STUDY SERIES
(Statistics #2021-02)

Empirical Study of Two Aspects of the Topdown Algorithm Output for Redistricting: Reliability & Variability (August 5, 2021 Update)

Tommy Wright,
Kyle Irimata

Center for Statistical Research & Methodology
Research and Methodology Directorate
U.S. Census Bureau
Washington, D.C. 20233

Report Issued: August 5, 2021

EMPIRICAL STUDY OF TWO ASPECTS OF THE TOPDOWN ALGORITHM OUTPUT FOR REDISTRICTING: RELIABILITY & VARIABILITY

(August 5, 2021 Update)

Tommy Wright and Kyle Irimata
Center for Statistical Research and Methodology
Research and Methodology Directorate
U.S. Bureau of the Census
Washington, D.C. 20233

Abstract

This two-part study provides an update of empirical results for ongoing research and development that were reported in [6]. In this update, data output from the same version of the *TopDown Algorithm* that will produce the "2020 Census Redistricting Data (Public Law 94-171) Summary File" are reported in the tables and figures. Except for wording changes due to changes in the data output, the wording throughout is the same as in [6]. The *TopDown Algorithm (TDA)* [1] is being used to protect the confidentiality of respondent data collected during the 2020 Census. Following the 2010 Census, the swapping methodology *(SWA)* [7] was applied to respondent data to protect confidentiality.

In Part I, we propose an empirically based solution to the question: "What is the minimum TOTAL population of a district to have reliable characteristics of various demographic groups?" To answer this question, we use data treated by the 2020 Census redistricting data production settings version ($\epsilon = 17.14$, for the person file) of the *TDA* for all block groups (proxy for districts) in the United States. We also consider "places and minor civil divisions (MCDs)" as proxies for districts. Empirical results suggest a minimum TOTAL that is between 450 and 499 people in a block group provides reliable characteristics of various demographic groups in a block group based on the *TDA*. Similarly, a minimum TOTAL that is between 200 and 249 is observed to provide reliable characteristics for places and MCDs. No Congressional or state legislative district failed our test for reliability. It is important to keep in mind that these results are comparisons to the swapped 2010 Census data. They do not evaluate the reliability relative to the actual enumeration in 2010 because the 2010 redistricting data contained statistical uncertainty due to swapping.

Part II is an update of our results reported in [6] where $\epsilon = 10.3$ with the difference being that this study uses $\epsilon = 17.14$. The objective here is to assess the variability of data results from application of the 2020 Census redistricting data production settings version *TDA* to the 2010 Census Edited File (2010 CEF) for Rhode Island and for three additional jurisdictions. Our approach has two parts: (1) to report observations on variability of results among 25 runs of the *TDA* and (2) to report observations on variability between the results among the 25 runs of the *TDA* and the published 2010 Census *Public Law 94-171* data. We observe that variability in data results from the *TDA* increases as we consider smaller pieces of geography and population. Variability with the 2020 Census redistricting data production settings version of the *TDA* ($\epsilon = 17.14$) tends to be less than what we reported in [6] with the 2021-04-28 version where $\epsilon = 10.3$.

Disclaimer and Acknowledgements: The views presented in this paper are those of the authors and not the U.S. Bureau of the Census. We are grateful to our colleague Mary Mulry for her reading of a draft of this study. We are also grateful for the many conversations with members of the Disclosure Avoidance System (DAS) Team. The statistics in this paper have been cleared by the Census Bureau Disclosure Review Board (DRB Clearance Number CBDRB-FY20-DSEP-001). *Corresponding author for comments:* tommy.wright@census.gov; (301) 763-1702.

.

COMMENT: Throughout Parts I and II, we compare *TDA* counts with published corresponding *SWA* counts from 2010 rather than with the “as enumerated” 2010 counts, i.e., counts in the 2010 Census Edited File (CEF). For a clean comparison, it would be better to compare *TDA* counts with the corresponding CEF counts. However, we share a few thoughts that provide some support for the path we take, to use the *SWA* counts as a reference for assessing the *TDA* counts. First, the *SWA* counts from 2010 are official; they have been used widely by the public for ten years; and we assume that they have generally been accepted as credible. The public is familiar with the *SWA* counts. In this spirit, we see some value in comparing *TDA* counts with *SWA* counts. This permits the public the opportunity to compare relatively easily and to possibly reproduce most of our results. This would be impossible if we had used the CEF counts, which are confidential. A primary objective in Part I is to convey a new data-based concept - “what we mean by declaring *TDA* counts reliable”. We don’t really need the CEF counts to discuss this concept. It should be noted that the *SWA* TOTAL counts and the corresponding CEF TOTAL counts at the block level were the same in 2010. The same is true for TOTAL18 counts for the 18 years and over population at the block level. It should also be noted that the “tuning” of the *TDA* makes use of the CEF counts rather than the *SWA* counts, and we understand that results are similar to what we share, especially with regard to the main question on reliability in Part I. Furthermore, had we used CEF counts, additional Disclosure Review Board clearance would have slowed the speed in sharing our study results.

TECHNICAL SUMMARY

The Census Bureau Data Stewardship Executive Policy Committee (DSEP) approved production settings for the 2020 Census Redistricting Data (Public Law 94-171) Summary File (hereafter "2020 Census redistricting data production settings") version of the *TopDown Algorithm (TDA)* [1] that will be applied to the 2020 Census Edited File (CEF), and the results will be used by jurisdictions in devising redistricting plans for selecting officials ranging from Members of the U.S. House of Representatives to local school boards. We also assume the results will be used for the analysis of such plans for compliance with Federal voting rights laws, including Section 2 of the *Voting Rights Act of 1965, 52 U.S.C. 10301.*

In Part I of this limited study, we attempt to take a closer look at reliability of characteristics of demographic groups inside smaller districts. For convenience, we consider "Census Block Groups, Minor Civil Divisions (MCDs), and Census Places" as proxies for smaller districts and seek to gain more insights regarding the following question:

> *"What is the minimum* TOTAL *(ideal*[a]*) population of a district to have reliable characteristics of various demographic groups?"*

For each of the 217,740 block groups and 21,591 MCDs and places in the United States, we desire to compare the closeness between the following two sets of population counts: (a) published *SWA* counts for twenty demographic groups based on the application of a Swapping Algorithm (*SWA*) to the 2010 CEF and (b) the corresponding *TDA* counts for the same twenty demographic groups based on application of the 2020 Census redistricting data production settings version of the *TDA* ($\epsilon = 17.14$) to the 2010 CEF. Our comparisons are facilitated by a measure called the **difference of ratios** *DR* (see Section I.1). We analyze data for block groups, MCDs, and places as proxies for districts to make reliability statements about *TDA* output. We also analyze all Congressional and state legislative districts. For block groups, MCDs, places, and legislative districts:

The Key Empirical Message on Reliability

> *"for any block group with a TOTAL count between 450 and 499 people, and for MCDs and places between 200 and 249, the difference between the TDA ratio of the largest demographic group (LDG) and the corresponding SWA ratio for the LDG is less than or equal to 5 percentage points at least 95% of the time". No Congressional or state legislative district fails this test; that is, for these districts, the 5 percentage point criterion holds 100% of the time.*

Part II of this study provides empirical results for ongoing research and development and provides an update of the data and results presented in [6] where $\epsilon = 10.3$; throughout this updated study, $\epsilon = 17.14$. The objective of this part of our study is to assess the variability of data results from application of the 2020 Census redistricting data production settings version of the *TDA* to the 2010 Census Edited File (2010 CEF) for Rhode Island and for three additional jurisdictions. *Because there has been more development of the TDA, a larger ϵ, and additional focus on how to allocate this ϵ, we tend to see less variability throughout.*

Part II of our study has two components: (1) report variability among 25 runs and (2) report variability of the 25 runs relative to the official published results from the 2010 Census (i.e., the 2010 Census Redistricting Data (Public Law 94-171) Summary File).

The first component of our study is a follow-up to earlier analyses [5, 6] for Rhode Island. For each of the given redistricting plans we studied for Rhode Island, we observe that counts and percentages put in place from swapping being applied to the 2010 CEF have very similar counts and percentages after the *TDA* is applied to the same 2010 CEF.

In the second component of our study, we repeat our analyses for three specific jurisdictions provided by the U.S. Department of Justice (DOJ). Our observations for these three smaller geographies and populations show similarities between swapping (SWA) and TDA results.

The key data analyses are presented

(i) in Tables 7, 8, 9, 10, 11, and 12 that contain SWA counts and percentages publicly released following the 2010 Census and corresponding released TDA counts and percentages; and

(ii) in Tables 7V, 8V, 9V, 10V, 11V, and 12V that contain measures of relative variability for the TDA as described in Section II.8 (APPENDIX B contains an illustration of the computations).

The Key Empirical Message on Variability

The two measures $AVERV(\cdot)$ and $MEDRV(\cdot)$, defined in Section II.7, summarize the key single empirical message for Part II of this study ($\epsilon = 17.14$). As we reported in [5, 6], relative variability in the TDA increases as we consider smaller pieces of geography and population. To see this empirical evidence, sequentially observe the values for $AVERV(\cdot)$ and $MEDRV(\cdot)$ on the last two rows of Tables 7V; 8V; 9V; 10V; 11V; and 12V; also see Figure 1. At a high level, *Figure 2 tends to show less relative variability using the 2020 Census redistricting data production settings version of the TDA than the 2021-04-28 version.*

Part I

THE MINIMUM TOTAL POPULATION OF A GEOGRAPHIC DISTRICT
TO HAVE RELIABLE CHARACTERISTICS OF VARIOUS DEMOGRAPHIC GROUPS

I.1. INTRODUCTION

Our earlier empirical study [6] assessed the variability of data results from application of the 2021-04-28 version of the *TopDown Algorithm (TDA)* to the 2010 Census Edited File (2010 CEF) for disclosure avoidance and confidentiality protection. It documented that it is the smaller geographic districts with smaller ideal[a] populations where we observed more variability among twenty-five different runs of the *TDA*. Indeed, it is the block level where redistricting takes place, where local people have some sense of "ground truth", and where some field checking seems possible to assess the reliability of *TDA* output. In Part I of this study, we attempt to take a closer look at variability for smaller districts (a level closer to the block level) and the reliability of counts of various demographic groups in these smaller districts based on the *TDA*. As a proxy for smaller districts, we consider Census block groups, Minor Civil Divisions (MCDs) and Census places and seek insights for the following question:

> *"What is the minimum* TOTAL *(ideal[a]) population of a district to have reliable characteristics of various demographic groups?"*

(A block group is a cluster of blocks and generally contains between 600 and 3,000 people. MCDs and places vary in size, but approximately half have population less than or equal to 2,100 people.)

For each of the 217,740 block groups in the United States and for each of the 21,591 MCDs and places, we desire to compare closeness between the following two sets of population counts: (a) published *SWA* counts for twenty demographic groups based on the application of a Swapping Algorithm (*SWA*) to the 2010 CEF and (b) the corresponding *TDA* counts for the same twenty demographic groups based on application of the DSEP-approved production settings for the 2020 Census Redistricting Data (Public Law 94-171) Summary File (hereafter "2020 Census redistricting data production settings") version of the *TDA* to the 2010 CEF. Our comparisons are facilitated by the **difference of ratios** (*DR*).

Definition 1: Let $C_{SWA}(g)$ and $C_{TDA}(g)$ be two competing counts of the demographic group g associated with a block group (more generally, geographic district) whose total population counts are C_{SWA} and C_{TDA}, respectively. The **difference of ratios** is the absolute value of the difference between the *SWA* ratio $\dfrac{C_{SWA}(g)}{C_{SWA}}$ and the *TDA* ratio $\dfrac{C_{TDA}(g)}{C_{TDA}}$, given by:

$$DR_g = \left| \frac{C_{SWA}(g)}{C_{SWA}} - \frac{C_{TDA}(g)}{C_{TDA}} \right|. \tag{1}$$

Small values of the difference of ratios DR_g imply that the ratios for a group g due to *SWA* and *TDA* in the block group, MCD, or place are close.

Definition 2: When DR_g is sufficiently small while comparing a $C_{SWA}(g)$ count and corresponding $C_{TDA}(g)$ count for a demographic group g associated with a given block group, MCD or place, we say that the $C_{TDA}(g)$ count (or ratio) provides a **reliable characteristic** for the block group, MCD, or place.

[a] The ideal population for each of K districts of a jurisdiction is the jurisdiction's total population divided by K.

I.2. ILLUSTRATION OF COMPUTATIONS FOR TWO BLOCK GROUPS

For a block group in Maryland, Table 1a provides differences of ratios for twenty demographic groups as used in the past for redistricting related analyses [6]. For definition of each demographic group, see APPENDIX A. For the demographic group g = ASIANNH18, $C_{SWA}(g) = 142$ and $C_{TDA}(g) = 146$ with difference of ratios $DR_g = 0.0003$. That is, the difference between the two ratios for demographic group g is 0.03 percentage points for this block group. (Note using Appendix A that $C_{SWA}(g) = 142 (= 130 + 12)$ where 130 is the count for all individuals 18 years of age or older who chose Asian singly and chose Not Hispanic; and 12 is the count for all individuals 18 years of age or older who chose Asian in combination with White and chose Not Hispanic.)

Note: When the counts being compared are for individuals of all ages for a block group, we take C_{SWA} = TOTAL count using SWA and C_{TDA} = TOTAL count using TDA; when the counts being compared for individuals 18 years and older for a block group, we take C_{SWA} = TOTAL18 count using SWA and C_{TDA} = TOTAL18 count using TDA.

Table 1a: Block Group 240317044041 (564 HUs) Characteristics

($C_{TDA}(g)$ counts result from 2020 Census Redistricting Data Production Settings ($\epsilon = 17.14$ for persons) version of TDA.)

Demographic Group $(g)^b$	$C_{SWA}(g)$	$C_{TDA}(g)$	$DR_g = \left\lvert \frac{C_{SWA}(g)}{C_{SWA}} - \frac{C_{TDA}(g)}{C_{TDA}} \right\rvert$
TOTAL	1,560	1,598	^{c}
TOTAL18	1,198	1,229	^{c}
TOTALHISP	133	141	$\lvert \frac{133}{1,560} - \frac{141}{1,598} \rvert = 0.0030$
TOTALNH	1,427	1,457	$\lvert \frac{1,427}{1,560} - \frac{1,457}{1,598} \rvert = 0.0030$
WHITENH	1,169	1,178	$\lvert \frac{1,169}{1,560} - \frac{1,178}{1,598} \rvert = 0.0122$
BLACKNH	36	54	$\lvert \frac{36}{1,560} - \frac{54}{1,598} \rvert = 0.0107$
AIANNH	10	8	$\lvert \frac{10}{1,560} - \frac{8}{1,598} \rvert = 0.0014$
ASIANNH	187	189	$\lvert \frac{187}{1,560} - \frac{189}{1,598} \rvert = 0.0016$
HPINH	5	2	$\lvert \frac{5}{1,560} - \frac{2}{1,598} \rvert = 0.0020$
OTHERNH	11	12	$\lvert \frac{11}{1,560} - \frac{12}{1,598} \rvert = 0.0005$
MLTMNNH	9	14	$\lvert \frac{9}{1,560} - \frac{14}{1,598} \rvert = 0.0030$
HISP18	93	95	$\lvert \frac{93}{1,198} - \frac{95}{1,229} \rvert = 0.0003$
NONHISP18	1,105	1,134	$\lvert \frac{1,105}{1,198} - \frac{1,134}{1,229} \rvert = 0.0003$
WHITENH18	914	923	$\lvert \frac{914}{1,198} - \frac{923}{1,229} \rvert = 0.0119$
BLACKNH18	29	42	$\lvert \frac{29}{1,198} - \frac{42}{1,229} \rvert = 0.0100$
AIANNH18	8	8	$\lvert \frac{8}{1,198} - \frac{8}{1,229} \rvert = 0.0002$
ASIANNH18	142	146	$\lvert \frac{142}{1,198} - \frac{146}{1,229} \rvert = 0.0003$
HPINH18	2	2	$\lvert \frac{2}{1,198} - \frac{2}{1,229} \rvert = 0.0000$
OTHERNH18	6	4	$\lvert \frac{6}{1,198} - \frac{4}{1,229} \rvert = 0.0018$
MLTMNNH18	4	9	$\lvert \frac{4}{1,198} - \frac{9}{1,229} \rvert = 0.0040$

[b] For definitions of the demographic groups, see APPENDIX A.

[c] Because $DR_g = 0.0000$ when g = TOTAL or g = TOTAL18 in Tables 1a, 1b, and 2, we leave the entries for DR_g empty. To see comparisons in these cases, one could take $|C_{SWA}(g) - C_{TDA}(g)|/C_{SWA}$ which is a special case of DR_g. (A similar approach could be taken for TOTAL18.)

