The Impact of the U.S. Census Disclosure Avoidance System on Redistricting and Voting Rights Analysis

Christopher T. Kenny
Department of Government
Harvard University
Cambridge, MA
christopherkenny@fas.harvard.edu

Shiro Kuriwaki
Department of Government
Harvard University
Cambridge, MA
kuriwaki@g.harvard.edu

Cory McCartan
Department of Statistics
Harvard University
Cambridge, MA
cmccartan@fas.harvard.edu

Evan Rosenman
Harvard Data Science Initiative
Harvard University
Cambridge, MA
erosenm@fas.harvard.edu

Tyler Simko
Department of Government
Harvard University
Cambridge, MA
tsimko@g.harvard.edu

Kosuke Imai *
Department of Government and Department of Statistics
Harvard University
Cambridge, MA
imai@harvard.edu

May 28, 2021

Abstract

The U.S. Census Bureau plans to protect the privacy of 2020 Census respondents through its Disclosure Avoidance System (DAS), which attempts to achieve differential privacy guarantees by adding noise to the Census microdata. By applying redistricting simulation and analysis methods to DAS-protected 2010 Census data, we find that the protected data are not of sufficient quality for redistricting purposes. We demonstrate that the injected noise makes it impossible for states to accurately comply with the *One Person, One Vote* principle. Our analysis finds that the DAS-protected data are biased against certain areas, depending on voter turnout and partisan and racial composition, and that these biases lead to large and unpredictable errors in the analysis of partisan and racial gerrymanders. Finally, we show that the DAS algorithm does not universally protect respondent privacy. Based on the names and addresses of registered voters, we are able to predict their race as accurately using the DAS-protected data as when using the 2010 Census data. Despite this, the DAS-protected data can still inaccurately estimate the number of majority-minority districts. We conclude with recommendations for how the Census Bureau should proceed with privacy protection for the 2020 Census.

***K*eywords** Census · Redistricting · BISG · Differential privacy · TopDown algorithm · One Person One Vote

*To whom correspondence should be addressed. We thank Ben Fifield and the ACLU for providing precinct-level state legislative assignments and election data for several states, and Bruce Willsie of L2, Inc for providing voterfiles.

Contents

1 Introduction

In preparation for the official release of the 2020 Census data, the United States Census Bureau has built the Disclosure Avoidance System (DAS) to prevent Census respondents from being linked to specific people [1]. The DAS is based on differential privacy technology, which adds a certain amount of random noise to the raw Census counts. The decision to use differential privacy for the 2020 Census has been controversial, with many scholars voicing concerns about the negative impacts of noisy data on public policy and social science research, which critically rely upon the Census data [2, 3].

In this paper, we empirically evaluate the impact of the DAS on redistricting and voting rights analysis. Once released as part of the 2020 Census data later this year, states will use the P.L. 94-171 redistricting data to redraw their district boundaries of Congressional and other federal and local electoral offices. It is therefore of paramount importance to examine how the DAS affects redistricting analysis and the map-drawing process.

The Census Bureau has requested public feedback on the "fitness-for-use" of the P.L. 94-171 data by making available the Privacy-Protected Microdata Files (PPMFs) based on the application of the DAS to the 2010 Census redistricting data. The Census Bureau released two PPMFs at different levels of *privacy loss budget,* ϵ, which controls the amount of noise. The DAS-12.2 data are based on a relatively high level of privacy loss budget ($\epsilon = 12.2$) to achieve the accuracy targets at the expense of greater privacy loss, whereas the DAS-4.5 data use a lower privacy loss budget at the expense of worse accuracy ($\epsilon = 4.5$). In addition, the Census Bureau post-processes the noisy data in order to ensure that the resulting public release data are self-consistent (e.g., no negative counts) and certain aggregate statistics such as state-level total population counts are accurate.

We examine the fitness-for-use of PPMFs through a variety of redistricting and voting rights analyses. In particular, we employ a set of recently developed simulation methods that can generate a large number of realistic redistricting maps under a set of legal and other relevant constraints, such as contiguity, compactness, population parity, and preservation of communities of interest and counties [4, 5, 6, 7, 8, 9, 10]. These simulation methods have been extensively used by expert witnesses in recent court cases on redistricting, including *Common Cause v. Lewis* (2020), *Rucho v. Common Cause* (2019), *Ohio A. Philip Randolph Institute v. Householder* (2020), *League of Women Voters of Michigan v. Benson* (2019), *League of Women Voters v. Pennsylvania* (2017), *Missouri State Conference of the NAACP v. Ferguson-Florissant School District* (2017), *Raleigh Wake Citizens Association v. Wake County Board of Elections* (2016), and *City of Greensboro v. Guilford County Board of Elections* (2015). These cases span all levels of government: local redistricting, state legislative redistricting, and congressional redistricting. We apply the simulation methods to the DAS-12.2 and DAS-4.5 data and compare the results with those obtained based on the 2010 Census data. This comparison reveals how the DAS affects the conclusions of redistricting analysis.

In addition, we examine the impact of DAS on the prediction accuracy of an individual voter's race. Redistricting analysis for voting rights cases often necessitates such individualized prediction because most states' voter lists do not include individual's race. One prominent prediction method combines the Census block-level proportion of each race with a voter's name and address [11, 12, 13]. This methodology played a key role in the most recent racial gerrymandering case, *NAACP, Spring Valley Branch et al. v. East Ramapo School District* (2020), in which the federal Court of Appeals for the Second Circuit upheld the district court's ruling that the school board elections violated the Voting Rights Act. We reanalyze this case using the DAS data and compare the results with those based on the 2010 Census data.

2 Overview of Analysis

For the purposes of evaluating the impact of the new DAS on redistricting plan-drawing and analysis, we generated eight sets of redistricting datasets for simulation, described in Table 1. We create precinct-level datasets that have three versions of total population counts: the original 2010 Census, the DAS-12.2 data, and the DAS-4.5 data.

In our modal analysis, we simulate realistic district plans under the scenario that population counts are given by each of the three datasets. All simulations were conducted with the SMC redistricting sampler of [9], except for the Louisiana House of Representatives Districts for East Baton Rouge, which were conducted with a Merge-Split-type MCMC sampler similar to that of [5, 6]. Both of these sampling algorithms are implemented in the open-source software package `redist` [10]. All sampling diagnostics, including the number of effective samples, indicated accurate sampling and adequate sample diversity.

The DAS-12.2 data yield precinct population counts that are roughly 1.0% different from the original Census, and the DAS-4.5 data are about 1.9% different. For the average precinct, this amounts to a discrepancy of 18 people (for DAS-12.2) or 33 people (for DAS-4.5) moving across precinct boundaries. Therefore, our main simulation results should be thought of as a study of how such precinct-level differences propagate into noise at the district-level by exploring redistricting plans.

2.1 Population Parity

Perhaps the strongest constraint on modern redistricting is the requirement that districts be nearly equal in population. Deviations in population between districts have the effect of diluting the power of voters in larger-population districts. The importance of this principle stems from a series of Supreme Court cases in the 1960s, beginning with *Gray v. Sanders* (1963), in which the court held that political equality comes via a standard known as *One Person, One Vote.* As for acceptable deviations from population equality, *Wesberry v. Sanders* (1964) set the basic terms by holding that the Constitution requires that "as nearly as is practicable, one person's vote in a congressional election is to be worth as much as another's." Even minute differences in population parity across congressional districts must be justified, even when smaller than the expected error in decennial Census figures (*Karcher v. Daggett* 1983). For state legislative districts, *Reynolds v. Sims* (1964) held that they must be drawn to near population equality. However, subsequent rulings stated that states may allow for small population deviations when seeking other legitimate interests (*Mahan v. Howell* 1972; *Gaffney v. Cummings* 1973).

When measuring population equality, states must rely on Census data, which was viewed as the most reliable source of population figures (*Kirkpatrick v. Preisler* 1969). We therefore empirically examine how the DAS affects the ability to draw redistricting maps that adhere to this equal population principle. We simulate

State	Office	Districts	Precincts	Total simulated plans
Pennsylvania	U.S. House	18	9,256	30k
Louisiana	State Senate	39	3,668	60k
Louisiana*	State House	15	361	1,700k
North Carolina	U.S. House	13	2,692	30k
South Carolina	U.S. House	7	2,122	30k
South Carolina	State House	124	2,122	30k
Mississippi§	State Senate	9	310	30k
New York†	School Board	9	1,207	10k

Table 1: *States and districts studied. We compared the Census 2010, DAS-12.2, and DAS-4.5 datasets in six states and three levels of elections.*
**Examines the Baton Rouge area.*
§*Examines District 22 and its 8 adjacent districts.*
†*Examines the East Ramapo school district, using Census blocks instead of voting precincts.*

realistic maps for Pennsylvania Congressional districts and Louisiana State Senate districts based on the DAS-4.5 and DAS-12.2 data under various levels of population parity. We then examine the degree to which the resulting maps satisfy the same population parity criteria using the 2010 Census data.

