)

Williams, D. R., & Collins, C. (2001). Racial residential segregation: A fundamental cause of racial disparities in health. *Public Health Reports*, *116*(5), 404–416. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1497358/ (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1497358/)

Yao, J., Wong, D. W. S., Bailey, N., & Minton, J. (2019). Spatial Segregation Measures: A Methodological Review. *Tijdschrift Voor Economische En Sociale Geografie*, *110*(3), 235–250. https://doi.org/10.1111/tesg.12305 (https://doi.org/10.1111/tesg.12305)

---

1. Readers interested in the history of the Bureau's disclosure avoidance techniques are directed to McKenna (McKenna, 2018)↩
2. Traditional disclosure avoidance techniques, such as swapping, target statistically unique individuals or households identified after generating the initial cross-tabulations. After identifying these individuals or households, the Bureau applies swapping and then re-creates the cross-tabulations from the swapped microdata. Swap rules and rates are secret and have never been released to the general public(boyd, 2019; McKenna, 2018). Hansen found evidence of swapping when he examined the demographic characteristics of the census block containing Liberty Island in New York City (Hansen, 2018), and the type and shape, along with the post-processing routine, influence the accuracy and utility of the published data and the privacy protection afforded respondents.↩
3. The DDP included housing unit and household data. A PLB of 2.0 was provided housing unit-based queries. The v20200527 and v20200917 only include person data.↩
4. The census tract group is not a standard unit in the Census Bureau's geographic hierarchy. It was created specifically for the disclosure avoidance system to control the number of child units for each county↩

5. The Bureau changed to the discrete Gaussian distribution to minimize the risk of drawing a large random value. For the Laplace or double geometric distribution, approximately 75% of the random values will fall within 1 standard deviation of the mean as compared to 68% of random values in the normal distribution. But the tails of the Laplace distribution are thicker than the normal distribution, increasing the probability of drawing a random value that is 3 or 4 times the standard deviation. Specifically, 98.56% of the random values for the Laplace distribution fall within 3 standard deviations compared to 99.73% of random values in the normal distribution.↩

6. Sensitivity is the value by which a query changes if we make a single modification to the database. Histogram queries have a sensitivity of 2 - if we increase the count in a cell by 1, we must decrease the count in another cell by 1.↩

7. American Indian/Alaska Native, Asian, Native Hawaiian or Pacific Islander, Some other race, Two or more races↩