Thus from Table 1a and for the difference of ratios for demographic group g = TOTALNH, $DR_g = 0.0030$; the difference between the two ratios is $0.0030 \times 100\% = 0.30$ percentage points.

Table 1b provides similar characteristics of demographic groups for a block group in Washington D.C. From Table 1b and for the difference of ratios for demographic group g = TOTALNH, the difference between the ratios is $0.0017 \times 100\% = 0.17$ percentage points.

Table 1b: Block Group 110010047012 (1,709 HUs) Characteristics

($C_{TDA}(g)$ counts result from 2020 Census Redistricting Data Production Settings ($\epsilon = 17.14$ for persons) version of *TDA*.)

Demographic Group (g)	$C_{SWA}(g)$	$C_{TDA}(g)$	$DR_g = \left\lvert \frac{C_{SWA}(g)}{C_{SWA}} - \frac{C_{TDA}(g)}{C_{TDA}} \right\rvert$
TOTAL	2,875	2,868	c
TOTAL18	2,261	2,244	c
TOTALHISP	92	87	0.0017
TOTALNH	2,783	2,781	0.0017
WHITENH	541	534	0.0020
BLACKNH	1,686	1,688	0.0021
AIANNH	12	10	0.0007
ASIANNH	515	524	0.0036
HPINH	1	0	0.0003
OTHERNH	3	3	0.0000
MLTMNNH	25	22	0.0010
HISP18	86	74	0.0051
NONHISP18	2,175	2,170	0.0051
WHITENH18	529	526	0.0004
BLACKNH18	1,151	1,150	0.0034
AIANNH18	12	10	0.0009
ASIANNH18	460	461	0.0020
HPINH18	1	0	0.0004
OTHERNH18	3	2	0.0004
MLTMNNH18	19	21	0.0010

I.3. CHARACTERISTICS OF TWELVE MORE BLOCK GROUPS

We extend our overview of block groups beyond those in Tables 1a and 1b by considering counts for the demographic groups for block groups with TOTAL that span from 82 (this block group is actually the complete Loving County, Texas) to 37,452 (this block group is the largest block group in population in the United States). Table 2 presents the characteristics we observe. Our analyses focus more on the larger demographic groups within each block group because they may play a larger role when thinking about reliable characteristics of actual districts. We highlight the counts and DR_g's for the following demographic groups {TOTAL, TOTAL18} and for some of the demographic groups {TOTALHISP, WHITENH, BLACKNH, AIANNH, ASIANNH, HPINH}. The superscripts ^{1}, ^{2}, and ^{3} represent, in order, the three largest demographic groups among TOTALHISP, WHITENH, BLACKNH, AIANNH, ASIANNH, and HPINH (based on $C_{TDA}(g)$ counts) for the block group. Clearly, as the count for the TOTAL demographic group increases across the twelve block groups in Table 2, corresponding values of highlighted DR_g values tend to decrease.

Motivating Example for Reliable Characteristics

Assume we stratify or partition the 12 block groups in Table 2 into 4 strata; the first three, then the next 3, the next three, and finally the last three with the following DR_g values for each stratum where g is the largest demographic group: {0.0494, 0.0239, 0.0032}; {0.0127, 0.0024, 0.0056}; {0.0012, 0.0001, 0.0010}; and {0.0004, 0.0000, 0.0000}. Assume the *TDA* count is considered a reliable characteristic for the largest demographic group if its $DR_g \leq 0.0050$. One of the block groups in stratum 1 would be reliable; 1 out of 3 (0.3333) of the block groups in stratum 2 would be reliable; all 3 (1.0000) of the block groups in stratum 3 would be reliable; and finally, again all 3 (1.0000) of the block groups in stratum 4 would be reliable. We build on this in Section I.4.

Table 2: Characteristics of Twelve Block Groups

($C_{TDA}(g)$ counts result from 2020 Census Redistricting Data Production Settings ($\epsilon = 17.14$ for persons) version of TDA.)

Demographic Group (g)	Block Group 483019501001 (TX)[d]			Block Group 010599729001 (AL)			Block Group 010059507002 (AL)			Block Group 040030008001 (AZ)		
	C_{SWA}	C_{TDA}	DR_g	C_{SWA}	C_{TDA}	DR_g	C_{SWA}	C_{TDA}	DR_g	C_{SWA}	C_{TDA}	DR_g
TOTAL	82	**85**	[c]	500	**514**	[c]	1,000	**980**	[c]	1,500	**1,543**	[c]
TOTAL18	73	**67**	[c]	386	**389**	[c]	745	**733**	[c]	1,035	**1,047**	[c]
TOTALHISP	18	**13**[2]	**0.0666**	18	**16**[2]	**0.0049**	30	**24**[3]	**0.0055**	1,237	**1,292**[1]	**0.0127**
TOTALNH	64	72	0.0666	482	498	0.0049	970	956	0.0055	263	251	0.0127
WHITENH	60	**58**[1]	**0.0494**	455	**480**[1]	**0.0239**	306	**301**[2]	**0.0011**	235	**215**[2]	**0.0173**
BLACKNH	0	0	0.0000	7	**9**[3]	**0.0035**	659	**649**[1]	**0.0032**	10	9	0.0008
AIANNH	4	2	0.0253	6	7	0.0016	4	1	0.0030	0	2	0.0013
ASIANNH	0	**2**[3]	**0.0235**	11	0	0.0220	0	2	0.0020	18	**20**[3]	**0.0010**
HPINH	0	2	0.0235	0	0	0.0000	0	0	0.0000	0	1	0.0006
OTHERNH	0	7	0.0824	1	1	0.0001	0	0	0.0000	0	2	0.0013
MLTMNNH	0	1	0.0118	2	1	0.0021	1	3	0.0021	0	2	0.0013
HISP18	14	3	0.1470	10	11	0.0024	21	22	0.0018	807	818	0.0016
NONHISP18	59	64	0.1470	376	378	0.0024	724	711	0.0018	228	229	0.0016
WHITENH18	55	53	0.0376	354	366	0.0238	255	250	0.0012	203	198	0.0070
BLACKNH18	0	0	0.0000	6	5	0.0027	464	458	0.0020	9	9	0.0001
AIANNH18	4	2	0.0249	5	7	0.0050	4	1	0.0040	0	2	0.0019
ASIANNH18	0	0	0.0000	9	0	0.0233	0	0	0.0000	16	17	0.0008
HPINH18	0	2	0.0299	0	0	0.0000	0	0	0.0000	0	1	0.0010
OTHERNH18	0	6	0.0896	0	0	0.0000	0	0	0.0000	0	1	0.0010
MLTMNNH18	0	1	0.0149	2	0	0.0052	1	2	0.0014	0	1	0.0010

[d] This block group is all of Loving County, Texas.

Table 2: Characteristics of Twelve Block Groups (continued)

($C_{TDA}(g)$ counts result from 2020 Census Redistricting Data Production Settings ($\epsilon = 17.14$ for persons) version of TDA.)

Demographic Group (g)	Block Group 040030017032 (AZ)			Block Group 051430110011 (AR)			Block Group 120210112023 (FL)			Block Group 131350505461 (GA)		
	C_{SWA}	C_{TDA}	DR_g	C_{SWA}	C_{TDA}	DR_g	C_{SWA}	C_{TDA}	DR_g	C_{SWA}	C_{TDA}	DR_g
TOTAL	2,000	**1,982**	[c]	3,000	**2,976**	[c]	5,001	**4,980**	[c]	10,000	**10,042**	[c]
TOTAL18	1,562	**1,549**	[c]	2,153	**2,125**	[c]	3,689	**3,689**	[c]	6,704	**6,716**	[c]
TOTALHISP	349	**323**[2]	**0.0115**	224	**221**[2]	**0.0004**	1,770	**1,773**[2]	**0.0021**	1,291	**1,298**[3]	**0.0002**
TOTALNH	1,651	1,659	0.0115	2,776	2,755	0.0004	3,231	3,207	0.0021	8,709	8,744	0.0002
WHITENH	1,308	**1,301**[1]	**0.0024**	2,580	**2,576**[1]	**0.0056**	2,891	**2,873**[1]	**0.0012**	3,565	**3,569**[2]	**0.0011**
BLACKNH	181	**171**[3]	**0.0042**	87	**80**[3]	**0.0021**	235	**223**[3]	**0.0022**	4,475	**4,495**[1]	**0.0001**
AIANNH	25	24	0.0004	65	64	0.0002	18	13	0.0010	30	32	0.0002
ASIANNH	106	120	0.0075	32	31	0.0003	59	65	0.0013	473	472	0.0003
HPINH	10	10	0.0000	1	0	0.0003	8	3	0.0010	2	2	0.0000
OTHERNH	3	9	0.0030	4	2	0.0007	7	17	0.0020	79	90	0.0011
MLTMNNH	18	24	0.0031	7	2	0.0017	13	13	0.0000	85	84	0.0001
HISP18	236	220	0.0091	110	111	0.0011	1,193	1,201	0.0022	783	787	0.0004
NONHISP18	1,326	1,329	0.0091	2,043	2,014	0.0011	2,496	2,488	0.0022	5,921	5,929	0.0004
WHITENH18	1,089	1,080	0.0000	1,931	1,913	0.0033	2,267	2,259	0.0022	2,630	2,628	0.0010
BLACKNH18	129	122	0.0038	40	37	0.0012	149	144	0.0014	2,868	2,876	0.0004
AIANNH18	20	18	0.0012	41	41	0.0003	14	6	0.0022	22	26	0.0006
ASIANNH18	72	86	0.0094	23	19	0.0017	50	58	0.0022	304	303	0.0002
HPINH18	4	4	0.0000	1	0	0.0005	4	3	0.0003	2	2	0.0000
OTHERNH18	2	6	0.0026	3	2	0.0005	5	10	0.0014	43	43	0.0000
MLTMNNH18	10	13	0.0020	4	2	0.0009	7	8	0.0003	52	51	0.0002

Table 2: Characteristics of Twelve Block Groups (continued)

($C_{TDA}(g)$ counts result from 2020 Census Redistricting Data Production Settings ($\epsilon = 17.14$ for persons) version of *TDA*.)

Demographic Group (g)	Block Group 130510107001 (GA)			Block Group 517100038001 (VA)			Block Group 121199112001 (FL)			Block Group 060730187001 (CA)		
	C_{SWA}	C_{TDA}	DR_g	C_{SWA}	C_{TDA}	DR_g	C_{SWA}	C_{TDA}	DR_g	C_{SWA}	C_{TDA}	DR_g
TOTAL	15,089	**15,101**	c	19,506	**19,512**	c	29,677	**29,672**	c	37,452	**37,453**	c
TOTAL18	11,561	**11,567**	c	19,486	**19,460**	c	29,214	**29,220**	c	28,368	**28,408**	c
TOTALHISP	1,066	**1,057**[3]	**0.0007**	2,599	**2,592**[3]	**0.0004**	502	**501**[2]	**0.0000**	8,192	**8,190**[2]	**0.0001**
TOTALNH	14,023	14,044	0.0007	16,907	16,920	0.0004	29,175	29,171	0.0000	29,260	29,263	0.0001
WHITENH	7,901	**7,923**[1]	**0.0010**	10,579	**10,590**[1]	**0.0004**	28,555	**28,550**[1]	**0.0000**	23,326	**23,328**[1]	**0.0000**
BLACKNH	5,281	**5,284**[2]	**0.0001**	4,972	**4,976**[2]	**0.0001**	276	**274**[3]	**0.0001**	3,040	**3,047**[3]	**0.0002**
AIANNH	54	51	0.0002	275	284	0.0005	58	62	0.0001	601	601	0.0000
ASIANNH	643	649	0.0004	776	782	0.0003	246	244	0.0001	1,422	1,427	0.0001
HPINH	17	12	0.0003	80	77	0.0002	7	7	0.0000	340	343	0.0001
OTHERNH	42	40	0.0001	45	43	0.0001	15	17	0.0001	89	91	0.0001
MLTMNNH	85	85	0.0000	180	168	0.0006	18	17	0.0000	442	426	0.0004
HISP18	693	693	0.0000	2,597	2,583	0.0005	460	460	0.0000	5,506	5,539	0.0009
NONHISP18	10,868	10,874	0.0000	16,889	16,877	0.0005	28,754	28,760	0.0000	22,862	22,869	0.0009
WHITENH18	6,404	6,409	0.0001	10,562	10,562	0.0007	28,186	28,189	0.0001	18,751	18,745	0.0011
BLACKNH18	3,849	3,860	0.0008	4,971	4,972	0.0004	247	249	0.0001	2,118	2,124	0.0001
AIANNH18	46	40	0.0005	275	284	0.0005	58	56	0.0001	436	439	0.0001
ASIANNH18	494	494	0.0000	776	780	0.0003	227	229	0.0001	1,032	1,036	0.0001
HPINH18	9	10	0.0001	80	77	0.0001	7	7	0.0000	261	261	0.0000
OTHERNH18	22	21	0.0001	45	42	0.0002	14	15	0.0000	62	60	0.0001
MLTMNNH18	44	40	0.0003	180	160	0.0010	15	15	0.0000	202	204	0.0001

I.4. THE QUESTION

More focused and concretely, we might proceed as follows to get an answer to our question at the national level (might also look at each state). To be more specific, imagine ordering the 217,740 block groups from smallest to largest C_{SWA} counts for the demographic group TOTAL (Later, we focus only on block groups where $50 \leq C_{SWA} \leq 2,499$). To each block group in this ordering, imagine attaching its Table (as given for example in Tables 1a, 1b, or 2) of counts and difference of ratios values for all of the twenty demographic groups. To respond to our question, we seek to determine a value C^*_{SWA} for the TOTAL block group such that for block groups whose TOTAL C_{SWA} value is less than C^*_{SWA}, the differences of ratios of the twenty demographic groups tend to be large, i.e., the counts (or characteristics) are not reliable; also for block groups whose TOTAL C_{SWA} values are greater than C^*_{SWA}, the differences of ratios of the twenty demographic groups tend to be small. See (2) below. (We use a similar ordering for MCDs and places, as well as for Congressional and state legislative districts.)

$$C_{SWA(1)} \leq C_{SWA(2)} \leq C_{SWA(3)} \leq \cdots \leq C^*_{SWA} \leq \cdots \leq C_{SWA(217{,}739)} \leq C_{SWA(217{,}740)}, \tag{2}$$

where the $C_{SWA(i)}$ counts are the counts for the TOTAL block group i, for $i = 1; 2; ...; 217,740$.

Table 3 reveals an empirical answer to our question. For each block group, we consider three criteria (others could be considered) for the expression "reliable characteristics" based on the largest demographic group's (LDG) $DR_g \leq 0.01$; the largest demographic group's (LDG) $DR_g \leq 0.03$; and the largest demographic group's (LDG) $DR_g \leq 0.05$. For each criterion (column), Table 3 gives proportions of the number of block groups that satisfy the criterion for different strata of block groups based on TOTAL C_{SWA} counts. For example, consider the 7,356 block groups in the stratum where "$700 \leq C_{SWA} \leq 749$" for the TOTAL demographic group. We consider three (3) different criteria and present the proportion of block groups that satisfy Criterion I, or Criterion II, or Criterion III. For Criterion I (LDG $DR_g \leq 0.01$), 0.5007 (or 50.07%) of the 7,356 block groups have $DR_g \leq 0.01$ for LDG counts. Because the proportions tend to increase as one goes down the Criterion I column, it seems that for each stratum below the stratum $700 \leq C_{SWA} \leq 749$ (i.e., those strata with larger block group TOTAL counts), one also tends to see that at least 0.5007 of the block groups have $DR_g \leq 0.01$ for LDG counts. We observe a similar trend for the other two Criterion columns. For Criterion III (LDG $DR_g \leq 0.05$), 0.9826 (or 98.26%) of the 7,356 block groups have $DR_g \leq 0.05$ for the block group's largest demographic group among TOTALHISP, WHITENH, BLACKNH, AIANNH, ASIANNH, and HPINH groups. We do not consider any block groups where the C_{SWA} count for TOTAL block group is less than 50 or greater than 2,499. (*Table 3′ of APPENDIX C gives analogous results as Table 3 for the 18 years and over population.*)

Table 3: Proportion of Block Groups in Each Stratum for Three Criteria

(Proportion computations result from 2020 Census Redistricting Data Production Settings ($\epsilon = 17.14$ for persons) version of *TDA*.)