2.2 Partisan Effects

If changes in reported population in precincts affect the districts in which they are assigned to, this has implications for which parties win those districts. While a change in population counts of about 1 percent may seem small, differences in vote counts of that magnitude can reverse some election outcomes. Across the five U.S. House elections during 2012 − 2020, 25 races were decided by a margin of less than a percentage point between the Republican and Democratic party's vote shares. And 228 state legislative races were decided by less than a percentage point from 2012–2016.

Partisan implications also raise the concern of gerrymandering, where political parties draw district boundaries to systematically favor their own voters. Many uses of redistricting simulation in redistricting litigation have been over partisan gerrymanders, including *Common Cause v. Lewis*, *Rucho v. Common Cause*, *Ohio A. Philip Randolph Institute v. Householder*, *League of Women Voters of Michigan v. Benson*, and *League of Women Voters v. Pennsylvania*. To evaluate the impact of the DAS on the analysis of potential partisan gerrymanders, we simulate 120,000 redistricting plans across the states of Pennsylvania, North Carolina, and South Carolina, and compare the partisan attributes of the simulated plans from the three data sources. We also analyze voting-related patterns in DAS-induced population count error at the precinct level, and connect these patterns to the statewide findings from the simulations.

2.3 Racial Effects

The Voting Rights Act of 1965, its subsequent amendments, and a series of Supreme Court cases all center race as an important feature of redistricting. A large number of these cases focus on the creation of majority-minority districts (MMDs) (e.g. *Thornburg v. Gingles* 1986, *Shaw v. Reno* 1993, *Miller v. Johnson* 1995, *Shelby County v. Holder* 2013). First, we analyze whether the DAS data systematically undercounts or overcounts certain areas across racial lines. In doing so, we focus attention on the potential consequences of the decision to target accuracy to the majority racial group in a given area [14]. We explore patterns with racial diversity in four states (Pennsylvania, Louisiana, North Carolina, South Carolina).

We also explicitly explore how DAS data can influence the creation of MMDs. To do so, we empirically examine how using the DAS data to create MMDs differs from the same process undertaken using the 2010 Census data. We simulate nearly two million maps in the Louisiana State House and examine the degree to which maps generated using the Census and DAS data lead to different results at the district and precinct levels.

2.4 Ecological Inference and Voting Rights Analysis

Social scientists have developed methods to predict the race and ethnicity of individual voters using Census data. Since *Gingles*, voting rights cases have required evidence that an individual's race is highly correlated with candidate choice. Statistical methods must therefore estimate this individual quantity from aggregate election results and aggregate demographic statistics [15, 16]. A key input to these methods is accurate racial information on voters. To produce this data, recent litigation has used Bayesian Improved Surname Geocoding (BISG) to impute race and ethnicity into a voter file [11, 12, 13]. This methodology is often used to improve classification of the degree of racially polarized voting and racial segregation.

To understand how DAS data influence these analyses, we look at the effect of DAS data on BISG accuracy across several states where race is recorded on the voter file. We then re-examine a recent voting rights case on a school board election in New York using the DAS-12.2 data and compare results to using the Census 2010 data.

3 Summary of Findings

Compared to the original Census 2010 data, we find that the DAS-protected data:

- **Prevent map drawers from creating districts of equal population, according to current statutory and judicial standards.** Actual deviations from equal population will generally be several times larger than as reported under the DAS data. The magnitude of this problem increases for smaller districts such as state legislative districts and school boards.
- **Transfer population from low-turnout, mixed-party areas to high-turnout, single-party areas.** This differential bias leads to different district boundaries, which in turn implies significant and unpredictable differences in election results. The discrepancy also degrades the ability of analysts to reliably identify partisan gerrymanders.
- **Transfer population from racially mixed areas to racially segregated areas.** This bias effectively means racially heterogeneous areas are under-counted. The degree of racial segregation can therefore be over-estimated, which can lead to a change in the number of majority-minority districts. It also creates significant precinct-level variability, which adds substantial unpredictability to whether or not a minority voter is included in a majority-minority district.
- **Alter individual-level race predictions constructed from voter names and addresses.** This leads to fewer estimated minority voters and majority-minority districts in a re-analysis of a recent Voting Rights Act case, *NAACP v. East Ramapo School District.* At a statewide level, however, the DAS data does not curb the ability of algorithms to identify the race of voters from names and addresses. Therefore, this casts doubt on the universal privacy protection guarantee of DAS data.

The subsequent sections deal with these findings and their accompanying methods and data in more detail.

4 Population Parity in Redistricting

Deviation from population parity across n_d districts is generally defined as

$$\text{deviation from parity} = \max_{1 \le k \le n_d} \frac{|P_k - \overline{P}|}{\overline{P}},$$

where P_k denotes the population of district k and $\overline{P}$ denotes the target district population. In other words, we track the percent difference in the district population P_k from the average district size $\overline{P}$, and report the maximum deviation. Our redistricting simulations generate plans that do not exceed a user-specified tolerance. After generating these plans, we then re-evaluate the deviation from parity using the precinct populations from the three data sources.

We find that the noise introduced by the DAS prevents the drawing of equal-population maps with commonly-used population deviation thresholds. Because only one dataset will be available in practice, redistricting practitioners who attempt to create equal-population districts with DAS data should expect the actual

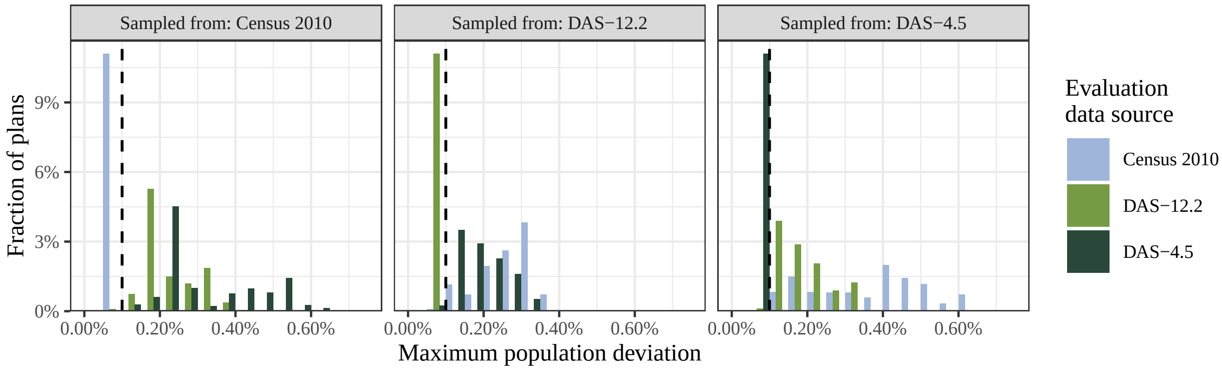

Figure 1: *Maximum deviation from population parity among Pennsylvania redistricting plans simulated from the three data sources. All plans were sampled with a population constraint of 0.1 percent, corresponding to the deviation measured from the Census 2010 precinct data, and marked with the dashed line. Deviation from parity was then evaluated using the three versions of population data.*

deviation from parity to be significantly larger than what they can observe in their data. This problem is more acute in state legislative districts, where there are more districts and each district is comprised of fewer precincts.

4.1 Congressional Districts in Pennsylvania

Figure 1 shows the maximum deviation from population parity for the 30,000 simulated redistricting plans in Pennsylvania, when evaluated according to the three different data sources.[2] Consistently, plans generated under one set of population data and drawn to have a maximum deviation of no more than 0.1% had much larger deviations when measured under a different set of population data. For example, of the 10,000 maps simulated using the DAS-12.2 data, 9,915 exceeded the maximum population deviation threshold, according to the Census 2010 data. While nearly every plan failed to meet the population deviation threshold, the exact amount of error varied significantly across the simulation set. As a result, redistricting practitioners who attempt to create equal-population districts according to similar thresholds can expect the actual deviation from parity to be significantly larger but of unknown magnitude.