Population: United States (50 States & DC)

		Reliable Characteristics Criteria		
Stratum for Block Groups Using C_{SWA} for TOTAL	Number of Block Groups	Criterion I LDG $DR_g \leq 0.01$	Criterion II LDG $DR_g \leq 0.03$	Criterion III LDG $DR_g \leq 0.05$
$50 \leq C_{SWA} \leq 99$	128	0.1250	0.3594	0.5156
$100 \leq C_{SWA} \leq 149$	99	0.1818	0.5253	0.7071
$150 \leq C_{SWA} \leq 199$	124	0.1694	0.5565	0.7581
$200 \leq C_{SWA} \leq 249$	154	0.2662	0.6234	0.7922
$250 \leq C_{SWA} \leq 299$	209	0.2919	0.6459	0.8565
$300 \leq C_{SWA} \leq 349$	264	0.3636	0.7348	0.8902
$350 \leq C_{SWA} \leq 399$	407	0.3366	0.7346	0.8698
$400 \leq C_{SWA} \leq 449$	569	0.4077	0.7750	0.9315
$450 \leq C_{SWA} \leq 499$	915	0.4087	0.8284	0.9552
$500 \leq C_{SWA} \leq 549$	1,699	0.4197	0.8458	0.9588
$550 \leq C_{SWA} \leq 599$	3,238	0.4546	0.8684	0.9654
$600 \leq C_{SWA} \leq 649$	5,131	0.4578	0.8827	0.9751
$650 \leq C_{SWA} \leq 699$	6,683	0.4718	0.8927	0.9753
$700 \leq C_{SWA} \leq 749$	7,356	0.5007	0.9082	0.9826
$750 \leq C_{SWA} \leq 799$	8,170	0.5160	0.9100	0.9845
$800 \leq C_{SWA} \leq 849$	8,213	0.5268	0.9293	0.9897
$850 \leq C_{SWA} \leq 899$	8,441	0.5517	0.9371	0.9914
$900 \leq C_{SWA} \leq 949$	8,657	0.5557	0.9409	0.9920
$950 \leq C_{SWA} \leq 999$	8,723	0.5849	0.9512	0.9952
$1{,}000 \leq C_{SWA} \leq 1{,}049$	8,398	0.6044	0.9582	0.9952
$1{,}050 \leq C_{SWA} \leq 1{,}099$	8,345	0.6192	0.9646	0.9965
$1{,}100 \leq C_{SWA} \leq 1{,}149$	7,950	0.6244	0.9701	0.9972
$1{,}150 \leq C_{SWA} \leq 1{,}199$	7,860	0.6422	0.9763	0.9977
$1{,}200 \leq C_{SWA} \leq 1{,}249$	7,451	0.6515	0.9757	0.9988
$1{,}250 \leq C_{SWA} \leq 1{,}299$	7,124	0.6645	0.9749	0.9978
$1{,}300 \leq C_{SWA} \leq 1{,}349$	6,714	0.6822	0.9812	0.9988
$1{,}350 \leq C_{SWA} \leq 1{,}399$	6,507	0.6859	0.9866	0.9989
$1{,}400 \leq C_{SWA} \leq 1{,}449$	5,911	0.7090	0.9866	0.9992
$1{,}450 \leq C_{SWA} \leq 1{,}499$	5,617	0.7002	0.9858	0.9995
$1{,}500 \leq C_{SWA} \leq 1{,}549$	5,390	0.7330	0.9900	0.9994
$1{,}550 \leq C_{SWA} \leq 1{,}599$	4,856	0.7341	0.9866	0.9994
$1{,}600 \leq C_{SWA} \leq 1{,}649$	4,508	0.7420	0.9918	0.9998
$1{,}650 \leq C_{SWA} \leq 1{,}699$	4,325	0.7489	0.9908	0.9998
$1{,}700 \leq C_{SWA} \leq 1{,}749$	4,093	0.7669	0.9922	1.0000
$1{,}750 \leq C_{SWA} \leq 1{,}799$	3,689	0.7650	0.9938	0.9997
$1{,}800 \leq C_{SWA} \leq 1{,}849$	3,469	0.7530	0.9925	1.0000
$1{,}850 \leq C_{SWA} \leq 1{,}899$	3,252	0.7811	0.9945	0.9997
$1{,}900 \leq C_{SWA} \leq 1{,}949$	3,008	0.7793	0.9947	1.0000
$1{,}950 \leq C_{SWA} \leq 1{,}999$	2,832	0.7970	0.9965	1.0000
$2{,}000 \leq C_{SWA} \leq 2{,}049$	2,573	0.8022	0.9965	1.0000
$2{,}050 \leq C_{SWA} \leq 2{,}099$	2,356	0.7975	0.9966	1.0000
$2{,}100 \leq C_{SWA} \leq 2{,}149$	2,307	0.8331	0.9957	1.0000
$2{,}150 \leq C_{SWA} \leq 2{,}199$	2,033	0.8170	0.9975	1.0000
$2{,}200 \leq C_{SWA} \leq 2{,}249$	1,999	0.8354	0.9990	1.0000
$2{,}250 \leq C_{SWA} \leq 2{,}299$	1,892	0.8494	0.9984	1.0000
$2{,}300 \leq C_{SWA} \leq 2{,}349$	1,666	0.8331	0.9982	1.0000
$2{,}350 \leq C_{SWA} \leq 2{,}399$	1,622	0.8453	0.9994	1.0000
$2{,}400 \leq C_{SWA} \leq 2{,}449$	1,421	0.8621	0.9993	1.0000
$2{,}450 \leq C_{SWA} \leq 2{,}499$	1,350	0.8600	1.0000	1.0000
Total	199,698			

Using Criterion II and searching from top to bottom for the first stratum whose proportion is at least 0.9500: From Table 3, take C^*_{SWA} to be between 950 and 999. For block groups whose TOTAL C_{SWA} count is at least 999, the difference of ratios between the C_{TDA} and C_{SWA} ratios for the LDG will tend to be less than or equal to 3% (using our data).

Using Criterion III and searching from top to bottom for the first stratum whose proportion is at least 0.9500: From Table 3, take C^*_{SWA} to be between 450 and 499. For block groups whose TOTAL C_{SWA} count is at least 499, the difference of ratios between the C_{TDA} and C_{SWA} ratios for the LDG will tend to be less than or equal to 5% (using our data).

Using data to be released to the public (one run of the 2020 Census redistricting data production settings version of *TDA*), we might say, empirically based on the data for the block groups used in our study, that

> *"for any block group with a TOTAL count between 450 and 499 people, the difference between the TDA ratio of the largest demographic group (LDG) and the corresponding SWA ratio for the LDG is less than or equal to 5 percentage points at least 95% of the time".*

The same version of the *TDA* was applied to the same underlying CEF data 25 independent times, i.e., for 25 additional runs. For each run, the stratum where we first observed that 0.9500 was exceeded is given in Table 4 for each run. (*Table 3a′ of APPENDIX C gives analogous results as Table 3a for the 18 years and over population.*)

Table 3a: For Each Run, the Stratum and Stratum Proportion When 0.9500 First Exceeded
(Proportion computations result from 2020 Census Redistricting Data Production Settings ($\epsilon = 17.14$ for persons) version of *TDA*.)
Population: United States (50 States & DC)

		Criterion III LDG $DR_g \leq 0.05$
TDA Run	Stratum for Block Groups	Proportion When 0.9500 First Exceeded
1	$450 \leq C_{SWA} \leq 499$	0.9716
2	$400 \leq C_{SWA} \leq 449$	0.9596
3	$450 \leq C_{SWA} \leq 499$	0.9661
4	$400 \leq C_{SWA} \leq 449$	0.9543
5	$400 \leq C_{SWA} \leq 449$	0.9561
6	$400 \leq C_{SWA} \leq 449$	0.9508
7	$350 \leq C_{SWA} \leq 399$	0.9509
8	$450 \leq C_{SWA} \leq 499$	0.9541
9	$450 \leq C_{SWA} \leq 499$	0.9617
10	$450 \leq C_{SWA} \leq 499$	0.9661
11	$450 \leq C_{SWA} \leq 499$	0.9596
12	$450 \leq C_{SWA} \leq 499$	0.9683
13	$400 \leq C_{SWA} \leq 449$	0.9525
14	$400 \leq C_{SWA} \leq 449$	0.9543
15	$350 \leq C_{SWA} \leq 399$	0.9558
16	$450 \leq C_{SWA} \leq 499$	0.9650
17	$450 \leq C_{SWA} \leq 499$	0.9607
18	$400 \leq C_{SWA} \leq 449$	0.9596
19	$450 \leq C_{SWA} \leq 499$	0.9727
20	$350 \leq C_{SWA} \leq 399$	0.9582
21	$450 \leq C_{SWA} \leq 499$	0.9617
22	$450 \leq C_{SWA} \leq 499$	0.9683
23	$350 \leq C_{SWA} \leq 399$	0.9558
24	$450 \leq C_{SWA} \leq 499$	0.9628
25	$450 \leq C_{SWA} \leq 499$	0.9519

Each "block group" represents a type of defined geography used by the Census Bureau which is among a series of statistical and legal geographic entities that have a nesting relationship with each other including: nation, state, county, tract, block group, and block. Many Census Bureau data products provide access to information about such nested geographies.

There are other types of defined geographies that are not a part of this nesting. These geographies (e.g., places, school districts, minor civil divisions,..) do not provide a complete national coverage and we consider them in this study as proxies for the yet to be defined electoral geography such as congressional, state legislative, and other electoral districts. [A Census Bureau designated place (CDP) is a statistical entity (geography) that is typically an unincorporated community, a concentration of population, housing, and commercial structures, identifiable by name, but not within an incorparated place. A Census Bureau incorporated place is a legally bounded entity, typically includes cities, towns (except in some states), villages, boroughs (except in New York and Alaska). A minor civil division (MCD) is a legally defined county subdivision. MCDs are the primary divisions of a county. They comprise both governmentally functioning entities–that is, those with elected or appointed officials who provide services and raise revenues–and nonfunctioning entities that exist primarily for administrative purposes, such as election districts. *Source: Census Bureau*]

Analysis of Places and MCDs

As with Tables 3 and 3a for block groups, we present analagous results in Tables 4 and 4a using results from a single run and an additional 25 runs for all "places and MCDs". Altogether, we make use of 21,591 places and minor civil divisions (including 6,607,533 blocks). Concerning the distribution of these places and MCDs using TOTAL counts, we note: Min = 0; 25^{th} percentile = 547; 50^{th} percentile = 2,065; mean = 11,743; 75^{th} percentile = 7,695; Max = 3,796,060. Again using Criterion III for all places and minor civil divisions in the United States, the stratum where we first observed that 0.9500 was exceeded is given in Table 4a for each run. Also, see details of a single run in Table 4.

.

Table 4: Proportion of Places and MCDs in Each Stratum for Three Criteria

(Proportion computations result from 2020 Census Redistricting Data Production Settings ($\epsilon = 17.14$ for persons) version of *TDA*.)

Population: United States (50 States & DC)

		Reliable Characteristics Criteria		
Stratum for Places and MCDs Using C_{SWA} for TOTAL	Number of Places and MCDs	Criterion I LDG $DR_g \leq 0.01$	Criterion II LDG $DR_g \leq 0.03$	Criterion III LDG $DR_g \leq 0.05$
$50 \leq C_{SWA} \leq 99$	573	0.2182	0.5969	0.7923
$100 \leq C_{SWA} \leq 149$	622	0.4051	0.7540	0.9116
$150 \leq C_{SWA} \leq 199$	645	0.3442	0.8109	0.9473
$200 \leq C_{SWA} \leq 249$	598	0.4197	0.8361	0.9632
$250 \leq C_{SWA} \leq 299$	500	0.4860	0.9000	0.9760
$300 \leq C_{SWA} \leq 349$	448	0.5379	0.9152	0.9844
$350 \leq C_{SWA} \leq 399$	417	0.5731	0.9233	0.9808
$400 \leq C_{SWA} \leq 449$	399	0.6416	0.9449	0.9975
$450 \leq C_{SWA} \leq 499$	344	0.6424	0.9680	0.9913
$500 \leq C_{SWA} \leq 549$	341	0.6716	0.9765	0.9971
$550 \leq C_{SWA} \leq 599$	291	0.7113	0.9691	0.9966
$600 \leq C_{SWA} \leq 649$	277	0.6859	0.9783	1.0000
$650 \leq C_{SWA} \leq 699$	306	0.7157	0.9902	1.0000
$700 \leq C_{SWA} \leq 749$	254	0.7165	0.9843	1.0000
$750 \leq C_{SWA} \leq 799$	233	0.7425	0.9914	1.0000
$800 \leq C_{SWA} \leq 849$	255	0.7569	0.9608	1.0000
$850 \leq C_{SWA} \leq 899$	222	0.7162	0.9955	1.0000
$900 \leq C_{SWA} \leq 949$	201	0.7562	0.9851	1.0000
$950 \leq C_{SWA} \leq 999$	210	0.7571	0.9952	1.0000
$1{,}000 \leq C_{SWA} \leq 1{,}049$	223	0.7982	0.9955	1.0000
$1{,}050 \leq C_{SWA} \leq 1{,}099$	157	0.8153	0.9873	1.0000
$1{,}100 \leq C_{SWA} \leq 1{,}149$	194	0.7423	0.9897	0.9948
$1{,}150 \leq C_{SWA} \leq 1{,}199$	178	0.8596	1.0000	1.0000
$1{,}200 \leq C_{SWA} \leq 1{,}249$	162	0.8395	1.0000	1.0000
$1{,}250 \leq C_{SWA} \leq 1{,}299$	174	0.8563	0.9885	1.0000
$1{,}300 \leq C_{SWA} \leq 1{,}349$	164	0.8659	0.9939	1.0000
$1{,}350 \leq C_{SWA} \leq 1{,}399$	166	0.8614	1.0000	1.0000
$1{,}400 \leq C_{SWA} \leq 1{,}449$	134	0.9030	0.9851	0.9925
$1{,}450 \leq C_{SWA} \leq 1{,}499$	153	0.8562	0.9935	1.0000
$1{,}500 \leq C_{SWA} \leq 1{,}549$	147	0.9320	1.0000	1.0000
$1{,}550 \leq C_{SWA} \leq 1{,}599$	135	0.8741	1.0000	1.0000
$1{,}600 \leq C_{SWA} \leq 1{,}649$	124	0.9516	1.0000	1.0000
$1{,}650 \leq C_{SWA} \leq 1{,}699$	139	0.9137	1.0000	1.0000
$1{,}700 \leq C_{SWA} \leq 1{,}749$	141	0.8794	1.0000	1.0000
$1{,}750 \leq C_{SWA} \leq 1{,}799$	127	0.8740	1.0000	1.0000
$1{,}800 \leq C_{SWA} \leq 1{,}849$	134	0.8881	1.0000	1.0000
$1{,}850 \leq C_{SWA} \leq 1{,}899$	117	0.8803	1.0000	1.0000
$1{,}900 \leq C_{SWA} \leq 1{,}949$	108	0.9259	0.9907	1.0000
$1{,}950 \leq C_{SWA} \leq 1{,}999$	120	0.9000	1.0000	1.0000
$2{,}000 \leq C_{SWA} \leq 2{,}049$	106	0.9340	1.0000	1.0000
$2{,}050 \leq C_{SWA} \leq 2{,}099$	100	0.8700	0.9900	1.0000
$2{,}100 \leq C_{SWA} \leq 2{,}149$	110	0.9000	1.0000	1.0000
$2{,}150 \leq C_{SWA} \leq 2{,}199$	105	0.9429	1.0000	1.0000
$2{,}200 \leq C_{SWA} \leq 2{,}249$	95	0.9474	1.0000	1.0000
$2{,}250 \leq C_{SWA} \leq 2{,}299$	77	0.9351	1.0000	1.0000
$2{,}300 \leq C_{SWA} \leq 2{,}349$	111	0.8919	1.0000	1.0000
$2{,}350 \leq C_{SWA} \leq 2{,}399$	109	0.9450	1.0000	1.0000
$2{,}400 \leq C_{SWA} \leq 2{,}449$	83	0.9398	1.0000	1.0000
$2{,}450 \leq C_{SWA} \leq 2{,}499$	94	0.9149	1.0000	1.0000
Total	199,698			

.

Table 4a: For Each Run, the Stratum and Stratum Proportion When 0.9500 First Exceeded
(Proportion computations result from 2020 Census Redistricting Data Production Settings ($\epsilon = 17.14$ for persons) version of *TDA*.)
Population: United States (50 States & DC)

		Criterion III LDG $DR_g \leq 0.05$
TDA Run	Stratum for Places & MCDs	Proportion When 0.9500 First Exceeded
1	$150 \leq C_{SWA} \leq 199$	0.9504
2	$200 \leq C_{SWA} \leq 249$	0.9548
3	$150 \leq C_{SWA} \leq 199$	0.9566
4	$150 \leq C_{SWA} \leq 199$	0.9504
5	$200 \leq C_{SWA} \leq 249$	0.9632
6	$200 \leq C_{SWA} \leq 249$	0.9632
7	$200 \leq C_{SWA} \leq 249$	0.9615
8	$200 \leq C_{SWA} \leq 249$	0.9615
9	$200 \leq C_{SWA} \leq 249$	0.9582
10	$150 \leq C_{SWA} \leq 199$	0.9643
11	$150 \leq C_{SWA} \leq 199$	0.9566
12	$150 \leq C_{SWA} \leq 199$	0.9504
13	$200 \leq C_{SWA} \leq 249$	0.9615
14	$150 \leq C_{SWA} \leq 199$	0.9550
15	$200 \leq C_{SWA} \leq 249$	0.9565
16	$150 \leq C_{SWA} \leq 199$	0.9550
17	$200 \leq C_{SWA} \leq 249$	0.9515
18	$150 \leq C_{SWA} \leq 199$	0.9519
19	$150 \leq C_{SWA} \leq 199$	0.9504
20	$200 \leq C_{SWA} \leq 249$	0.9532
21	$200 \leq C_{SWA} \leq 249$	0.9615
22	$200 \leq C_{SWA} \leq 249$	0.9548
23	$150 \leq C_{SWA} \leq 199$	0.9566
24	$150 \leq C_{SWA} \leq 199$	0.9550
25	$150 \leq C_{SWA} \leq 199$	0.9519

Using the data that will be released to the public (one run of the 2020 Census Production Settings version of *TDA*), we might say (as we did with block groups), empirically based on the data for the MCDs and places used in our study, that

> *"for any MCD or place with a TOTAL count between 200 and 249 people, the difference between the TDA ratio of the largest demographic group (LDG) and the corresponding SWA ratio for the LDG is less than or equal to 5 percentage points at least 95% of the time".*

Analysis of Congressional and State Legislative Districts

Another type of defined geography that is not a part of this nesting includes Congressional districts and state legislative districts. As we will see with Rhode Island in Part II of this study report, each state has Congressional district(s) (CD), state legislative districts in an upper chamber (SLDU), and state legislative districts in a lower chamber (SLDL).

As with the summary display in Table 3a for block groups and the summary display in Table 4a for places and MCDs, we use results from the 25 runs for all "Congressional and state legislative districts". Altogether, we make use of all 7,167 (= 436 + 1,946 + 4,785) Congressional and state legislative districts in the United States. The Table below gives a few parameters for the national accounting of these districts.

	CD	SLDU	SLDL
Number of Districts	436	1,946	4,785
Min Population	526,283	13,629	3,173
Median Population	705,831	121,212	41,713
Mean Population	708,132	158,656	64,016
Max Population	989,415	940,612	470,325

Again using Criterion III for all Congressional and state legislative districts in the United States, the stratum, where we first observed that 0.9500 was exceeded is given in Table 5 for each run. We display the entire table to emphasize that for each and every one of these districts, the size is sufficiently large to believe that the *TDA* counts are reliable for the largest demographic group (LDG) "all" of the time (based on our data).

Using the data that will be released to the public (one run of the 2020 Census redistricting data production settings version of *TDA*), we might say (as we did with block groups, also with MCDs and places) based on Table 5, that

> *"for all Congressional and state legislative districts, the difference between the TDA ratio of the largest demographic group (LDG) and the corresponding SWA ratio for the LDG is less than or equal to 5 percentage points 100% of the time".*

Table 5: For Each Run, the Stratum and Stratum Proportion When 0.9500 First Exceeded
(Proportion computations result from 2020 Census Redistricting Data Production Settings ($\epsilon = 17.14$ for persons) version of *TDA*.)
Population: United States (50 States & DC)

		Criterion III LDG $DR_g \leq 0.05$
TDA Run	Stratum for Congressional & State Legislative Districts	Proportion When 0.9500 First Exceeded
1	$3,150 \leq C_{SWA} \leq 3,199$	1.0000
2	$3,150 \leq C_{SWA} \leq 3,199$	1.0000
3	$3,150 \leq C_{SWA} \leq 3,199$	1.0000
4	$3,150 \leq C_{SWA} \leq 3,199$	1.0000
5	$3,150 \leq C_{SWA} \leq 3,199$	1.0000
6	$3,150 \leq C_{SWA} \leq 3,199$	1.0000
7	$3,150 \leq C_{SWA} \leq 3,199$	1.0000
8	$3,150 \leq C_{SWA} \leq 3,199$	1.0000
9	$3,150 \leq C_{SWA} \leq 3,199$	1.0000
10	$3,150 \leq C_{SWA} \leq 3,199$	1.0000
11	$3,150 \leq C_{SWA} \leq 3,199$	1.0000
12	$3,150 \leq C_{SWA} \leq 3,199$	1.0000
13	$3,150 \leq C_{SWA} \leq 3,199$	1.0000
14	$3,150 \leq C_{SWA} \leq 3,199$	1.0000
15	$3,150 \leq C_{SWA} \leq 3,199$	1.0000
16	$3,150 \leq C_{SWA} \leq 3,199$	1.0000
17	$3,150 \leq C_{SWA} \leq 3,199$	1.0000
18	$3,150 \leq C_{SWA} \leq 3,199$	1.0000
19	$3,150 \leq C_{SWA} \leq 3,199$	1.0000
20	$3,150 \leq C_{SWA} \leq 3,199$	1.0000
21	$3,150 \leq C_{SWA} \leq 3,199$	1.0000
22	$3,150 \leq C_{SWA} \leq 3,199$	1.0000
23	$3,150 \leq C_{SWA} \leq 3,199$	1.0000
24	$3,150 \leq C_{SWA} \leq 3,199$	1.0000
25	$3,150 \leq C_{SWA} \leq 3,199$	1.0000

I.5. CONCLUDING REMARKS FOR PART I

Remark 1: Within each of the criterion columns across Table 3, the values of the proportions tend to increase (though not always) as we go from the stratum with the smallest block groups to the stratum with the largest block groups using the *SWA* TOTAL counts. Also, the values of the proportions within a stratum (row) do increase as we go from Criterion I to Criterion III. From Table 3, we believe that a value for C^*_{SWA} can be produced (which is based on the data used in this study). This C^*_{SWA} is an empirical result. We can make similar statements relating to MCDs and places using Table 4, as well statements relating to Congressional and state legislative districts using Table 5.

Remark 2: Much of our focus in Part I has been in the context of the total population characteristics for block groups, MCDs and places, and Congressional and state legislative districts. In Table 3′ of APPENDIX C, we performed an analysis for the over 18 years and over population characteristics for block groups similar to what was done in Table 3 for the total population characteristics. We observed that the 5 percentage point criterion is reached 95% of the time for TOTAL18 in block groups whose size range between 500 and 549 people.

Remark 3: While small demographic groups are important, in the context of redistricting, it is the largest among the demographic groups that have the potential to form electoral districts where sufficiently large (and compact) minority groups have the opportunity "to elect representatives of their choice". We believe that support for consideration of the largest demographic group(s) is as noted in Section 2 of the *Voting Rights Act of 1965 (as amended)* and is called for by one of the three Gingles Requirements in the U.S. Supreme Court case *Thornburg v. Gingles (1986)* when establishing a violation of Section 2.

We understand that the potential for creating an electoral district that provides minority citizens with the opportunity to elect candidates of their choice is not necessarily limited to those block groups in which that group is the "largest demographic group". For example, a demographic group could comprise the second largest population group in two or more contiguous, randomly-created block groups. A different configuration of constituent blocks could result in that group being the basis of a district that affords the requisite opportunity to elect a candidate of their choice.

Part II

VARIABILITY ASSESSMENT OF DATA TREATED BY THE TOPDOWN ALGORITHM

II.1. INTRODUCTION

Part II is an update of our results in [6] where $\epsilon = 10.3$ and the 2021-04-28 version of *TDA* was used; whereas, throughout this study, $\epsilon = 17.14$ and further advances have been made resulting in the 2020 Census redistricting data production settings version of *TDA*. We reuse wording from [6] in many places; we do this in an attempt to repeat some of what we feel is important and in making this a more complete document. Of course, specific data results will differ.

As in [6], the specific focus of Part II is whether the explicitly acknowledged randomness used in the *TDA* for disclosure avoidance in the 2020 Census delivers official data that are fit for the development and analysis of redistricting plans. That randomness is characterized in this paper by measures of the variability observed in 25 runs of the same version of the *TDA* using the same allocation of the privacy-loss budget in each run ($\epsilon = 17.14$). The variability inherent in the official 2010 *PL-94-171* redistricting data resulted primarily from disclosure avoidance via household swapping. The parameters defining the rule(s) used in swapping that resulted in the official 2010 redistricting data are confidential and no estimates of the resulting variability have ever been published, including in this paper. Our approach (in the rest of this study as was the case in our earlier study [6]) has two parts: (1) to report observations on variability of results among 25 runs of the *TDA* [1] for Rhode Island and (2) to report observations on variability between the results among the 25 runs of the *TDA* and the published 2010 Census *Public Law 94-171* data for Rhode Island. In Part II, we also repeat these two-part analyses for three specific cases provided by the DOJ.

2010 Census Data for Rhode Island

The *TDA* was applied to data in the 2010 CEF for Rhode Island twenty-five different times, which we refer to as twenty-five runs of the *TDA*. For each run and for each of the 25,181 blocks in Rhode Island in the 2010 Census, various demographic variables report counts of various combinations of race, ethnicity (Hispanic or not Hispanic), and age.

Rhode Island has two (2) Congressional districts (CD), 38 state legislative districts (SLDU) in its upper legislative chamber, and 75 state legislative districts (SLDL) in its lower legislative chamber. These form the foundation of our case study for Rhode Island.

2010 Census Data for Three Cases Provided by DOJ

For three cases (jurisdictions) provided by DOJ, we conduct similar analyses of data in Section II.6 as just described for Rhode Island. The three cases are Panola County, Mississippi (MS) (2,180 blocks); Tate County (School District), MS (784 blocks); and Tylertown (Walthall County), MS (136 blocks).

Overview of Part II

An overview of Part II follows. In Section II.2 of this report, we present data for the two Congressional districts of Rhode Island and using formatted data tables as shown in Table 6. Section II.3 visually compares 2010 CEF data treated by the disclosure avoidance method (swapping [6]) with randomly selected runs of the same 2010 CEF data treated by the *TDA* method (i.e., differential privacy) being planned for use by the 2020 Census. Section II.4 is similar to Section II.3 except the visual comparisons are for four of Rhode Island's Upper Chamber Districts. Section II.5 is similar to Sections II.3 and II.4 except the visual comparisons are for four of Rhode Island's Lower Chamber Districts. Section II.6 investigates three cases provided by DOJ using varying (mainly smaller) total population and varying group composition selected for comparisons similar to those of previous Sections for CDs, SLDUs, and SLDLs. Section II.7 defines and looks at variablility among the 25 *TDA* runs of Rhode Island data using the planned *TDA* method of 2020, and it also looks at variability among the 25 *TDA* runs in comparison with the public data for Rhode Island from 2010 (this section also presents similar tables for the three cases provided by DOJ). The insert following Table 6 gives a suggestion for reviewing the tables of counts and percentages. The key empirical message on variability is given in the last paragraph of Section II.7. Section II.8 provides some concluding remarks based on the tables. The APPENDICES follow Section II.8.

II.2. FORMAT OF COUNTS & PERCENTAGES TABLES USED IN OUR STUDY

Table 6 shows the redistricting plan (POST-2010) adopted by Panola County, Mississippi. Panola County, with five (5) districts, has an overall population (TOTAL) of 34,707 people based on the 2010 Census. The average population per district (IDEAL POPULATION) is 34,707/5 = 6,941 people. Using the POST-2010 plan, the deviations from the IDEAL POPULATION for each of the 5 districts (DEV) are 33, -392, 133, 164, and 64, respectively; and the corresponding percent deviations (DEV = DEV/6941) × 100% are respectively: 0.48%, -5.65%, -1.92%, 2.36%, and 0.92%. From Table 6, it is noteworthy that the demographic group of WHITENH has 16,981 people which is WHITENHP = 48.93% of the county's population while the demographic group BLACKNH has 16,899 people which is BLACKNHP = 48.69% of the county's population. Other demographic group characteristics in Table 6 are given for the 18 years and over population (TOTAL18).

Table 6. POST-2010 Census Demographics, Counts, & Percentages: Panola County, Mississippi

Demographics		Counts & Percentages by District (POST-2010)				
DIST-ID	Panola	1	2	3	4	5
TOTAL	34,707	6,974	6,549	7,074	7,105	7,005
DEV		33	-392	133	164	64
DEVP		0.48	-5.65	1.92	2.36	0.92
TOTAL18	25,363	5,214	4,732	5,171	5,345	4,901
TOTALHISP	494	66	75	85	120	148
TOTALHISPP	1.42	0.95	1.15	1.20	1.69	2.11
TOTALNH	34,213	6,908	6,474	6,989	6,985	6,857
TOTALNHP	98.58	99.05	98.85	98.8	98.31	97.89
WHITENH	16,981	2,419	2,096	4,030	5,250	3,186
WHITENHP	48.93	34.69	32.00	56.97	73.89	45.48
BLACKNH	16,899	4,427	4,332	2,925	1,658	3,557
BLACKNHP	48.69	63.48	66.15	41.35	23.34	50.78
AIANNH	148	26	20	15	38	49
AIANNHP	0.43	0.37	0.31	0.21	0.53	0.70
ASIANNH	89	8	7	5	17	52
ASIANNHP	0.26	0.11	0.11	0.07	0.24	0.74
HPINH	4	0	0	0	2	2
HPINHP	0.01	0.00	0.00	0.00	0.03	0.03
OTHERNH	19	7	5	1	3	3
OTHERNHP	0.05	0.10	0.08	0.01	0.04	0.04
MLTMNNH	73	21	14	13	17	8
MLTMNNHP	0.21	0.30	0.21	0.18	0.24	0.11
HISP18	298	44	44	52	63	95
HISP18P	1.17	0.84	0.93	1.01	1.18	1.94
NONHISP18	25,065	5,170	4,688	5,119	5,282	4,806
NONHISP18P	98.83	99.16	99.07	98.99	98.82	98.06
WHITENH18	13,455	2,025	1,732	3,072	4,115	2,511
WHITENH18P	53.05	38.84	36.6	59.41	76.99	51.23
BLACKNH18	11,394	3,099	2,928	2,024	1,118	2,225
BLACKNH18P	44.92	59.44	61.88	39.14	20.92	45.40
AIANNH18	115	21	16	11	29	38
AIANNH18P	0.45	0.40	0.34	0.21	0.54	0.78
ASIANNH18	54	8	5	2	12	27
ASIANNH18P	0.21	0.15	0.11	0.04	0.22	0.55
HPINH18	2	0	0	0	1	1
HPINH18P	0.01	0.00	0.00	0.00	0.02	0.02
OTHERNH18	5	1	0	1	2	1
OTHERNH18P	0.02	0.02	0.00	0.02	0.04	0.02
MLTMNH18	40	16	7	9	5	3
MLTMNH18P	0.16	0.31	0.15	0.17	0.09	0.06

Source: U.S. Department of Justice, Washington, D.C.

A Suggestion from the Authors for Reviewing Each Table

When we inspect the various tables that follow in this study, we first look at the column of overall counts and percentages for the various demographic groups in a jurisdiction (e.g., state or county or school district) and then ask how these counts and percentages are distributed over the various districts.

II.3. EXAMINATION OF RHODE ISLAND CONGRESSIONAL DISTRICT DATA

Table 7 shows results from three randomly chosen runs of the twenty-five runs of the *TDA* for Congressional Districts CD-01 and CD-02 for Rhode Island (last six columns) and displays them with the counts from the 2010 Census (alternately referred to as swapping or Summary File 1 (SF1) in this part of our study) relative to the boundaries for the 113^{th} Congress. These three runs provide a taste of what variability might be expected among the various runs of the *TDA*. Throughout this report, we use the same value of $\epsilon = 17.14$, and exactly the same implementation code and parameters, for all discussed runs of the *TDA*.

In Table 7, we also compare the results for CD-01 and CD-02 from each of the three *TDA* runs with the corresponding published results (2010 Census, SF1) for CD-01 and CD-02.

From Table 7, while the corresponding counts for each demographic group (on each row) vary among the runs as well as relative to the released 2010 Census counts, the corresponding percentages displayed differ by less than 0.5 of a percentage point for all demographic groups. The fact that the DEV values for the three runs differ from -0.5 and 0.5 should be of no concern because the 2020 Congressional redistricting would use the noise-infused block level counts to create Congressional districts where the DEV values differ by no more than 1 person. In general, state legislative districts are allowed to deviate by more than 1 person.