4.2 State Legislative Districts in Louisiana

We expect smaller districts such as state legislative districts to be more prone to discrepancies in population parity. For example, the average Louisiana Congressional district comprises about 600 precincts, but a State Senate district comprises about 90 and a State House district only 35. Therefore, deviations due to DAS are more likely to result in larger percent deviations from the average. To test this, we compared 60,000 Louisiana State Senate plans generated from the three data sources and population parity constraints ranging from 0.1% to 50%, measuring the plans' population deviation against the three different data sources.[3] Figure 2 plots the results of this comparison.

As expected, we see complete acceptance for plans measured with the dataset from which they were generated. However, plans generated under one dataset can be invalid under another. Specifically, plans generated under DAS data can be very likely to be invalid when evaluated using the true Census data. The rate of invalid plans grows as the tolerance becomes more precise.

Also noteworthy is the fact that even at the population parity tolerances as generous as 1.0%, all generated plans are invalid in some cases. Compared to Pennsylvania, with a parity tolerance of 0.1%, this is as a result

[2] 10,000 plans were simulated from each data source, with every plan satisfying a 0.1% population parity constraint. The simulation algorithm also ensured that no more than 17 counties were split across the entire state, reflecting the requirement in Pennsylvania that district boundaries align with the boundaries of political subdivisions to the greatest extent possible.

[3] 2,500 plans were simulated for each data source/population parity pair.

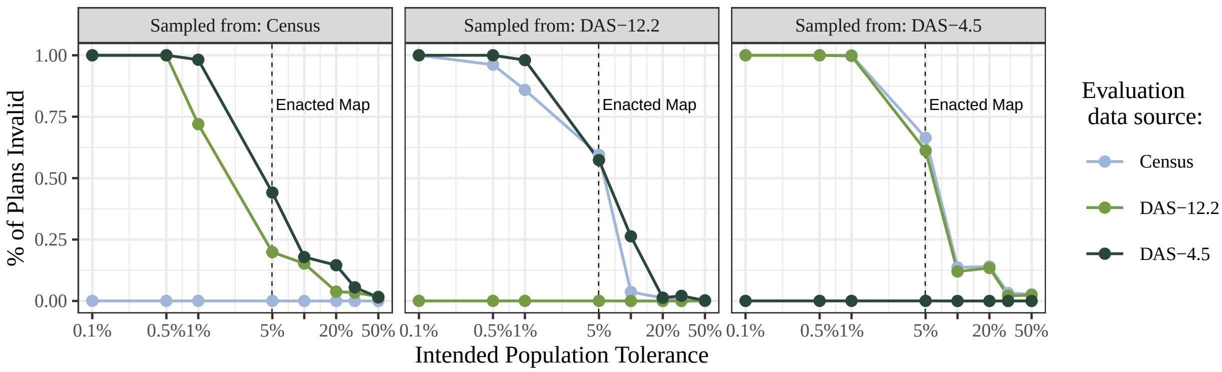

Figure 2: *Fraction of Louisiana State Senate plans simulated under one data source which are invalid when measured under another. The dashed line shows the parity of the enacted 2010 map.*

of the smaller district sizes in the Louisiana State Senate—the DAS-added noise is relatively larger at smaller scales.

5 Partisan Effects on Redistricting

To analyze the partisan implications of a redistricting plan using a set of simulated redistricting plans, practitioners generate hypothetical district-level election results for the simulated plans and for the plan to be analyzed. Plans which are partisan gerrymanders stand out from the simulated ensemble as yielding more seats for one party over the other.

In computing a party's expected vote share for each congressional district, we use data from statewide elections to avoid the variation in uncontested races and any incumbency effects in U.S. House races. In Pennsylvania, we use the two party vote share averaged across all statewide and Presidential races, 2004–2008, and adjust to match 2008 turnout levels. In South Carolina we use the 2018 gubernatorial election, in North Carolina we use the 2012 gubernatorial election, and in Louisiana we use the 2019 Secretary of State election.

5.1 Partisan Patterns in DAS-induced Population Error

We first examine the electoral correlates of population change induced by the DAS. By the nature of the noise injection of the DAS, there is significant variation in the population error, even among similar precincts, and as a result it is difficult to discern systematic patterns by observation alone. Consequently, we fit a generalized additive model (GAM) to the precinct-level population errors to understand the degree to which different factors influence these errors, on average. The GAM regresses the difference in precinct population between the DAS-12.2 and the Census data on a tensor product cubic regression smooth of precinct turnout, two-way Democratic vote, and log population density, and thin-plate regression splines of the fraction of voters who are White and the racial Herfindahl-Hirschman index [17, 18]. We fit the GAM on precincts in Pennsylvania, North Carolina, South Carolina, and Louisiana. The model explained about 9–12 percent of the overall variance in population errors.

Figure 3 plots the fitted values from this model against Democratic vote share for each of the four states. Perhaps unexpectedly, several consistent patterns emerge. First, higher-turnout precincts are on average assigned more population under the DAS than they should otherwise have, according to the 2010 Census. Second, moderately Democratic precincts are on average assigned less population under the DAS. These effects are on the order of 5–15 voters per precinct, on average, though some are larger.[4] Aggregated across the hundreds of precincts that comprise the average district, however, the errors may become substantial. In Pennsylvania's 2nd and 3rd Congressional Districts, for example, which cover Philadelphia County and are majority-minority, the accumulated population error in each district is on average 3,000 voters across the set of simulated plans.

[4] Not shown is the equivalent figure for the DAS-4.5 data, which displayed an identical pattern but with roughly double the magnitude of fitted error.

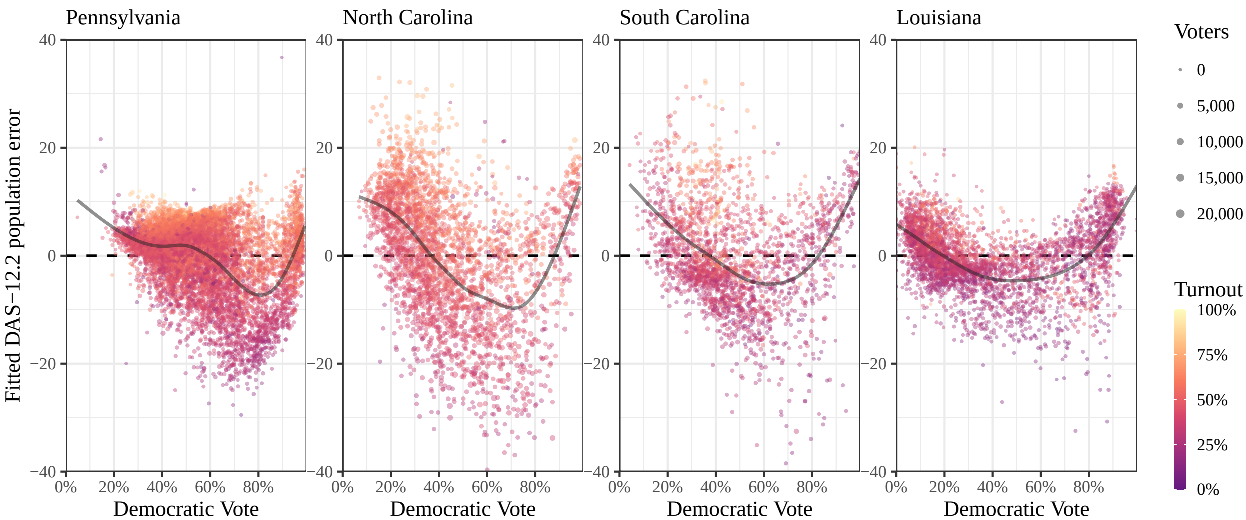

Figure 3: *Model-smoothed error in precinct populations by Democratic two-party vote share, with color indicating turnout. A GAM smooth is overlaid to show the mean error by Democratic share.*