In Table 7, note that CD-01 has smaller counts for WHITENH than CD-02 using the 2010 Census counts. As a consequence, CD-01 has comparatively larger counts for most minority demographic groups than CD-02. This observation is true for the total population group counts as well as for the 18 and older population groups. This observation tends to also hold for each of the three *TDA* runs. (The same holds true for WHITENH18 and most minority groups in the 18 and older population.)

II.4. EXAMINATION OF RHODE ISLAND's 38 UPPER CHAMBER DISTRICTS

There are 38 districts with one legislator each in Rhode Island's Upper Chamber. Therefore, the IDEAL POPULATION for each State Upper Chamber District is $\dfrac{1,052,567}{38} = 27,699.1$. Columns 2-5 of Table 8 give 2010 Census counts and percentages for the State Upper Chamber Districts (SLDU) 01, 02, 03, and 04. Columns 6-9 of Table 8 give corresponding counts and percentages from the same *TDA* Run A noted in Table 7.

For the 2010 Census counts as well as the counts for the *TDA* Run A, SLDU-02 has relatively high percentages for both TOTALHISPP and HISP18P. Similarly, for the 2010 Census counts as well as for the *TDA* Run A, SLDU-03 and SLDU-04 each has relatively high percentages for both WHITENHP and WHITENH18P. SLDU-01 has a relatively high percentage total for TOTALHISPP and BLACKNHP. The same holds true in SLDU-01 for HISP18P and BLACKNH18P.

Table 7. Rhode Island: Three of Twenty-five Runs of the *TDA* by Congressional Districts (CDs) for the 113^{th} Congress

($C_{TDA}(g)$ counts result from 2020 Census Redistricting Data Production Settings ($\epsilon = 17.14$ for persons) version of *TDA*.)

		2010 Census, SF1 (*PL 94-171*)(2013) Counts & Percentages POST-2010 Plan		Counts & Percentages, 113^{th} Congress 3 Out of 25 Runs of the *TDA*					
Demographics		113^{th} Congress		*TDA*-Run A		*TDA*-Run B		*TDA*-Run C	
DIST-ID	Rhode Island	CD-01	CD-02	CD-01	CD-02	CD-01	CD-02	CD-01	CD-02
TOTAL	1,052,567	526,283	526,284	526,125	526,442	526,120	526,447	526,140	526,427
DEV		-0.5	0.5	-158.5	158.5	-163.5	163.5	-143.5	143.5
DEVP		0.00	0.00	-0.03	0.03	-0.03	0.03	-0.03	0.03
TOTAL18	828,611	412,778	415,833	412,604	416,009	412,633	415,956	412,648	415,951
TOTALHISP	130,655	76,100	54,555	76,050	54,607	76,081	54,575	76,070	54,603
TOTALHISPP	12.41	14.46	10.37	14.45	10.37	14.46	10.37	14.46	10.37
TOTALNH	921,912	450,183	471,729	450,075	471,835	450,039	471,872	450,070	471,824
TOTALNHP	87.59	85.54	89.63	85.55	89.63	85.54	89.63	85.54	89.63
WHITENH	803,685	377,109	426,576	377,055	426,626	377,035	426,649	377,057	426,628
WHITENHP	76.35	71.66	81.05	71.67	81.04	71.66	81.04	71.66	81.04
BLACKNH	57,927	37,627	20,300	37,600	20,340	37,678	20,255	37,619	20,301
BLACKNHP	5.50	7.15	3.86	7.15	3.86	7.16	3.85	7.15	3.86
AIANNH	6,839	3,142	3,697	3,161	3,694	3,119	3,739	3,084	3,759
AIANNHP	0.65	0.60	0.70	0.60	0.70	0.59	0.71	0.59	0.71
ASIANNH	34,194	17,705	16,489	17,664	16,518	17,676	16,516	17,722	16,473
ASIANNHP	3.25	3.36	3.13	3.36	3.14	3.36	3.14	3.37	3.13
HPINH	655	383	272	392	263	411	258	361	294
HPINHP	0.06	0.07	0.05	0.07	0.05	0.08	0.05	0.07	0.06
OTHERNH	10,296	8,492	1,804	8,484	1,818	8,487	1,802	8,517	1,763
OTHERNHP	0.98	1.61	0.34	1.61	0.35	1.61	0.34	1.62	0.33
MLTMNNH	8,316	5,725	2,591	5,719	2,576	5,633	2,653	5,710	2,606
MLTMNNHP	0.79	1.09	0.49	1.09	0.49	1.07	0.50	1.09	0.50
HISP18	84,715	49,303	35,412	49,146	35,562	49,240	35,469	49,237	35,487
HISP18P	10.22	11.94	8.52	11.91	8.55	11.93	8.53	11.93	8.53
NONHISP18	743,896	363,475	380,421	363,458	380,447	363,393	380,487	363,411	380,464
NONHISP18P	89.78	88.06	91.48	88.09	91.45	88.07	91.47	88.07	91.47
WHITENH18	660,823	312,240	348,583	312,217	348,607	312,174	348,648	312,219	348,605
WHITENH18P	79.75	75.64	83.83	75.67	83.80	75.65	83.82	75.66	83.81
BLACKNH18	39,485	25,402	14,083	25,392	14,105	25,458	14,029	25,405	14,081
BLACKNH18P	4.77	6.15	3.39	6.15	3.39	6.17	3.37	6.16	3.39
AIANNH18	4,963	2,332	2,631	2,344	2,626	2,298	2,684	2,306	2,657
AIANNH18P	0.60	0.56	0.63	0.57	0.63	0.56	0.65	0.56	0.64
ASIANNH18	25,333	13,276	12,057	13,252	12,072	13,235	12,090	13,280	12,051
ASIANNH18P	3.06	3.22	2.90	3.21	2.90	3.21	2.91	3.22	2.90
HPINH18	500	307	193	305	189	326	186	302	195
HPINH18P	0.06	0.07	0.05	0.07	0.05	0.08	0.04	0.07	0.05
OTHERNH18	7,290	6,061	1,229	6,069	1,227	6,113	1,177	6,070	1,208
OTHERNH18P	0.88	1.47	0.30	1.47	0.29	1.48	0.28	1.47	0.29
MLTMNH18	5,502	3,857	1,645	3,879	1,621	3,789	1,673	3,829	1,667
MLTMNH18P	0.66	0.93	0.40	0.94	0.39	0.92	0.40	0.93	0.40

Source: Data from 3 Runs of the *TDA*, U. S. Bureau of the Census, Washington, D.C.

Selected observations for Table 7:

1: Corresponding percentages between the 2010 Census data and the *TDA* data on each row displayed in Table 7 differ by less than 0.5 of a percentage point for all demographic groups.

2: CD-01 has lower counts for WHITENH (also WHITENH18) than CD-02 when using the 2010 Census counts. As a consequence, CD-01 has comparatively larger counts for most minority demographic groups than CD-02. The same relationships between the CD-01 and CD-02 data hold for these demographic groups within the 18 and older population groups. This observation also tends to hold for each of the three *TDA* runs.

Table 8. Rhode Island Run A of Twenty-five Runs of the *TDA*
for State Upper Chamber Districts (SLDU) 01, 02, 03, and 04 (4 of 38 Districts)
($C_{TDA}(g)$ counts result from 2020 Census Redistricting Data Production Settings ($\epsilon = 17.14$ for persons) version of *TDA*.)

	2010 Census, SF1 (*PL 94-171*) (2013) Counts & Percentages POST-2010 Plan				Counts & Percentages, 2013 Run A of the *TDA*			
Demographics								
DIST-ID	SLDU-01	SLDU-02	SLDU-03	SLDU-04	SLDU-01	SLDU-02	SLDU-03	SLDU-04
TOTAL	28,161	28,079	28,398	28,201	28,095	28,151	28,614	28,196
DEV	461.9	379.9	698.9	501.9	395.9	451.9	914.9	496.9
DEVP	1.64	1.35	2.46	1.78	1.41	1.61	3.20	1.76
TOTAL18	20,914	19,846	25,361	23,599	20,876	19,853	25,435	23,579
TOTALHISP	10,282	16,288	1,409	3,217	10,254	16,312	1,510	3,177
TOTALHISPP	36.51	58.01	4.96	11.41	36.50	57.94	5.28	11.27
TOTALNH	17,879	11,791	26,989	24,984	17,841	11,839	27,104	25,019
TOTALNHP	63.49	41.99	95.04	88.59	63.50	42.06	94.72	88.73
WHITENH	10,222	3,553	22,028	21,210	10,192	3,510	22,018	21,247
WHITENHP	36.30	12.65	77.57	75.21	36.28	12.47	76.95	75.35
BLACKNH	4,862	4,332	1,124	2,348	4,888	4,337	1,175	2,345
BLACKNHP	17.27	15.43	3.96	8.33	17.40	15.41	4.11	8.32
AIANNH	283	216	135	172	267	213	137	201
AIANNHP	1.00	0.77	0.48	0.61	0.95	0.76	0.48	0.71
ASIANNH	1,526	3,032	3,262	826	1,541	3,071	3,273	806
ASIANNHP	5.42	10.80	11.49	2.93	5.48	10.91	11.44	2.86
HPINH	25	11	16	14	19	18	19	12
HPINHP	0.09	0.04	0.06	0.05	0.07	0.06	0.07	0.04
OTHERNH	457	189	224	241	454	208	243	236
OTHERNHP	1.62	0.67	0.79	0.85	1.62	0.74	0.85	0.84
MLTMNNH	504	458	200	173	480	482	239	172
MLTMNNHP	1.79	1.63	0.70	0.61	1.71	1.71	0.84	0.61
HISP18	6,458	11,014	1,241	2,097	6,430	10,991	1,244	2,067
HISP18P	30.88	55.50	4.89	8.89	30.80	55.36	4.89	8.77
NONHISP18	14,456	8,832	24,120	21,502	14,446	8,862	24,191	21,512
NONHISP18P	69.12	44.50	95.11	91.11	69.20	44.64	95.11	91.23
WHITENH18	9,131	3,062	19,682	18,839	9,111	3,055	19,679	18,850
WHITENH18P	43.66	15.43	77.61	79.83	43.64	15.39	77.37	79.94
BLACKNH18	3,309	3,027	973	1,599	3,335	3,008	1,005	1,612
BLACKNH18P	15.82	15.25	3.84	6.78	15.98	15.15	3.95	6.84
AIANNH18	197	154	110	136	185	146	107	151
AIANNH18P	0.94	0.78	0.43	0.58	0.89	0.74	0.42	0.64
ASIANNH18	1,170	2,135	2,989	611	1,163	2,156	3,001	600
ASIANNH18P	5.59	10.76	11.79	2.59	5.57	10.86	11.80	2.51
HPINH18	20	11	14	13	17	18	16	9
HPINH18P	0.10	0.06	0.06	0.06	0.08	0.09	0.06	0.04
OTHERNH18	326	125	186	178	332	134	203	167
OTHERNH18P	1.56	0.63	0.73	0.75	1.59	0.67	0.80	0.71
MLTMNH18	303	318	166	126	303	345	180	123
MLTMNH18P	1.45	1.60	0.65	0.53	1.45	1.74	0.71	0.52

Source: Data from Run A of the *TDA*, U. S. Bureau of the Census, Washington, D.C.

Selected observations for Table 8:

1: SLDU-01 has percentage total $\geq$ 50% for TOTALHISPP and BLACKNHP (also HISP18P and BLACKNH18P) for 2010 Census and the *TDA* run.

2: SLDU-02 has percentages $\geq$ 50% for both TOTALHISPP and HISP18P for 2010 Census and the *TDA* run.

3: SLDU-03 and SLDU-04 each has a percentage $\geq$ 50% for both WHITENHP and WHITENH18P for the 2010 Census and the *TDA* run.

II.5. EXAMINATION OF RHODE ISLAND's 75 LOWER CHAMBER DISTRICTS

There are 75 districts with one legislator each in Rhode Island's Lower Chamber. Therefore, the IDEAL POPULATION for each State Lower Chamber District (SLDL) is $\frac{1,052,567}{75} = 14,034.2$. As with Table 8 for Rhode Island's Upper Chamber, Columns 2-5 of Table 9 give 2010 Census counts and percentages for the State Lower Chamber Districts 01, 02, 03, and 04. Columns 6-9 of Table 9 give corresponding counts and percentages from the same *TDA* Run A noted in Table 7.

For the 2010 Census counts as well as for the *TDA* Run A, note the SLDL-03 has a relatively high percentage total for TOTALHISPP and BLACKNHP as well as a high percentage total for HISP18P and BLACKNH18P. Similarly for the 2010 Census counts as well as for the *TDA* Run A, note that SLDL-01, SLDL-02, and SLDL-04 each has relatively high percentages for both WHITENHP and WHITENH18P.

Unlike in Table 7 for the congressional districts, the corresponding percentages for the demographic groups in the Lower Chamber Districts differ by approximately 1 percentage point. Thus we see more variability for lower levels of geography.

II.6. EXAMINATION OF THREE CASES PROVIDED BY DOJ

To examine variability for each of the cases provided by DOJ, we proceed for each as we did with Rhode Island. A high level overview of the three cases follows

	Jurisdiction	2010 Census Population	Number of Districts	Number of Blocks Overall	Number of Blocks by Districts
1.	Panola County, MS	34,707	5	2,180	(458; 492; 413; 443; 374)
2.	Tate County, MS (School District)	18,823	5	784	(168; 204; 139; 178; 95)
3.	Tylertown, MS (Walthall County)	1,609	4	136	(35; 42; 42; 17)

Panola County, MS: In Table 10, the 2010 Census data show, WHITENHP = 48.93% and BLACKNHP = 48.69% for the overall county as noted earlier in Table 6. For the same data, and for districts 01, 02, and 05, we see BLACKNHP values of 63.48%, 66.15%, and 50.78%, respectively; for districts 03 and 04, we see WHITENHP values of 56.97% and 73.89%, respectively. We see similar corresponding percentages for the results from the *TDA*.

Tate County (School District), MS: In Table 11, the 2010 Census data show WHITENHP = 68.22% and BLACKNHP = 28.63% for the overall county. In addition, the 2010 Census data for districts 01, 03, 04, and 05 show WHITENHP values of 86.31%, 78.04%, 62.02%, and 73.40%, respectively; for district 02, we see BLACKNHP = 54.94%. We see similar corresponding percentages for the results from the *TDA*.

Tylertown (Walthall County), MS: In Table 12, the 2010 Census data show WHITENHP = 53.45% and BLACKNHP = 42.20% for Tylertown (the county seat of Walthall County) overall. For the same data, and for districts 01, 02, and 03, we see WHITENHP values of 91.60%, 53.88%, and 62.92%, respectively; for district 04, we see BLACKNHP = 89.13%. We see less similar corresponding percentages for the results from the *TDA* for Tylertown than we see for Panola and Tate.

Table 9. Rhode Island Run A of Twenty-five Runs of the *TDA*
for State Lower Chamber Districts (SLDL) 01, 02, 03, and 04 (4 of 75 Districts)
($C_{TDA}(g)$ counts result from 2020 Census Production Redistricting Data Settings ($\epsilon = 17.14$ for persons) version of *TDA*.)

	2010 Census, SF1 (*PL 94-171*) (2013) Counts & Percentages POST-2010 Plan				Counts & Percentages, 2013 Run A of the *TDA*			
Demographics								
DIST-ID	SLDL-01	SLDL-02	SLDL-03	SLDL-04	SLDL-01	SLDL-02	SLDL-03	SLDL-04
TOTAL	13,881	13,821	13,949	13,713	14,056	13,729	13,790	13,617
DEV	-153.2	-213.2	-85.2	-321.2	21.8	-305.2	-244.2	-417.2
DEVP	-1.10	-1.54	-0.61	-2.34	0.15	-2.22	-1.77	-3.06
TOTAL18	12,835	12,800	9,607	11,205	12,874	12,712	9,589	11,157
TOTALHISP	1,002	1,768	5,905	1,049	1,068	1,715	5,887	1,041
TOTALHISPP	7.22	12.79	42.33	7.65	7.60	12.49	42.69	7.64
TOTALNH	12,879	12,053	8,044	12,664	12,988	12,014	7,903	12,576
TOTALNHP	92.78	87.21	57.67	92.35	92.40	87.51	57.31	92.36
WHITENH	9,922	8,714	3,465	9,539	9,922	8,709	3,428	9,529
WHITENHP	71.48	63.05	24.84	69.56	70.59	63.44	24.86	69.98
BLACKNH	581	1,125	3,015	1,495	605	1,124	3,001	1,480
BLACKNHP	4.19	8.14	21.61	10.90	4.30	8.19	21.76	10.87
AIANNH	46	104	189	126	60	89	158	112
AIANNHP	0.33	0.75	1.35	0.92	0.43	0.65	1.15	0.82
ASIANNH	2,175	1,776	794	792	2,204	1,763	788	782
ASIANNHP	15.67	12.85	5.69	5.78	15.68	12.84	5.71	5.74
HPINH	12	16	12	1	11	10	8	3
HPINHP	0.09	0.12	0.09	0.01	0.08	0.07	0.06	0.02
OTHERNH	57	148	257	396	74	143	248	382
OTHERNHP	0.41	1.07	1.84	2.89	0.53	1.04	1.80	2.81
MLTMNNH	86	170	312	315	112	176	272	288
MLTMNNHP	0.62	1.23	2.24	2.30	0.80	1.28	1.97	2.12
HISP18	951	1,475	3,518	693	939	1,423	3,524	685
HISP18P	7.41	11.52	36.62	6.18	7.29	11.19	36.75	6.14
NONHISP18	11,884	11,325	6,089	10,512	11,935	11,289	6,065	10,472
NONHISP18P	92.59	88.48	63.38	93.82	92.71	88.81	63.25	93.86
WHITENH18	9,081	8,339	3,040	8,119	9,059	8,351	3,029	8,117
WHITENH18P	70.75	65.15	31.64	72.46	70.37	65.69	31.59	72.75
BLACKNH18	560	972	1,971	1,144	576	958	1,982	1,141
BLACKNH18P	4.36	7.59	20.52	10.21	4.47	7.54	20.67	10.23
AIANNH18	45	82	129	101	52	69	114	96
AIANNH18P	0.35	0.64	1.34	0.90	0.40	0.54	1.19	0.86
ASIANNH18	2,052	1,655	575	635	2,082	1,649	567	631
ASIANNH18P	15.99	12.93	5.99	5.67	16.17	12.97	5.91	5.66
HPINH18	10	14	11	1	11	7	7	3
HPINH18P	0.08	0.11	0.11	0.01	0.09	0.06	0.07	0.03
OTHERNH18	51	126	190	280	65	122	193	268
OTHERNH18P	0.40	0.98	1.98	2.50	0.50	0.96	2.01	2.40
MLTMNH18	85	137	173	232	90	133	173	216
MLTMNH18P	0.66	1.07	1.80	2.07	0.70	1.05	1.80	1.94

Source: Data from Run A of the *TDA*, U. S. Bureau of the Census, Washington, D.C.

Selected observations for Table 9:

1: SLDL-01, SLDL-02, and SLDL-04 each has a percentage $\geq$ 50% for both WHITENHP and WHITENH18P for 2010 Census and the *TDA* run.

2: SLDL-03 has a percentage total $\geq$ 50% for TOTALHISPP and BLACKNHP jointly, as well as a percentage total $\geq$ 50% for HISP18P and BLACKNH18P jointly for 2010 Census and the *TDA* run.

Table 10. Panola County, MS Run A of Twenty-five Runs of the *TDA* for County Districts 01, 02, 03, 04, and 05

($C_{TDA}(g)$ counts result from 2020 Census Redistricting Data Production Settings ($\epsilon = 17.14$ for persons) version of *TDA*.)

$$\text{2010 Census IDEAL POPULATION} = \frac{34,707}{5} = 6,941.4 \qquad \textit{TDA} \text{ IDEAL POPULATION} = \frac{34,710}{5} = 6,942.0$$

	2010 Census, SF1 (*PL 94-171*) Counts & Percentages POST-2010 Plan						Counts & Percentages Run A of the *TDA*					
Demographics												
DIST-ID	Panola	01	02	03	04	05	Panola	01	02	03	04	05
TOTAL	34,707	6,974	6,549	7,074	7,105	7,005	34,710	7,015	6,524	7,029	7,119	7,023
DEV		32.6	-392.4	132.6	163.6	63.6		73	-418	87	177	81
DEVP		0.47	-5.99	1.87	2.30	0.91		1.04	-6.41	1.24	2.49	1.15
TOTAL18	25,363	5,214	4,732	5,171	5,345	4,901	25,383	5,229	4,711	5,155	5,364	4,924
TOTALHISP	494	66	75	85	120	148	501	79	61	82	120	159
TOTALHISPP	1.42	0.95	1.15	1.20	1.69	2.11	1.44	1.13	0.94	1.17	1.69	2.26
TOTALNH	34,213	6,908	6,474	6,989	6,985	6,857	34,209	6,936	6,463	6,947	6,999	6,864
TOTALNHP	98.58	99.05	98.85	98.80	98.31	97.89	98.56	98.87	99.06	98.83	98.31	97.74
WHITENH	16,981	2,419	2,096	4,030	5,250	3,186	16,986	2,426	2,116	3,995	5,262	3,187
WHITENHP	48.93	34.69	32.00	56.97	73.89	45.48	48.94	34.58	32.43	56.84	73.91	45.38
BLACKNH	16,899	4,427	4,332	2,925	1,658	3,557	16,888	4,437	4,311	2,910	1,660	3,570
BLACKNHP	48.69	63.48	66.15	41.35	23.34	50.78	48.65	63.25	66.08	41.40	23.32	50.83
AIANNH	148	26	20	15	38	49	153	35	20	17	38	43
AIANNHP	0.43	0.37	0.31	0.21	0.53	0.70	0.44	0.50	0.31	0.24	0.53	0.61
ASIANNH	89	8	7	5	17	52	84	9	3	8	14	50
ASIANNHP	0.26	0.11	0.11	0.07	0.24	0.74	0.24	0.13	0.05	0.11	0.20	0.71
HPINH	4	0	0	0	2	2	9	0	1	5	2	1
HPINHP	0.01	0.00	0.00	0.00	0.03	0.03	0.03	0.00	0.02	0.07	0.03	0.01
OTHERNH	19	7	5	1	3	3	12	3	1	2	4	2
OTHERNHP	0.05	0.10	0.08	0.01	0.04	0.04	0.03	0.04	0.02	0.03	0.06	0.03
MLTMNNH	73	21	14	13	17	8	77	26	11	10	19	11
MLTMNNHP	0.21	0.30	0.21	0.18	0.24	0.11	0.22	0.37	0.17	0.14	0.27	0.16
HISP18	298	44	44	52	63	95	302	49	35	46	74	98
HISP18P	1.17	0.84	0.93	1.01	1.18	1.94	1.19	0.94	0.74	0.89	1.38	1.99
NONHISP18	25,065	5,170	4,688	5,119	5,282	4,806	25,081	5,180	4,676	5,109	5,290	4,826
NONHISP18P	98.83	99.16	99.07	98.99	98.82	98.06	98.81	99.06	99.26	99.11	98.62	98.01
WHITENH18	13,455	2,025	1,732	3,072	4,115	2,511	13,460	2,028	1,726	3,058	4,129	2,519
WHITENH18P	53.05	38.84	36.60	59.41	76.99	51.23	53.03	38.78	36.64	59.32	76.98	51.16
BLACKNH18	11,394	3,099	2,928	2,024	1,118	2,225	11,397	3,111	2,922	2,017	1,113	2,234
BLACKNH18P	44.92	59.44	61.88	39.14	20.92	45.40	44.90	59.50	62.03	39.13	20.75	45.37
AIANNH18	115	21	16	11	29	38	116	22	15	16	30	33
AIANNH18P	0.45	0.40	0.34	0.21	0.54	0.78	0.46	0.42	0.32	0.31	0.56	0.67
ASIANNH18	54	8	5	2	12	27	48	5	2	4	8	29
ASIANNH18P	0.21	0.15	0.11	0.04	0.22	0.55	0.19	0.10	0.04	0.08	0.15	0.59
HPINH18	2	0	0	0	1	1	9	0	1	5	2	1
HPINH18P	0.01	0.00	0.00	0.00	0.02	0.02	0.04	0.00	0.02	0.10	0.04	0.02
OTHERNH18	5	1	0	1	2	1	3	0	0	1	0	2
OTHERNH18P	0.02	0.02	0.00	0.02	0.04	0.02	0.01	0.00	0.00	0.02	0.00	0.04
MLTMNH18	40	16	7	9	5	3	48	14	10	8	8	8
MLTMNH18P	0.16	0.31	0.15	0.17	0.09	0.06	0.19	0.27	0.21	0.16	0.15	0.16

Source: Data from Run A of the *TDA*, U.S. Bureau of the Census, Washington, D.C.

Selected observations for Table 10:

1: Panola has WHITENHP = 48.93% and BLACKNHP = 48.69% for the 2010 Census; and WHITENHP = 48.94% and BLACKNHP = 48.65% for the *TDA* run. For 18^+ population, WHITENH18P $\geq$ 50.00% for the 2010 Census and for the *TDA* run.

2: Districts 01 and 02 each has a percentage $\geq$ 50% for BLACKNHP (also BLACKNH18P) for both the 2010 Census and the *TDA* run. District 05 has a BLACKNHP (also BLACKNH18P) percentage close to 50.00% for both the 2010 Census and the *TDA* run.

Table 11. Tate County School Districts (SD), MS Run A of Twenty-five Runs of the *TDA* for School Districts 01, 02, 03, 04, and 05

($C_{TDA}(g)$ counts result from 2020 Census Redistricting Data Production Settings ($\epsilon = 17.14$ for persons) version of *TDA*.)

$$\text{2010 Census IDEAL POPULATION} = \frac{18,823}{5} = 3,764.6 \qquad \textit{TDA}\ \text{IDEAL POPULATION} = \frac{18,813}{5} = 3,762.6$$

	2010 Census, SF1 (*PL 94-171*) Counts & Percentages POST-2010 Plan						Counts & Percentages Run A of the *TDA*					
Demographics												
DIST-ID	Tate	01	02	03	04	05	Tate	01	02	03	04	05
TOTAL	18,823	3,914	3,893	3,665	3,697	3,654	18,813	3,954	3,834	3,683	3,710	3,632
DEV		149.4	128.4	-99.6	-67.6	-110.6		191.4	71.4	-79.6	-52.6	-130.6
DEVP		3.82	3.30	-2.72	-1.83	-3.03		4.84	1.86	-2.16	-1.42	-3.60
TOTAL18	13,893	2,780	2,826	2,799	2,755	2,733	13,894	2,818	2,787	2,796	2,773	2,720
TOTALHISP	399	87	63	110	32	107	408	93	63	104	43	105
TOTALHISPP	2.12	2.22	1.62	3.00	0.87	2.93	2.17	2.35	1.64	2.82	1.16	2.89
TOTALNH	18,424	3,827	3,830	3,555	3,665	3,547	18,405	3,861	3,771	3,579	3,667	3,527
TOTALNHP	97.88	97.78	98.38	97.00	99.13	97.07	97.83	97.65	98.36	97.18	98.84	97.11
WHITENH	12,841	3,378	1,628	2,860	2,293	2,682	12,799	3,403	1,579	2,857	2,280	2,680
WHITENHP	68.22	86.31	41.82	78.04	62.02	73.40	68.03	86.06	41.18	77.57	61.46	73.79
BLACKNH	5,389	400	2,139	666	1,349	835	5,408	392	2,147	701	1,350	818
BLACKNHP	28.63	10.22	54.94	18.17	36.49	22.85	28.75	9.91	56.00	19.03	36.39	22.52
AIANNH	103	32	26	19	11	15	97	38	19	13	15	12
AIANNHP	0.55	0.82	0.67	0.52	0.30	0.41	0.52	0.96	0.50	0.35	0.40	0.33
ASIANNH	47	14	16	6	7	4	45	17	10	4	10	4
ASIANNHP	0.25	0.36	0.41	0.16	0.19	0.11	0.24	0.43	0.26	0.11	0.27	0.11
HPINH	3	2	0	0	0	1	4	2	0	0	0	2
HPINHP	0.02	0.05	0.00	0.00	0.00	0.03	0.02	0.05	0.00	0.00	0.00	0.06
OTHERNH	9	1	5	1	1	1	13	1	8	0	1	3
OTHERNHP	0.05	0.03	0.13	0.03	0.03	0.03	0.07	0.03	0.21	0.00	0.03	0.08
MLTMNNH	32	0	16	3	4	9	39	8	8	4	11	8
MLTMNNHP	0.17	0.00	0.41	0.08	0.11	0.25	0.21	0.20	0.21	0.11	0.30	0.22
HISP18	215	47	34	63	16	55	218	43	39	57	24	55
HISP18P	1.55	1.69	1.20	2.25	0.58	2.01	1.57	1.53	1.40	2.04	0.87	2.02
NONHISP18	13,678	2,733	2,792	2,736	2,739	2,678	13,676	2,775	2,748	2,739	2,749	2,665
NONHISP18P	98.45	98.31	98.80	97.75	99.42	97.99	98.43	98.47	98.60	97.96	99.13	97.98
WHITENH18	9,747	2,438	1,278	2,219	1,755	2,057	9,724	2,450	1,250	2,210	1,751	2,063
WHITENH18P	70.16	87.70	45.22	79.28	63.70	75.27	69.99	86.94	44.85	79.04	63.14	75.85
BLACKNH18	3,790	261	1,471	498	965	595	3,807	274	1,464	514	972	583
BLACKNH18P	27.28	9.39	52.05	17.79	35.03	21.77	27.40	9.72	52.53	18.38	35.05	21.43
AIANNH18	79	23	21	13	9	13	75	29	19	10	11	6
AIANNH18P	0.57	0.83	0.74	0.46	0.33	0.48	0.54	1.03	0.68	0.36	0.40	0.22
ASIANNH18	35	8	13	4	6	4	34	12	8	2	9	3
ASIANNH18P	0.25	0.29	0.46	0.14	0.22	0.15	0.24	0.43	0.29	0.07	0.32	0.11
HPINH18	3	2	0	0	0	1	3	2	0	0	0	1
HPINH18P	0.02	0.07	0.00	0.00	0.00	0.04	0.02	0.07	0.00	0.00	0.00	0.04
OTHERNH18	4	1	1	1	0	1	7	0	3	0	1	3
OTHERNH18P	0.03	0.04	0.04	0.04	0.00	0.04	0.05	0.00	0.11	0.00	0.04	0.11
MLTMNH18	20	0	8	1	4	7	26	8	4	3	5	6
MLTMNH18P	0.14	0.00	0.28	0.04	0.15	0.26	0.19	0.28	0.14	0.11	0.18	0.22

Source: Data from Run A of the *TDA*, U.S. Bureau of the Census, Washington, D.C.

Selected observations for Table 11:

1: Tate Schools has WHITENHP = 68.22% and BLACKNHP = 28.63% for the 2010 Census; and WHITENHP = 68.03% and BLACKNHP = 28.75% for the *TDA* run. Similar results for 18^+ population.

2: School District 02 is the only district with a WHITENHP (also WHITENH18P) percentage lower than 50.00% in both the 2010 Census and the *TDA* run.