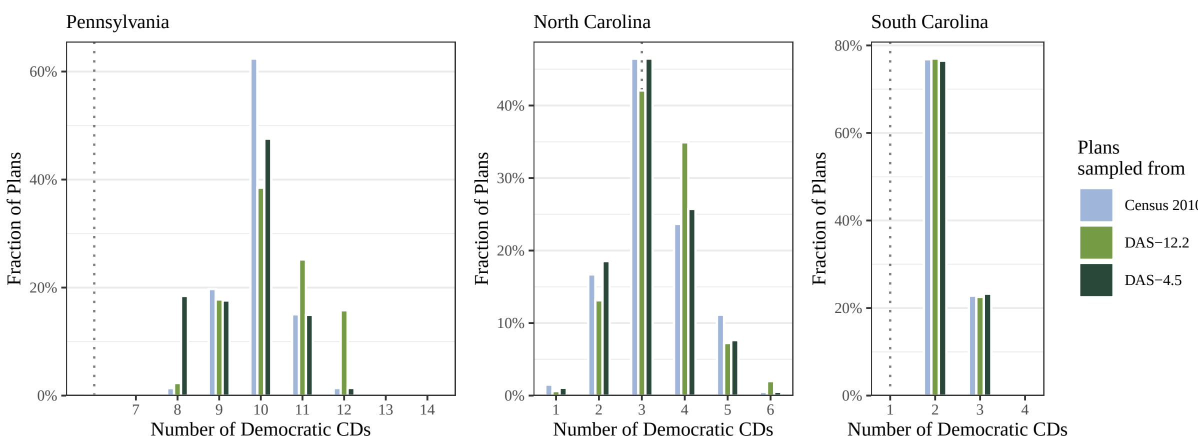

Figure 4: *Distribution of Democratic-majority congressional districts, by data source and state. The vertical dashed lines indicated the number of Democratic-majority seats under the plans enacted by the state legislatures.*

Some of these partisan effects may be explained by racial patterns, as shown in Figure 6 and discussed below in Section 6. It is difficult to know exactly these partisan and racial biases arise without more detail on the DAS post-processing system and parameters. Regardless, the presence of differential bias in the precinct populations according to partisanship and turnout is concerning. These precinct-level biases may aggregate in unexpected ways, leading to potentially large unknown biases in statewide analyses, as we discuss next.

5.2 Effects of Partisan Patterns on Aggregate Results

The spatial distribution of these types of precincts, and the details of the DAS post-processing, critically determine the overall effect once these precincts are aggregated into larger districts. Given the results of Figure 3, we would expect that aggregation to districts may not cancel out DAS-induced noise entirely. Indeed, for the 44 congressional districts in the four states we examine, the average district's population changes by 292 people (or 1%) by DAS-12.2 data, but in three Pennsylvania congressional districts in and around Philadelphia, the population changes by 1,311 people on average. Two congressional races in these

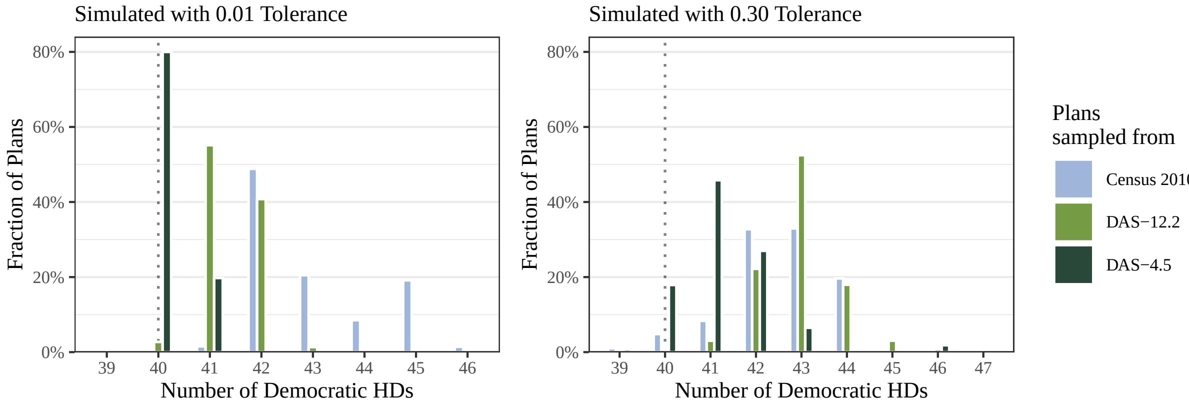

Figure 5: *Distribution of Democratic-majority South Carolina State House districts.*

four states have been decided by less than a percentage point during 2012-2020: NC-07 in 2012 and NC-09 in 2018.

We find that the DAS leads to unpredictable differences in the distribution state-level party outcomes under the three data sources. Figure 4 compares the distribution of the number of congressional districts in which the Democratic Party's candidate wins over 50% of the two-party vote.[5] In Pennsylvania and North Carolina, plans simulated with DAS-12.2 tend to favor the Democratic party more than plans simulated with DAS-4.5 or the original Census. The implied number of Democratic seats in the *enacted* plans, shown in the dotted line, tend to be on the lower end of the simulated reference distribution, although our simulations here do not impose constraints required by the Voting Rights Act.

Interestingly, with congressional districts, the DAS-4.5 data tend to produce a distribution of Democratic seats closer to the 2010 Census, even though it is noisier than DAS-12.2 on average. We caution that the number of congressional districts with majority Democratic vote is a coarse measure and can mask more subtle differences. For example, in South Carolina, the overall distribution of Democratic seats does not differ, but this may mask differences captured by other continuous metrics like mean-median difference in voteshares.

Differences between data sources are likely more stark for state legislative districts, which are composed from fewer precincts than the congressional districts. In Figure 5 we show simulations from the state legislative districts in South Carolina. We show two simulations with different tolerances for deviations from population parity.

Once again, there are significant differences in the distribution of Democratic seats across the three data sources, but the pattern in location and scale changes are not monotonic with the level of noise. Notably, at a 1% population parity constraint, the enacted legislative map is an outlier under the Census 2010 and DAS-12.2 simulations, but is the modal outcome under the DAS-4.5 data. A discrepancy of this magnitude could change the factual findings regarding the presence or absence of a partisan gerrymander in redistricting litigation.

6 Racial Effects on Redistricting

We also investigate the potential impact of privacy-protected data on the role of race in redistricting. We begin by the analysis of racial patterns in the population errors induced by the DAS. We then examine how those racial biases affect redistricting outcomes.

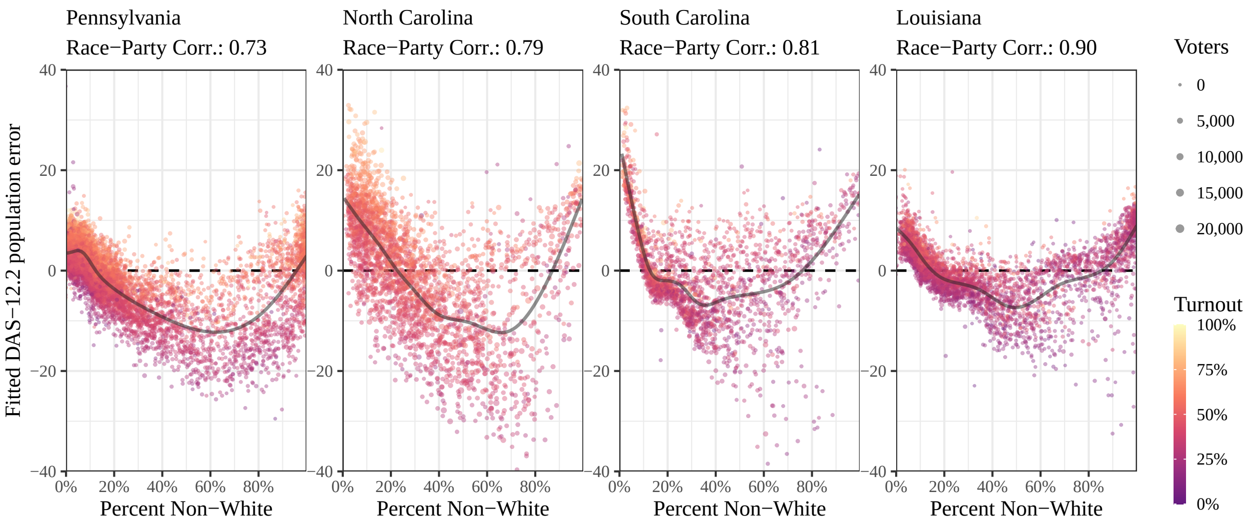

Figure 6: *Model-smoothed error in precinct populations by the minority fraction of voters, with color indicating turnout. A GAM smooth is overlaid to show the mean error by minority share.*

6.1 Racial Patterns in DAS-induced Population Error

In the previous section, we demonstrated that the population error introduced by the DAS procedure overcounts the most homogeneous Republican and Democratic precincts in high-turnout areas and undercounts heterogeneous, low-turnout areas. Race is highly correlated with partisanship in American politics, and we find that the same pattern of differential error by race and turnout levels holds for race as well as partisanship. Figure 6 shows this pattern across the states we have analyzed so far (PA, LA, NC, and SC). The results imply that in terms of population error, mixed White/nonwhite precincts lose the most population relative to more homogeneous precincts. Figure 7 more clearly shows this pattern with homogeneous precincts. We plot the error against the Herfindahl-Hirschman Index and find that the fitted error in estimated population steeply declines as the precinct becomes more racially diverse.