Table 12. Tylertown (Walthall County), MS Run A of Twenty-five Runs of the *TDA* for Districts 01, 02, 03, and 04

($C_{TDA}(g)$ counts result from 2020 Census Redistricting Data Production Settings ($\epsilon = 17.14$ for persons) version of *TDA*.)

$$\text{2010 Census IDEAL POPULATION} = \frac{1,609}{4} = 402.25 \qquad \textit{TDA} \text{ IDEAL POPULATION} = \frac{1,614}{4} = 403.50$$

	2010 Census, SF1 (*PL 94-171*) Counts & Percentages POST-2010 Plan					Counts & Percentages Run A of the *TDA*				
Demographics										
DIST-ID	Tylertown	01	02	03	04	Tylertown	01	02	03	04
TOTAL	1,609	405	399	391	414	1,614	405	417	396	396
DEV		2.8	-3.2	-11.2	11.8		1.5	13.5	-7.5	-7.5
DEVP		0.68	-0.81	-2.88	2.84		0.37	3.24	-1.89	-1.89
TOTAL18	1,233	327	320	313	273	1,226	327	327	301	271
TOTALHISP	42	12	7	9	14	42	5	11	18	8
TOTALHISPP	2.61	2.96	1.75	2.30	3.38	2.60	1.23	2.64	4.55	2.02
TOTALNH	1,567	393	392	382	400	1,572	400	406	378	388
TOTALNHP	97.39	97.04	98.25	97.70	96.62	97.40	98.77	97.36	95.45	97.98
WHITENH	860	371	215	246	28	850	370	217	236	27
WHITENHP	53.45	91.60	53.88	62.92	6.76	52.66	91.36	52.04	59.60	6.82
BLACKNH	679	17	174	119	369	682	17	183	128	354
BLACKNHP	42.20	4.20	43.61	30.43	89.13	42.26	4.20	43.88	32.32	89.39
AIANNH	14	5	3	3	3	14	4	2	4	4
AIANNHP	0.87	1.23	0.75	0.77	0.72	0.87	0.99	0.48	1.01	1.01
ASIANNH	12	0	0	12	0	17	6	2	7	2
ASIANNHP	0.75	0.00	0.00	3.07	0.00	1.05	1.48	0.48	1.77	0.51
HPINH	0	0	0	0	0	1	0	1	0	0
HPINHP	0.00	0.00	0.00	0.00	0.00	0.06	0.00	0.24	0.00	0.00
OTHERNH	0	0	0	0	0	4	2	0	1	1
OTHERNHP	0.00	0.00	0.00	0.00	0.00	0.25	0.49	0.00	0.25	0.25
MLTMNNH	2	0	0	2	0	4	1	1	2	0
MLTMNNHP	0.12	0.00	0.00	0.51	0.00	0.25	0.25	0.24	0.51	0.00
HISP18	27	7	4	8	8	25	2	5	12	6
HISP18P	2.19	2.14	1.25	2.56	2.93	2.04	0.61	1.53	3.99	2.21
NONHISP18	1,206	320	316	305	265	1,201	325	322	289	265
NONHISP18P	97.81	97.86	98.75	97.44	97.07	97.96	99.39	98.47	96.01	97.79
WHITENH18	723	302	188	210	23	713	303	191	196	23
WHITENH18P	58.64	92.35	58.75	67.09	8.42	58.16	92.66	58.41	65.12	8.49
BLACKNH18	462	14	127	81	240	467	10	130	90	237
BLACKNH18P	37.47	4.28	39.69	25.88	87.91	38.09	3.06	39.76	29.90	87.45
AIANNH18	10	4	1	3	2	10	3	1	2	4
AIANNH18P	0.81	1.22	0.31	0.96	0.73	0.82	0.92	0.31	0.66	1.48
ASIANNH18	10	0	0	10	0	7	6	0	1	0
ASIANNH18P	0.81	0.00	0.00	3.19	0.00	0.57	1.83	0.00	0.33	0.00
HPINH18	0	0	0	0	0	0	0	0	0	0
HPINH18P	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
OTHERNH18	0	0	0	0	0	3	2	0	0	1
OTHERNH18P	0.00	0.00	0.00	0.00	0.00	0.24	0.61	0.00	0.00	0.37
MLTMNH18	1	0	0	1	0	1	1	0	0	0
MLTMNH18P	0.08	0.00	0.00	0.32	0.00	0.08	0.31	0.00	0.00	0.00

Source: Data from Run A of the *TDA*, U.S. Bureau of the Census, Washington, D.C.

Selected observations for Table 12:

1: Tylertown has WHITENHP = 53.45% and BLACKNHP = 42.20% for the 2010 Census; and WHITENHP = 52.66% and BLACKNHP = 42.26% for the *TDA* run. Similar results hold for 18^{+} population.

2: District 04 has a BLACKNHP (also BLACKNH18P) percentage $\geq$ 75% in both the 2010 Census and the *TDA* run.

3: District 02 has WHITENHP = 53.88% in the 2010 Census and WHITENHP = 52.04% for the *TDA* run.

II.7. VARIATION DUE TO THE TopDown ALGORITHM

Definitions of Redistricting Measures of Variation. The measures defined here are all for a specific ϵ. Henceforth, and to simplify notation, we use S for SWA and T for TDA. Let

$$G \equiv \text{the number of demographic groups;}$$
$$C_S(g) \equiv \text{the population of group } g \text{ (2010 Census, SF1), for } g = 1, ..., G\text{; and}$$
$$C_{Ti}(g) \equiv \text{the population of group } g \text{ resulting from the } i^{th} \text{ } TDA \text{ run, for } i = 1, ..., 25.$$

We have the following measures including two types of variation among the 25 TDA runs within group g: one relative to $\bar{C}_T(g)$ (see below) and another relative to $C_S(g)$.

(i) The **average population of group** g over the 25 TDA runs is

$$\bar{C}_T(g) \equiv \frac{C_{T1}(g) + C_{T2}(g) + \cdots + C_{T,25}(g)}{25}.$$

(ii) The **variation(1) among the population of group** g over the 25 TDA runs is

$$V(1)_g \equiv \frac{[C_{T1}(g) - \bar{C}_T(g)]^2 + [C_{T2}(g) - \bar{C}_T(g)]^2 + \cdots + [C_{T,25}(g) - \bar{C}_T(g)]^2}{25}.$$

(iii) The **relative variation(1) among the population of group** g over the 25 TDA runs is

$$RV(1)_g \equiv \frac{\sqrt{V(1)_g}}{\bar{C}_T(g)}.$$

(iv) The **average relative variation(1) among the population over the** G **groups** (essentially a coefficient of variation) is

$$AVERV(1) \equiv \frac{RV(1)_1 + RV(1)_2 + \cdots + RV(1)_G}{G}.$$

(v) Denote the **median relative variation(1) among the population over the** G **groups** by $MEDRV(1)$.

(vi) The **variation(2) among the population of group** g over the 25 TDA runs is

$$V(2)_g \equiv \frac{[C_{T1}(g) - C_S(g)]^2 + [C_{T2}(g) - C_S(g)]^2 + \cdots + [C_{T,25}(g) - C_S(g)]^2}{25}.$$

(vii) The **relative variation(2) among the population of group** g over the 25 TDA runs is

$$RV(2)_g \equiv \frac{\sqrt{V(2)_g}}{C_S(g)}.$$

($viii$) The **average relative variation(2) among the population over the** G **groups** is

$$AVERV(2) \equiv \frac{RV(2)_1 + RV(2)_2 + \cdots + RV(2)_G}{G}.$$

(ix) Denote the **median relative variation(2) among the population over the** G **groups** by $MEDRV(2)$.

$V(1)_g$ is an empirical variance measuring variation among the 25 TDA runs for group g; and $V(2)_g$ is an empirical mean square error measuring variation and any potential bias (i.e., $(bias)^2$) relative to $C_S(g)$ for the 25 TDA runs for group g.

Tables 7V, 8V, 9V, 10V, 11V, and 12V are companion tables for Tables 7, 8, 9, 10, 11, and 12 respectively. The formats among the Tables 7V, 8V, 9V, 10V, 11V, and 12V are the same, so we make a few comments about Table 7V which also hold for the others. For each demographic group g in each district (Rhode Island and CD for Table 7V; SLDU for Table 8V; SLDL for Table 9V; etc.), we provide two sets of three quantities. The first set of quantities gives the average count ($\bar{C}_T(g)$) over the 25 *TDA* runs and two associated measures of variation ($\sqrt{V(1)_g}$ and $RV(1)_g$) relative to $\bar{C}_T(g)$, while the second set of quantities gives the 2010 Census (swapping) count ($C_S(g)$) and two associated measures of variation ($\sqrt{V(2)_g}$ and $RV(2)_g$) relative to $C_S(g)$. It is worth noting that $\sqrt{V(2)_g}$ and $RV(2)_g$ are not measures of variability in the swapped data. It is also worth noting that the unit is "persons" for each of the quantities $\bar{C}_T(g)$, $\sqrt{V(1)_g}$, $C_S(g)$, and $\sqrt{V(2)_g}$, while the quantities $RV(1)_g$ and $RV(2)_g$ are unitless. So for example, we consider the demographic group g = ASIANNH of CD-01 in Table 7V. We observe: $\bar{C}_T(g) = 17,680$ persons; $\sqrt{V(1)_g} = 20$ persons; and $RV(1)_g = 0.001$. We also observe: $C_S(g) = 17,705$ persons; $\sqrt{V(2)_g} = 32$ persons; and $RV(2)_g = 0.002$. The detailed computations for these quantities are illustrated in APPENDIX B. In the tables to follow, a few presented results are rounded. In such cases, especially when there is division, one may not be able to obtain other related presented results exactly.

Selected observations for Table 7V:

1: $RV(1)_g$ and $RV(2)_g$ are largest for the groups g = HPINH and HPINH18 which have the smallest counts. In general, groups with smaller counts tend to have more relative variation.

2: For a given group g, there is a tendency for $RV(2)_g \geq RV(1)_g$. While this may not be surprising given the definitions of the two measures of variation, this inequality need not hold in all cases, as standardized measures of variation insert different measures of total in the denominator.

3: We observe that $RV(1)_g$ and $RV(2)_g$ for counts of groups in CD-02 tend to be larger than for corresponding groups in CD-01. This may be because the districts formed in 2013 resulted in fewer members of minority groups being included in CD-02 than in the corresponding groups in CD-01.

Notice that the computations for $AVERV(1)$ and $AVERV(2)$ each only average over the relative variations for the counts in a column. Similarly, $MEDRV(1)$ and $MEDRV(2)$ are each the median over the relative variations for the counts in a column.

The Key Empirical Message on Variability

The two measures $AVERV(\cdot)$ and $MEDRV(\cdot)$ summarize the key single empirical message of this study ($\epsilon = 17.14$):

> *Relative variability in the TDA increases as we consider smaller pieces of geography and population - from state (RI POP = 1,052,567); to Congressional district (RI-CD IDEAL POP = 526,283.5); to upper chamber district (RI-SLDU IDEAL POP = 27,699.1); to lower chamber district (RI-SLDL IDEAL POP = 14,034.2); to Panola County, MS (DISTRICT IDEAL POP = 6,941.4); to Tate County, MS (SCHOOL DISTRICT IDEAL POP = 3,764.6); and finally to Tylertown (Walthall County), MS (DISTRICT IDEAL POP = 402.25).*

To see this empirical evidence, sequentially observe the values for $AVERV(\cdot)$ and $MEDRV(\cdot)$ on the last two rows of Tables 7V; 8V; 9V; 10V; 11V; 12V. We highlight some of this using *DISTRICT IDEAL POPULATION* and $AVERV(1)$ in Figure 1.

Figure 1

Jurisdiction	District	IDEAL POPULATION	*AVERV*(1)
Rhode Island	CD-01	526,283.50	0.006
Rhode Island	CD-02	526,283.50	0.008
Rhode Island	SLDU-01	27,699.10	0.036
Rhode Island	SLDU-02	27,699.10	0.051
Rhode Island	SLDU-03	27,699.10	0.045
Rhode Island	SLDU-04	27,699.10	0.048
Rhode Island	SLDL-01	14,034.20	0.067
Rhode Island	SLDL-02	14,034.20	0.069
Rhode Island	SLDL-03	14,034.20	0.057
Rhode Island	SLDL-04	14,034.20	0.104
Panola County, MS	D-01	6,941.40	0.362
Panola County, MS	D-02	6,941.40	0.353
Panola County, MS	D-03	6,941.40	0.344
Panola County, MS	D-04	6,941.40	0.369
Panola County, MS	D-05	6,941.40	0.280
Tate County Schools, MS	D-01	3,764.60	0.335
Tate County Schools, MS	D-02	3,764.60	0.355
Tate County Schools, MS	D-03	3,764.60	0.493
Tate County Schools, MS	D-04	3,764.60	0.449
Tate County Schools, MS	D-05	3,764.60	0.376
Tylertown, MS	D-01	402.25	0.748
Tylertown, MS	D-02	402.25	0.622
Tylertown, MS	D-03	402.25	0.596
Tylertown, MS	D-04	402.25	1.265

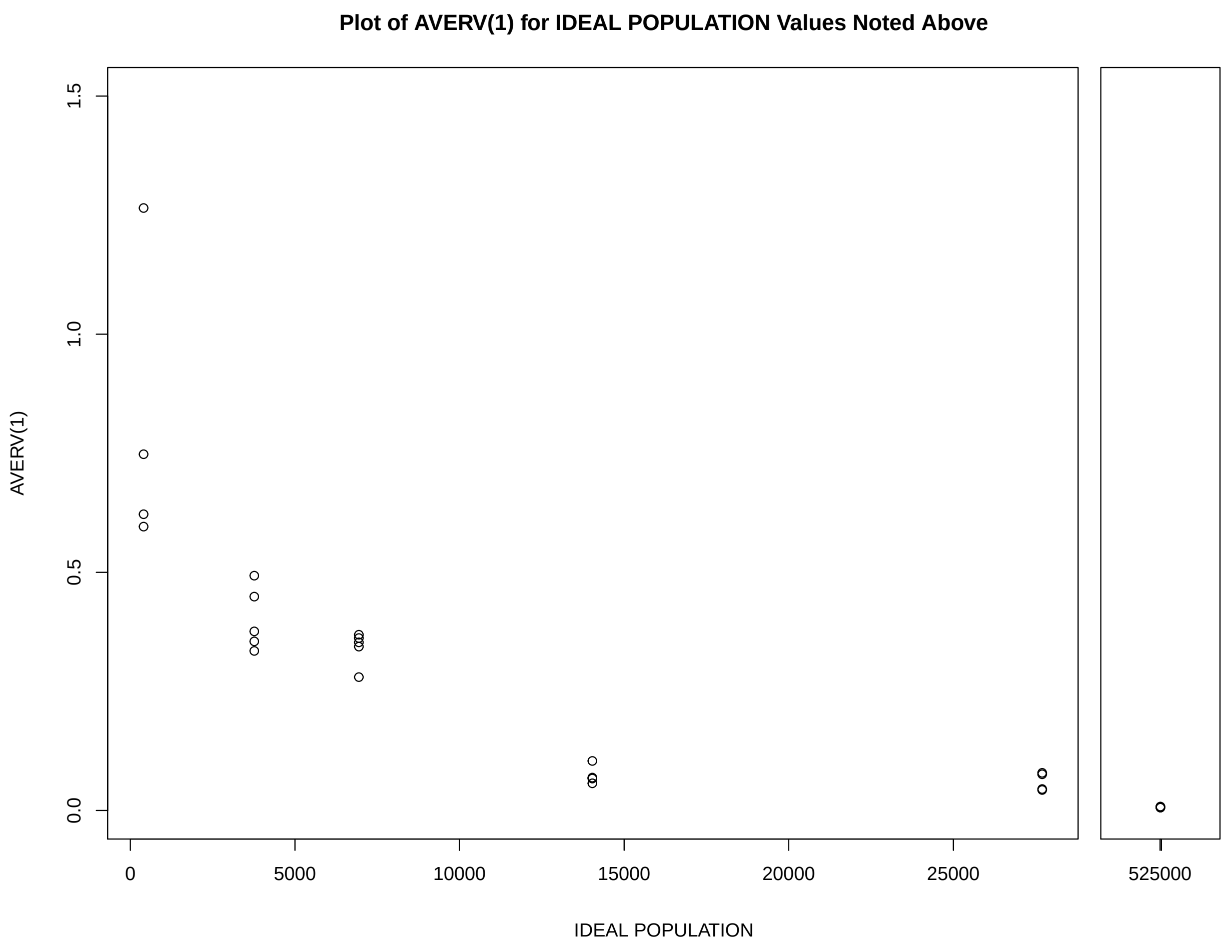


Table 7V. Counts & Measures of Variation for Rhode Island Twenty-five Runs of the *TDA* for Congressional Districts (CD) 01, and 02 (2013)
($C_T(g)$ counts result from 2020 Census Redistricting Data Production Settings ($\epsilon = 17.14$ for persons) version of *TDA*.)