These patterns are likely partially explained by the adopted DAS targets [14], which prioritize accuracy for the largest racial group in a given area. By doing so, the DAS procedure appears to undercount heterogeneous areas where the population differences between racial groups are relatively small. As precincts are the building blocks of political districts, our results demonstrate that precincts that are heterogeneous along racial and partisan lines would lose electoral power under the DAS. In aggregate, the movement of population from heterogeneous to homogeneous precincts would tend to increase the apparent spatial segregation by race.

6.2 Effects of Racial Patterns on Aggregate and Precinct-level Results

As with the partisan patterns of DAS population bias, the racial patterns of bias may not necessarily cancel upon aggregation. To evaluate the impact of these biases, we compare the distribution of the number of majority-minority districts (MMDs) across the simulations from the three data sources. MMDs are a primary focus in voting rights litigation and the analysis of race in redistricting.

Figure 8 shows the effects of the DAS on the number of MMDs in the South Carolina state House and Mississippi state Senate.[6] Ten thousand plans simulated from both 2010 Census and DAS-12.2 data were evaluated for MMDs under both data sources. There are two types of discrepancies. Not visible in the figure is the fact that while generally the DAS and 2010 Census data agree on the presence of an MMD given a set of simulated plans, the DAS data slightly but systematically understate the number of such districts in

[5]Simulations in North Carolina and South Carolina shown here satisfy a 1% population parity constraint, and ensure that no more than 12 and 6 counties, respectively, are split in each plan. Data for North Carolina was obtained from the North Carolina General Assembly Redistricting Archives.

[6]The Mississippi plans were generated to satisfy a 5.0% population parity constraint, reflecting the 4.98% population parity deviation of the currently enacted plan. Data for Mississippi was obtained from the Mississippi Automated Resource Information System.

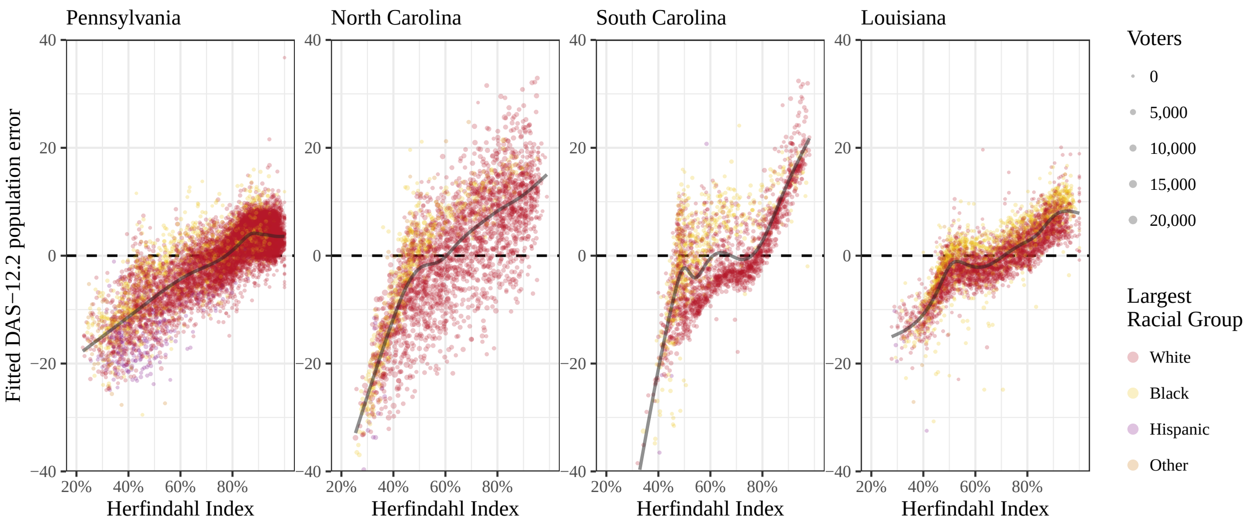

Figure 7: *Model-smoothed error in precinct populations by the Herfindahl-Hirschman Index. A Herfindahl-Hirschman Index of 100 percent indicates that the precinct is comprised of only one racial group.*

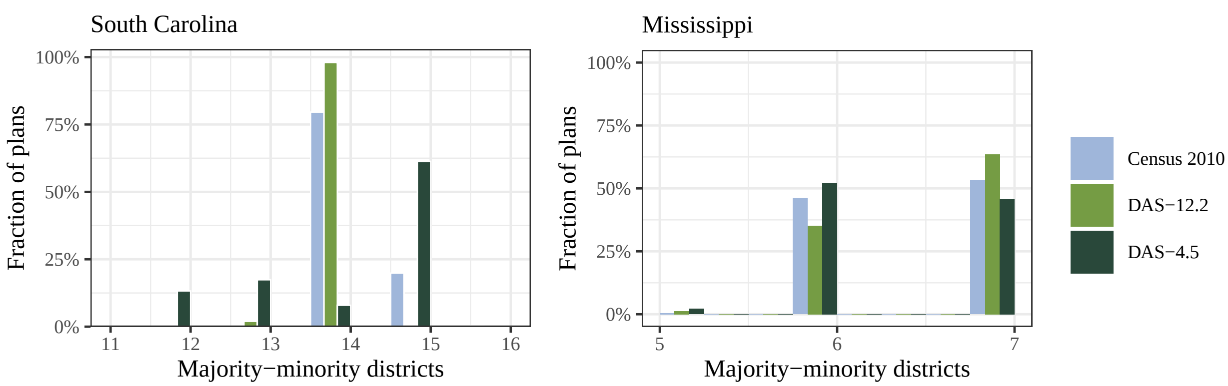

Figure 8: *Distribution of majority-minority districts in South Carolina and Mississippi, by simulation data source.*

South Carolina and overstate the number of such districts in Mississippi. For example, in South Carolina, among the 1,986 plans simulated from 2010 Census data that had 15 MMDs, 4.7% had only 14 MMDs when evaluated with DAS-12.2 data.

What is more concerning is that the overall distribution of the number of MMDs is significantly different across data sources. In Mississippi, the DAS-12.2 data generates far fewer plans with 6 MMDs compared to the 2010 Census data. In South Carolina, meanwhile, there are no simulated plans with 15 MMDs under DAS-12.2 data, but such plans make up nearly 20% of the 2010 Census-based simulations. As a result, a legislature-adopted plan drawn with 15 MMDs according to DAS-protected data could be improperly classified as an extreme outlier and might even be struck down as a racial gerrymander.

If these differences between DAS-based and 2010 Census-based summary statistics were of predictable magnitude, it might be possible for states or analysis to adjust to the additional noise. However, as with the partisan effects, we find that the DAS-induced distortions are not necessarily consistent across states. Our primary case for this purpose is Louisiana's East Baton Rouge Parish and the surrounding area. We chose this area because the city of Baton Rouge includes a large Black population represented by multiple MMDs in the state's lower and upper houses. From the 15 lower house districts in this area (each with approximately 40,000 population) comprising 361 precincts, we simulate 500,000 plans under each of the three data sources. We simulate each plan with a maximum 5% population parity constraint to match the enacted map. For each

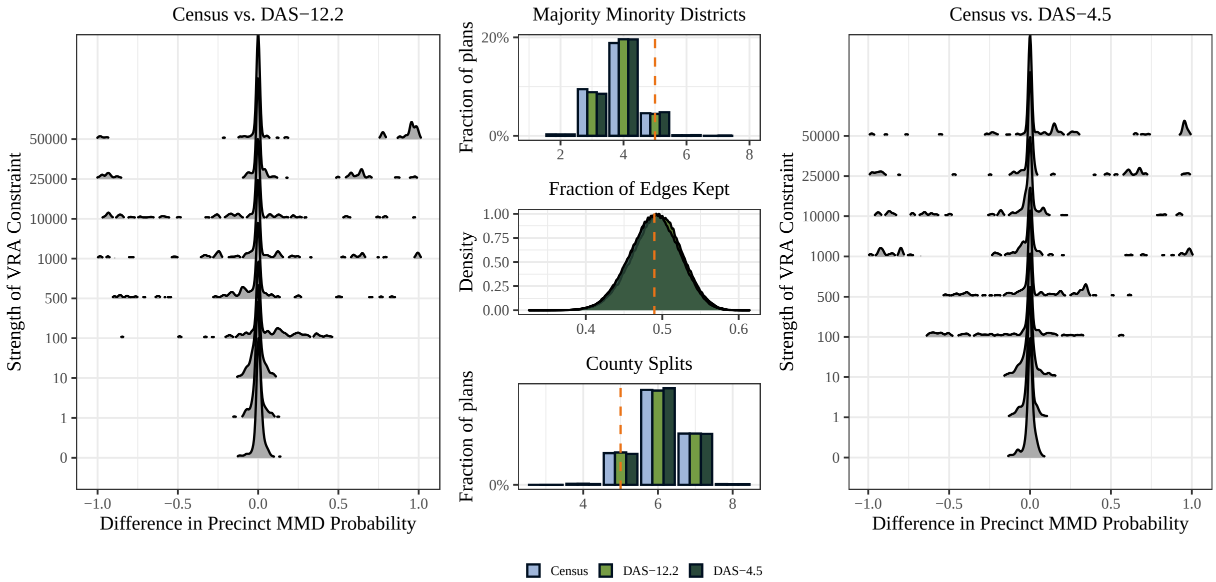

Figure 9: *The center column shows district-level comparisons between 500,000 plans generated under 2010 Census data, DAS-4.5 data, and DAS-12.2 data. Few aggregate-level differences are seen across three commonly used metrics—the number of majority-minority districts, the number of parish (county) splits, and the compactness of the districts. However, the left and right columns show that precinct-level assignments can differ substantially between the 2010 Census and DAS data. Here, the calculated probability of being assigned to a majority-minority district can be much higher or lower for individual precincts, and these differences grow as a constraint encouraging the formation of MMDs is strengthened.*

of these plans, we measure three commonly used metrics in redistricting—the number of resulting MMDs, the number of parish splits, and the compactness of the plan.

The middle column of Figure 9 finds few district-level differences between plans generated using 2010 Census data versus DAS data. Plans generated under all three datasets have essentially identical distributions of MMDs, parish splits, and compactness.

However, these aggregate distributions mask the variability around which individual precincts are included in majority minority districts. In the left and right columns of Figure 9, we show the results of 10,000 simulations of the Merge–Split-type MCMC sampler with various levels of a Voting Rights Act (VRA) constraint. This constraint, which we did not apply in the previous sections, encourages the formation of majority-minority districts. We then calculate the probability that each precinct is assigned to a majority-minority district (as defined by Black population). Finally, we calculate the difference between these probabilities for the Census versus DAS-12.2 and Census versus DAS-4.5.

With no VRA constraint, each precinct has similar probabilities of being in a MMD, regardless of the dataset used. However, as the strength of this constraint increases (making the algorithm search for MMDs more aggressively), we see that the noise introduced to the DAS data systematically alters the district membership of individual precincts. A precinct with a value of 1 or -1 in the left and right columns of Figure 9 indicates that those precincts are never in a MMD under one dataset but are always in a MMD when the same mapmaking process is done with a different dataset.

7 Ecological Inference and Voting Rights Analysis

Inferring the racial and ethnic composition of potential voters and their candidate choice is a key element of voting rights analysis in redistricting. Recent court cases have relied on Bayesian Improved Surname Geocoding (BISG) to predict the race and ethnicity of individual voters in a voter file [11, 12, 13]. This methodology combines the names and addresses of registered voters with block-level racial composition data from the Census.

We first examine how the accuracy of prediction changes between the DAS and original Census data. Since this is exactly the type of analysis from which the DAS is supposed to protect individual Census respondents, we expect the prediction accuracy to dramatically decline when using the DAS-protected data. We then revisit the most recent court case about the East Ramapo school board election and investigate whether this change in racial prediction alters the conclusions of the racial redistricting analysis.

7.1 Prediction of Individual Voter's Race and Ethnicity

We first compare the accuracy of predicting individual voters' race and ethnicity using the original 2010 Census data, the DAS-12.2 data, and the DAS-4.5 data. To obtain the benchmark, we use the North Carolina voter file obtained in February 2021.[7] In several southern states including North Carolina,[8] the voter files contain the self-reported race of each registered voter. This information can then be used to assess the accuracy of the BISG prediction methodology.

Our approach follows [19]. We denote by E_i the ethnicity of voter i, N_i as the surname of voter i, and G_i as the geography in which voter i resides. For each choice of ethnicity $e \in \mathcal{E} = \{\text{White, Black, Hispanic, Asian, Other}\}$, Bayes' rule implies

$$P(E_i = e \mid N_i = n, G_i = g) \;=\; \frac{\Pr(N_i = n \mid E_i = e)\Pr(E_i = e \mid G_i = g)}{\sum_{e' \in \mathcal{E}} \Pr(N_i = n \mid E_i = e')\Pr(E_i = e' \mid G_i = g)},$$

where we have assumed the conditional independence between the surname of a voter and their geolocation within each racial category, i.e., $N_i \perp\!\!\!\perp G_i \mid E_i$.

In the presence of multiple names—e.g. first name f, middle name m, and surname s—we make the further conditional independence assumption [20]

$$\Pr(N_i = \{f, m, s\} \mid E_i = e) \;=\; \Pr(F_i = f \mid E_i = e)\Pr(M_i = m \mid E_i = e)\Pr(S_i = s \mid E_i = e),$$

where F_i, M_i, and S_i represent individual i's first, middle, and surnames respectively.

We compare estimates by changing the data source from which the geographic prior, $\Pr(E_i = e \mid G_i = g)$, is estimated, from the 2010 Census to each of the two DAS datasets. Estimates of the other race prediction probabilities are obtained by merging three sources: the 2010 Census surname list [21], the Spanish surname list from the Census, and the voter files from six states in the U.S. South, where state governments collect racial and ethnic data about registered voters for Voting Rights Act compliance. The middle and first name probabilities are derived exclusively from the voter files.

We evaluate the accuracy of the BISG methodology on approximately 5.8 million registered voters included in the North Carolina February 2021 voter file. Among them, approximately 70% are White and 22.5% are Black, with smaller contingents of Hispanic (3.4%), Asian (1.5%), and Other (2.4%) voters.

Figure 10 summarizes the accuracy of the race prediction with the area under the Receiver Operating Characteristic curve (AUROC). The AUROC ranges from 0 (perfect misclassification) to 1 (perfect classification). Across all racial and ethnic groups except Hispanics, we find the same surprising pattern: relative to the 2010 Census data, the DAS-12.2 data yield a small improvement in prediction performance while the DAS-4.5 data give a slight degradation. Among Hispanics, both forms of DAS-protected data result in slightly improved predictions over the original Census data.

The strong performance of the DAS-12.2 data in this setting is counter-intuitive. It is possible that the noise added to the underlying data has somehow mirrored the true patterns of population shift from 2010 to 2021; or that this noise makes the DAS-12.2 data more reflective of the voter population relative to the voting-age population. Additionally, the DAS may degrade or attenuate individual probabilities without having a significant impact on the overall ability to classify, something that AUROC is not designed to measure [22].

Results are substantively similar if we consider the classification error, under the heuristic that we assign each individual to the ethnic group with the highest posterior probability. Using the true census data to establish

[7]We obtain the voter files used in this paper through L2, Inc., which is a leading national nonpartisan firm that supplies voter data and related technology.

[8]The other states are Alabama, Florida, Georgia, Louisiana, and South Carolina

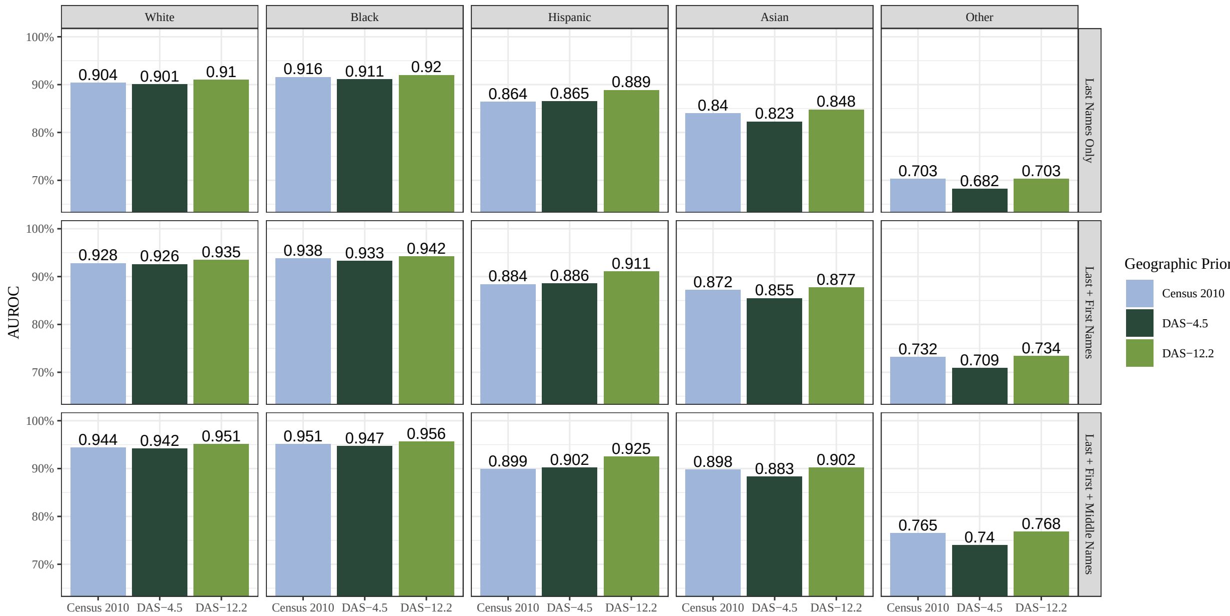

Figure 10: *Area under the Receiver Operating Characteristic Curve (AUROC) percentage values for the prediction of individual voter's race and ethnicity using North Carolina voter file. Bars represent AUROC with geographic priors given by each of three datasets: 2010 Census, DAS-4.5, and DAS-12.2.*

geographic priors, we achieve posterior misclassification rates of 15.1%, 12.1%, and 10.0% when using the last name; last name and first name; and last, first, and middle names for prediction, respectively. The analogous misclassification rates are slightly higher for the DAS-4.5 priors—15.6%, 12.5%, and 10.3%—but the same or slightly lower for the DAS-12.2 priors: 15.1%, 12.0%, and 9.9%.

Our analysis shows that across three main racial and ethnic groups, the predictions based on the DAS data appear to be as accurate as those based on the 2010 Census data. The finding suggests that the DAS data may not provide universal privacy protection.

7.2 Ecological Inference in the Voting Rights Analysis

The BISG methodology played a central role in the most recent court case regarding Section 2 of the Voting Rights Act, *NAACP of Spring Valley v. East Ramapo Central School District* (2020). The East Ramapo Central School District (ERCSD) nine-member school board was elected using at-large elections. This often led to an all White school board, despite 35% of the voter eligible population being Black or Hispanic. Yet, within the district, nearly all White school children attend private yeshivas, whereas nearly all Black and Hispanic children attend the ERCSD public schools. As a result of this case, the district moved to a ward system.

We re-examine the remedy of this case, focusing on effective majority-minority districts (MMDs) based on a voter file with individual race and ethnicity imputed using the DAS-12.2 and Census 2010 data. To approximate the data used by an expert witness who testified in the court case, we obtain the New York voter file (as of November 16, 2020) from the state Board of Elections. We subset the voters to active voters with addresses in Rockland County, where ERCSD is located. Using the R package `censusxy`, which interfaces with the Census Bureau's batch geocoder, we match each voter to a block and subset the voters to those who live within the geographic bounds of ERCSD [23, 24]. This leaves 58,253 voters, for whom we impute races using the same machinery behind the R package `wru` [25], as described in [19]. This process nearly exactly mimics the one used in the original case.

We examine how the predictions of individual race and ethnicity based on the 2010 Census and DAS-12.2 data result in different redistricting outcomes. Figure 11 compares these two predictions using the proportions of (predicted) Whites, Black, and Hispanic registered voters for each Census block. We find that the predictions based on the DAS-12.2 tend to produce blocks with more White voters than those based on the original

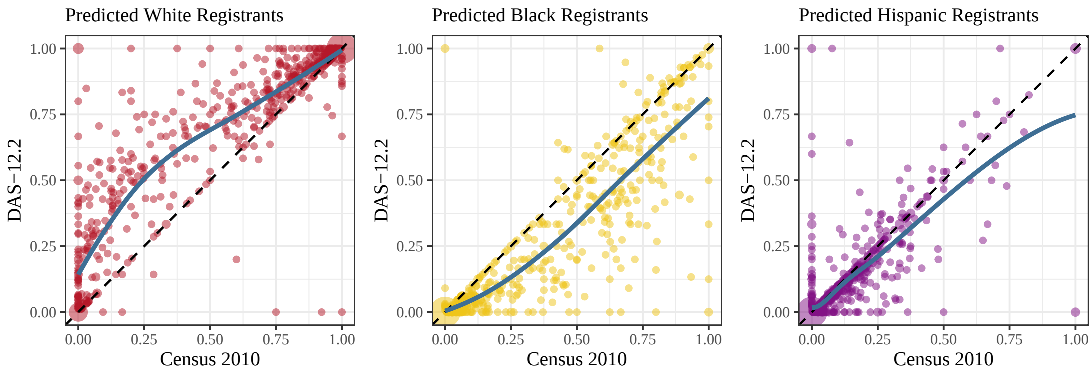

Figure 11: *Imputed Racial Registrants by Census Blocks. The x-axis represents the percent of a group, as measured by the most likely race from racial imputation using the Census 2010 data. The y-axis represents the corresponding imputation using the DAS-12.2 data.*

Table 2: *East Ramapo MMDs under Census 2010 and DAS-12.2 data. The noise introduced in the DAS-12.2 leads us to undercount the number of majority minority districts in many plans, but never to overcount them.*

	Number of MMDs from DAS-12.2				
Census 2010	0	1	2	3	Plans
0	**100%**	0	0	0	2
1	2	**98**	0	0	3,581
2	2	40	**59**	0	6,311
3	6	76	18	**0**	106

Note: Percentages add to 100% by row.

Census data. As a consequence, the predicted proportions of Black and Hispanic registrants are much smaller, especially in the blocks where they form a majority group.

The precise reason for these biases is unclear. The DAS tends to introduce more error for minority groups than for White voters, and even more error for voters who are in a minority group for their Census block, which is more common for minority voters as well. This additional noise, when carried through a nonlinear transformation such as the Bayes' rule calculation for racial imputation, may introduce some bias. In addition, the large bias for White and Black voters relative to Hispanic voters suggests that the similarity of surnames between the White and Black populations, compared to the Hispanic population, may also be a factor. Regardless, it is clear that the DAS-injected noise differentially biases voter race imputations at the block level. This pattern may not always yield greater inaccuracies when aggregated to the statewide level—as seen in the prior section—but it is especially prevalent within the ERCSD.

We next investigate whether these systematic differences in racial prediction lead to different redistricting outcomes. Specifically, we simulate 10,000 redistricting plans using DAS-12.2 population and a 5% population parity tolerance. We find that the systematic differences in racial prediction identified above results in the underestimation of the number of MMDs in these plans. As in the original court case, an MMD is defined as a district, in which more than 50% of its registered voters are either Black or Hispanic. Table 2 clearly shows that the number of MMDs based on the DAS-12.2 data never exceeds that based on the 2010 Census for all simulated plans. For example, among 6,311 plans that are estimated to yield 3 MMDs according to the Census data, nearly 60% of them are predicted to have 2 MMDs.

While one should not extrapolate from this single case study, our analysis implies that in small electoral districts such as those of school board elections, the DAS can generate bias that may favor one racial group over another. Although the number of MMDs is underestimated under the DAS data in this case, the direction and magnitude of racial effects are difficult to predict, as they depend on how the choice of tuning parameters in the DAS algorithm interact with a number of geographical and other factors. At a minimum, this poses a serious challenge in ensuring the effective number of MMDs using DAS-protected data.

8 Recommendations

These empirical findings lead to our primary recommendation: release Census P.L. 94-171 data without using the current Disclosure Avoidance System (DAS), and instead rely on a swapping method similar to that applied to the 2010 Census data. Over the past half century, the Supreme Court has firmly established the principle of *One Person, One Vote*, requiring states to minimize the population difference across districts based on the Census data. Our analysis shows that the DAS makes it impossible to follow this basic principle. The only solution is to make Census-block populations invariant, but doing so within the current DAS would, in the Bureau's own admission, require injecting far too much noise into Census tabulations other than total population [26].

We also find that the DAS introduces partisan and racial biases into local data, which may aggregate into large and unpredictable biases at the state level. Since many federal and local elections have narrow margins of victory, relatively small changes to the Census data can result in redistricting plans that produce favorable electoral outcomes for certain candidates and parties. Similarly, these changes affect the number of majority minority districts, either hampering or artificially inflating the voting power of minority groups.

One may argue that the protection of privacy is a worthy cause, and outweighs these potentially negative consequences. Unfortunately, the DAS algorithm fails to universally protect respondent privacy. We are able to predict the individual race of registered voters at least as accurately using the DAS-protected data as when using the original Census data. In sum, we find that the DAS negatively impacts the redistricting process and voting rights of minority groups without providing clear benefits.

If the Census Bureau decides to apply the current DAS to Census P.L. 94-171 data, our recommendation is to increase the privacy loss budget and allocate the increase to improving redistricting outcomes. In addition, the Bureau may consider publishing fewer block-level cross-tabulations in other Census products to ensure more accuracy in the P.L. 94-171 files. In allocating any increased privacy loss budget, we recommend minimizing the change in population at the voting tabulation district (VTD) level. Ensuring that population is accurate at this off-spine geography would help minimize population deviations among the overwhelming majority states which rely on these geographies to draw their districts. This would not fix the problem of ensuring near-exact population equality, but it would help to minimize extreme outliers. In our VTD-level population tabulations, we find that there is around a 1% average deviation in the DAS-12.2 data compared to the 2010 Census data. We recommend aiming for at most a 0.1% average deviation.

Furthermore, we recommend adjusting the parameters of the DAS to address the current demonstrated bias against racially integrated, diverse blocks, and low-turnout areas. Without more detail on the current parameters and workings of the DAS post-processing system, it is difficult to provide more specific recommendations. However, it is vital for the Bureau to ensure that it is not injecting racial and partisan bias into the privacy-protected data.

Finally, should the DAS be used, the Bureau should publish additional information on the known inaccuracies. The current information provided by the Census Bureau with the April PPMF data release provides only marginal distributions of variables, with a focus on total population data. For example, root mean squared error (RMSE) for urban and rural block populations is reported, but these statistics are not cross-tabulated by race or other relevant variables. Reports on inaccuracies and impossibilities must reflect the important relationship that this data has with race, age, population density, and total population. The burden of privacy must not be paid fully by some races or age groups.

References

[1] John M. Abowd, Gary L. Benedetto, Simson L. Garfinkel, Scot A. Dahl, Aref N. Dajani, Matthew Graham, Michael B. Hawes, Vishesh Karwa, Daniel Kifer, Hang Kim, Philip Leclerc, Aashwin Machanavajjhala, Jerome P. Reiter, Rolando Rodriguez, Ian M. Schmutte, William N. Sexton, Phyllis E. Singer, and Lars Vilhuber. The modernization of statistical disclosure limitation at the U.S. Census Bureau. 2020.

[2] Steven Ruggles, Catherine Fitch, Diana Magnuson, and Jonathan Schroeder. Differential privacy and census data: Implications for social and economic research. *AEA papers and proceedings*, 109:403–408,

2019.

[3] David Van Riper, Tracy Kugler, and Steven Ruggles. Disclosure avoidance in the Census Bureau's 2010 demonstration data product. In *Privacy in Statistical Databases*, Lecture Notes in Computer Science, pages 353–368. Springer International Publishing, Cham, 2020.

[4] Jowei Chen and Jonathan Rodden. Unintentional gerrymandering: Political geography and electoral bias in legislatures. *Quarterly Journal of Political Science*, 8(3):239–269, 2013.

[5] Daniel Carter, Gregory Herschlag, Zach Hunter, and Jonathan Mattingly. A merge-split proposal for reversible Monte Carlo Markov chain sampling of redistricting plans. *arXiv preprint arXiv:1911.01503*, 2019.

[6] Daryl DeFord, Moon Duchin, and Justin Solomon. Recombination: A family of Markov chains for redistricting. *arXiv preprint arXiv:1911.05725*, 2019.

[7] Eric Autry, Daniel Carter, Gregory Herschlag, Zach Hunter, and Jonathan Mattingly. Multi-scale merge-split Markov chain Monte Carlo for redistricting. *arXiv preprint arXiv:2008.08054*, 2020.

[8] Benjamin Fifield, Michael Higgins, Kosuke Imai, and Alexander Tarr. Automated redistricting simulator using Markov chain Monte Carlo. *Journal of Computational and Graphical Statistics*, 29(4):715–728, 2020.

[9] Cory McCartan and Kosuke Imai. Sequential Monte Carlo for sampling balanced and compact redistricting plans. *arXiv preprint arXiv:2008.06131*, 2020.

[10] Christopher T. Kenny, Cory McCartan, Ben Fifield, and Kosuke Imai. redist: Simulation methods for legislative redistricting. Available at The Comprehensive R Archive Network (CRAN), 2021.

[11] Kevin Fiscella and Allen M. Fremont. Use of geocoding and surname analysis to estimate race and ethnicity. *Health Services Research*, 41(4p1):1482–1500, August 2006.

[12] Marc N. Elliott, Peter A. Morrison, Allen Fremont, Daniel F. McCaffrey, Philip Pantoja, and Nicole Lurie. Using the Census Bureau's surname list to improve estimates of race/ethnicity and associated disparities. *Health Services and Outcomes Research Methodology*, 9(2):69–83, 2009.

[13] Kosuke Imai and Kabir Khanna. Improving ecological inference by predicting individual ethnicity from voter registration record. *Political Analysis*, 24(2):263–272, Spring 2016.

[14] United States Census Bureau. Meeting redistricting data requirements: Accuracy targets. `https://content.govdelivery.com/accounts/USCENSUS/bulletins/2cb745b`, 2021.

[15] Leo A. Goodman. Ecological regressions and behavior of individuals. *American Sociological Review*, 18(6):663–664, 1953.

[16] Gary King, Ori Rosen, and Matin Tanner, editors. *Ecological Inference: New Methodological Strategies*. Cambridge University Press, Cambridge, 2004.

[17] Orris C Herfindahl. *Concentration in the steel industry*. PhD thesis, Columbia University, 1950.

[18] Albert O Hirschman. *National power and the structure of foreign trade*, volume 105. Univ of California Press, 1945.

[19] Kosuke Imai and Kabir Khanna. Improving ecological inference by predicting individual ethnicity from voter registration records. *Political Analysis*, pages 263–272, 2016.

[20] Ioan Voicu. Using first name information to improve race and ethnicity classification. *Statistics and Public Policy*, 5(1):1–13, 2018.

[21] United States Census Bureau. Frequently Occurring Surnames from the 2010 Census. `https://www.census.gov/topics/population/genealogy/data/2010_surnames.html`, 2016.

[22] Jorge M Lobo, Alberto Jiménez-Valverde, and Raimundo Real. AUC: a misleading measure of the performance of predictive distribution models. *Global ecology and Biogeography*, 17(2):145–151, 2008.

[23] Christopher Prener and Branson Fox. *censusxy: Access the U.S. Census Bureau's Geocoding A.P.I. System*, 2021. R package version 1.0.1.

[24] Christopher G. Prener and Branson Fox. Creating open source composite geocoders: Pitfalls and opportunities. *Transactions in GIS*, n/a(n/a), 2021.

[25] Kabir Khanna and Kosuke Imai. *wru: Who are You? Bayesian Prediction of Racial Category Using Surname and Geolocation*, 2021. R package version 0.1-12.

[26] John M. Abowd. Email to James Whitehorne. Exhibit 1 of Plaintiffs' reply in support of combined motion for a preliminary injunction and petition for a writ of mandamus, *Alabama v. U.S. Dep't of Commerce*, No. 3: 21-cv-211-RAH-KFP, 2021.