	(Counts & Measures of Variation) (2013)					
DIST-ID	Rhode Island	Rhode Island	CD-01	CD-01	CD-02	CD-02
	$\bar{C}_T(g)$	$C_S(g)$	$\bar{C}_T(g)$	$C_S(g)$	$\bar{C}_T(g)$	$C_S(g)$
	$\sqrt{V(1)_g}$	$\sqrt{V(2)_g}$	$\sqrt{V(1)_g}$	$\sqrt{V(2)_g}$	$\sqrt{V(1)_g}$	$\sqrt{V(2)_g}$
Demographic (g)	$RV(1)_g$	$RV(2)_g$	$RV(1)_g$	$RV(2)_g$	$RV(1)_g$	$RV(2)_g$
TOTAL	1,052,567	1,052,567	526,149	526,283	526,418	526,284
	0	0	43	141	43	141
	0.000	0.000	0.000	0.000	0.000	0.000
TOTAL18	828,610	828,611	412,690	412,778	415,920	415,833
	15	15	85	122	80	119
	0.000	0.000	0.000	0.000	0.000	0.000
TOTALHISP	130,660	130,655	76,098	76,100	54,562	54,555
	13	14	51	51	57	58
	0.000	0.000	0.001	0.001	0.001	0.001
TOTALNH	921,907	921,912	450,051	450,183	471,856	471,729
	13	14	39	137	48	136
	0.000	0.000	0.000	0.000	0.000	0.000
WHITENH	803,685	803,685	377,045	377,109	426,640	426,576
	8	8	22	68	20	68
	0.000	0.000	0.000	0.000	0.000	0.000
BLACKNH	57,929	57,927	37,641	37,627	20,288	20,300
	10	10	18	23	19	22
	0.000	0.000	0.000	0.001	0.001	0.001
AIANNH	6,848	6,839	3,113	3,142	3,734	3,697
	9	12	19	35	18	41
	0.001	0.002	0.006	0.011	0.005	0.011
ASIANNH	34,193	34,194	17,680	17,705	16,513	16,489
	9	9	20	32	18	30
	0.000	0.000	0.001	0.002	0.001	0.002
HPINH	654	655	383	383	271	272
	11	11	13	13	11	11
	0.016	0.016	0.033	0.033	0.041	0.042
OTHERNH	10,292	10,296	8,503	8,492	1,789	1,804
	12	13	19	22	18	23
	0.001	0.001	0.002	0.003	0.010	0.013
MLTMNNH	8,306	8,316	5,686	5,725	2,620	2,591
	18	21	30	49	29	41
	0.002	0.003	0.005	0.009	0.011	0.016
HISP18	84,717	84,715	49,283	49,303	35,434	35,412
	14	14	52	56	51	55
	0.000	0.000	0.001	0.001	0.001	0.002
NONHISP18	743,894	743,896	363,407	363,475	380,486	380,421
	12	12	69	96	65	92
	0.000	0.000	0.000	0.000	0.000	0.000
WHITENH18	660,821	660,823	312,208	312,240	348,616	348,583
	8	8	29	43	28	44
	0.000	0.000	0.000	0.000	0.000	0.000
BLACKNH18	39,485	39,485	25,403	25,402	14,081	14,083
	9	9	27	27	27	27
	0.000	0.000	0.001	0.001	0.002	0.002
AIANNH18	4,965	4,963	2,313	2,332	2,652	2,631
	9	9	23	30	24	32
	0.002	0.002	0.010	0.013	0.009	0.012
ASIANNH18	25,330	25,333	13,257	13,276	12,073	12,057
	6	7	22	29	22	27
	0.000	0.000	0.002	0.002	0.002	0.002
HPINH18	499	500	303	307	196	193
	8	8	12	12	8	9
	0.017	0.017	0.039	0.040	0.043	0.047
OTHERNH18	7,285	7,290	6,079	6,061	1,207	1,229
	11	12	19	26	18	28
	0.001	0.002	0.003	0.004	0.015	0.023
MLTMNH18	5,505	5,502	3,844	3,857	1,661	1,645
	15	15	29	32	29	33
	0.003	0.003	0.008	0.008	0.018	0.020
$AVERV(\cdot)$	0.002	0.002	0.006	0.006	0.008	0.010
$MEDRV(\cdot)$	0.000	0.000	0.001	0.001	0.002	0.002

Source: Data from 25 Runs of the *TDA*, U. S. Bureau of the Census, Washington, D.C.

Table 8V. Counts & Measures of Variation for Rhode Island Twenty-five Runs of the *TDA* for State Upper Chamber Districts (SLDU) 01, 02, 03, and 04 (4 of 38 Districts, 2013)
($C_T(g)$ counts result from 2020 Census Redistricting Data Production Settings ($\epsilon = 17.14$ for persons) version of *TDA*.)

	(Measures of Variation) (2013)							
DIST-ID	SLDU-01	SLDU-01	SLDU-02	SLDU-02	SLDU-03	SLDU-03	SLDU-04	SLDU-04
	$\bar{C}_T(g)$	$C_S(g)$	$\bar{C}_T(g)$	$C_S(g)$	$\bar{C}_T(g)$	$C_S(g)$	$\bar{C}_T(g)$	$C_S(g)$
	$\sqrt{V(1)_g}$	$\sqrt{V(2)_g}$	$\sqrt{V(1)_g}$	$\sqrt{V(2)_g}$	$\sqrt{V(1)_g}$	$\sqrt{V(2)_g}$	$\sqrt{V(1)_g}$	$\sqrt{V(2)_g}$
Demographic (g)	$RV(1)_g$	$RV(2)_g$	$RV(1)_g$	$RV(2)_g$	$RV(1)_g$	$RV(2)_g$	$RV(1)_g$	$RV(2)_g$
TOTAL	28,109	28,161	28,147	28,079	28,600	28,398	28,228	28,201
	35	62	26	73	26	204	47	54
	0.001	0.002	0.001	0.003	0.001	0.007	0.002	0.002
TOTAL18	20,906	20,914	19,892	19,846	25,431	25,361	23,600	23,599
	29	30	31	55	31	76	45	45
	0.001	0.001	0.002	0.003	0.001	0.003	0.002	0.002
TOTALHISP	10,229	10,282	16,341	16,288	1,494	1,409	3,172	3,217
	30	61	28	60	18	87	27	52
	0.003	0.006	0.002	0.004	0.012	0.062	0.008	0.016
TOTALNH	17,881	17,879	11,806	11,791	27,106	26,989	25,055	24,984
	34	34	23	28	28	120	35	79
	0.002	0.002	0.002	0.002	0.001	0.004	0.001	0.003
WHITENH	10,212	10,222	3,517	3,553	22,023	22,028	21,260	21,210
	21	23	14	39	15	15	30	58
	0.002	0.002	0.004	0.011	0.001	0.001	0.001	0.003
BLACKNH	4,886	4,862	4,339	4,332	1,153	1,124	2,346	2,348
	28	37	13	15	15	33	17	18
	0.006	0.008	0.003	0.003	0.013	0.029	0.007	0.007
AIANNH	266	283	209	216	148	135	195	172
	10	20	9	12	6	15	10	25
	0.039	0.072	0.044	0.054	0.042	0.108	0.052	0.148
ASIANNH	1,554	1,526	3,055	3,032	3,269	3,262	805	826
	14	31	12	26	8	10	16	27
	0.009	0.021	0.004	0.008	0.002	0.003	0.020	0.032
HPINH	25	25	14	11	15	16	13	14
	5	5	4	5	4	4	3	3
	0.208	0.210	0.282	0.475	0.234	0.228	0.231	0.229
OTHERNH	456	457	200	189	253	224	240	241
	10	10	7	13	9	30	10	10
	0.022	0.022	0.033	0.068	0.035	0.136	0.041	0.041
MLTMNNH	482	504	473	458	244	200	197	173
	13	26	17	22	15	47	12	27
	0.028	0.052	0.035	0.049	0.062	0.235	0.060	0.156
HISP18	6,431	6,458	11,043	11,014	1,249	1,241	2,068	2,097
	27	38	31	42	17	19	24	38
	0.004	0.006	0.003	0.004	0.014	0.015	0.012	0.018
NONHISP18	14,474	14,456	8,849	8,832	24,182	24,120	21,532	21,502
	30	35	19	26	29	68	30	43
	0.002	0.002	0.002	0.003	0.001	0.003	0.001	0.002
WHITENH18	9,137	9,131	3,046	3,062	19,682	19,682	18,864	18,839
	16	17	11	19	13	13	25	35
	0.002	0.002	0.004	0.006	0.001	0.001	0.001	0.002
BLACKNH18	3,312	3,309	3,030	3,027	986	973	1,598	1,599
	25	25	14	15	13	18	18	18
	0.008	0.008	0.005	0.005	0.013	0.019	0.011	0.011
AIANNH18	187	197	150	154	117	110	146	136
	9	14	10	11	10	12	10	14
	0.050	0.069	0.068	0.071	0.085	0.111	0.067	0.101
ASIANNH18	1,183	1,170	2,149	2,135	3,001	2,989	600	611
	14	19	12	19	9	15	15	19
	0.012	0.016	0.006	0.009	0.003	0.005	0.025	0.030
HPINH18	20	20	11	11	13	14	10	13
	5	5	5	5	3	4	3	4
	0.237	0.237	0.430	0.445	0.262	0.254	0.297	0.313
OTHERNH18	327	326	134	125	198	186	175	178
	12	12	7	11	9	15	8	9
	0.038	0.038	0.053	0.090	0.045	0.080	0.047	0.050
MLTMNH18	307	303	329	318	185	166	139	126
	14	15	14	18	14	23	11	17
	0.045	0.048	0.043	0.055	0.076	0.140	0.079	0.136
$AVERV(\cdot)$	0.036	0.041	0.051	0.068	0.045	0.072	0.048	0.065
$MEDRV(\cdot)$	0.008	0.012	0.004	0.009	0.013	0.024	0.016	0.024

Source: Data from 25 Runs of the *TDA*, U. S. Bureau of the Census, Washington, D.C.

Table 9V. Counts & Measures of Variation for Rhode Island Twenty-five Runs of the *TDA* for State Lower Chamber Districts (SLDL) 01, 02, 03, and 04 (4 of 75 Districts, 2013) ($C_T(g)$ counts result from 2020 Census Redistricting Data Production Settings ($\epsilon = 17.14$ for persons) version of *TDA*.)

	(Measures of Variation) (2013)							
DIST-ID	SLDL-01	SLDL-01	SLDL-02	SLDL-02	SLDL-03	SLDL-03	SLDL-04	SLDL-04
	$\bar{C}_T(g)$	$C_S(g)$	$\bar{C}_T(g)$	$C_S(g)$	$\bar{C}_T(g)$	$C_S(g)$	$\bar{C}_T(g)$	$C_S(g)$
	$\sqrt{V(1)_g}$	$\sqrt{V(2)_g}$	$\sqrt{V(1)_g}$	$\sqrt{V(2)_g}$	$\sqrt{V(1)_g}$	$\sqrt{V(2)_g}$	$\sqrt{V(1)_g}$	$\sqrt{V(2)_g}$
Demographic (g)	$RV(1)_g$	$RV(2)_g$	$RV(1)_g$	$RV(2)_g$	$RV(1)_g$	$RV(2)_g$	$RV(1)_g$	$RV(2)_g$
TOTAL	14,098	13,881	13,721	13,821	13,776	13,949	13,594	13,713
	67	237	44	109	48	180	37	125
	0.005	0.016	0.003	0.008	0.003	0.013	0.003	0.009
TOTAL18	12,939	12,835	12,648	12,800	9,562	9,607	11,127	11,205
	54	118	36	156	34	56	32	84
	0.004	0.009	0.003	0.012	0.004	0.006	0.003	0.008
TOTALHISP	1,098	1,002	1,723	1,768	5,844	5,905	1,023	1,049
	31	100	27	53	35	70	11	28
	0.028	0.100	0.016	0.030	0.006	0.012	0.011	0.027
TOTALNH	13,001	12,879	11,999	12,053	7,932	8,044	12,571	12,664
	45	130	32	63	40	119	35	99
	0.003	0.010	0.003	0.005	0.005	0.015	0.003	0.008
WHITENH	9,916	9,922	8,693	8,714	3,452	3,465	9,546	9,539
	33	33	22	30	20	24	26	27
	0.003	0.003	0.002	0.003	0.006	0.007	0.003	0.003
BLACKNH	616	581	1,133	1,125	2,979	3,015	1,472	1,495
	13	37	19	20	22	42	9	25
	0.021	0.064	0.016	0.018	0.007	0.014	0.006	0.017
AIANNH	62	46	98	104	156	189	112	126
	6	17	7	9	10	34	5	15
	0.095	0.369	0.073	0.088	0.061	0.180	0.048	0.121
ASIANNH	2,186	2,175	1,767	1,776	802	794	785	792
	18	21	22	24	14	16	14	15
	0.008	0.010	0.012	0.013	0.017	0.020	0.017	0.019
HPINH	10	12	12	16	13	12	3	1
	3	3	4	6	4	4	3	3
	0.257	0.261	0.384	0.394	0.324	0.349	0.801	3.429
OTHERNH	82	57	136	148	248	257	382	396
	10	27	8	14	10	13	6	15
	0.121	0.477	0.061	0.098	0.038	0.051	0.016	0.039
MLTMNNH	128	86	159	170	281	312	271	315
	15	45	10	15	12	33	11	46
	0.113	0.518	0.065	0.088	0.042	0.106	0.012	0.145
HISP18	988	951	1,393	1,475	3,514	3,518	671	693
	27	16	24	86	27	27	17	28
	0.028	0.019	0.017	0.058	0.008	0.008	0.025	0.040
NONHISP18	11,951	11,884	11,256	11,325	6,048	6,089	10,456	10,512
	38	77	33	77	29	50	33	65
	0.003	0.006	0.003	0.007	0.005	0.008	0.003	0.006
WHITENH18	9,080	9,081	8,317	8,339	3,040	3,040	8,127	8,119
	25	25	23	32	17	17	19	21
	0.003	0.003	0.003	0.004	0.006	0.006	0.002	0.003
BLACKNH18	575	560	967	972	1,950	1,971	1,139	1,144
	12	19	15	16	19	28	10	11
	0.020	0.034	0.016	0.016	0.010	0.014	0.009	0.009
AIANNH18	53	45	77	82	110	129	93	101
	6	10	8	10	10	21	6	11
	0.104	0.217	0.109	0.120	0.087	0.163	0.070	0.105
ASIANNH18	2,063	2,052	1,646	1,655	582	575	631	635
	18	21	19	21	11	13	11	12
	0.009	0.010	0.011	0.012	0.019	0.023	0.017	0.018
HPINH18	9	10	9	14	10	11	3	1
	2	3	4	6	4	4	2	3
	0.260	0.251	0.462	0.454	0.362	0.346	0.906	2.742
OTHERNH18	65	51	112	126	183	190	268	280
	9	17	9	16	10	12	8	15
	0.132	0.327	0.076	0.127	0.053	0.063	0.031	0.053
MLTMNH18	105	85	126	137	172	173	196	232
	14	24	7	13	15	15	12	38
	0.130	0.284	0.054	0.092	0.086	0.085	0.059	0.164
$AVERV(\cdot)$	0.067	0.151	0.069	0.082	0.057	0.074	0.104	0.348
$MEDRV(\cdot)$	0.024	0.056	0.016	0.024	0.013	0.018	0.016	0.023

Source: Data from 25 Runs of the *TDA*, U. S. Bureau of the Census, Washington, D.C.

Table 10V. Counts & Measures of Variation for Panola County, MS Twenty-five Runs of the *TDA*
for County Districts 01, 02, 03, 04, 05
($C_T(g)$ counts result from 2020 Census Redistricting Data Production Settings ($\epsilon = 17.14$ for persons) version of *TDA*.)

	(Measures of Variation)											
DIST-ID	Panola County		01	01	02	02	03	03	04	04	05	05
	$\bar{C}_T(g)$	$C_S(g)$	$\bar{C}_T(g)$	$C_S(g)$	$\bar{C}_T(g)$	$C_S(g)$	$\bar{C}_T(g)$	$C_S(g)$	$\bar{C}_T(g)$	$C_S(g)$	$\bar{C}_T(g)$	$C_S(g)$
	$\sqrt{V(1)_g}$	$\sqrt{V(2)_g}$	$\sqrt{V(1)_g}$	$\sqrt{V(2)_g}$	$\sqrt{V(1)_g}$	$\sqrt{V(2)_g}$	$\sqrt{V(1)_g}$	$\sqrt{V(2)_g}$	$\sqrt{V(1)_g}$	$\sqrt{V(2)_g}$	$\sqrt{V(1)_g}$	$\sqrt{V(2)_g}$
Demographic (g)	$RV(1)_g$	$RV(2)_g$	$RV(1)_g$	$RV(2)_g$	$RV(1)_g$	$RV(2)_g$	$RV(1)_g$	$RV(2)_g$	$RV(1)_g$	$RV(2)_g$	$RV(1)_g$	$RV(2)_g$
TOTAL	34,708	34,707	6,996	6,974	6,523	6,549	7,062	7,074	7,114	7,105	7,012	7,005
	2	2	31	38	30	40	27	29	30	31	32	33
	0.000	0.000	0.004	0.005	0.005	0.006	0.004	0.004	0.004	0.004	0.005	0.005
TOTAL18	25,369	25,363	5,217	5,214	4,726	4,732	5,172	5,171	5,353	5,345	4,902	4,901
	10	12	24	24	30	31	26	26	24	25	27	27
	0.000	0.000	0.005	0.005	0.006	0.007	0.005	0.005	0.004	0.005	0.005	0.005
TOTALHISP	503	494	75	66	71	75	81	85	121	120	155	148
	9	13	8	12	7	8	6	7	8	8	8	10
	0.018	0.026	0.104	0.184	0.094	0.105	0.075	0.087	0.062	0.063	0.051	0.071
TOTALNH	34,205	34,213	6,921	6,908	6,452	6,474	6,982	6,989	6,993	6,985	6,857	6,857
	8	12	30	33	30	37	25	26	28	29	29	29
	0.000	0.000	0.004	0.005	0.005	0.006	0.004	0.004	0.004	0.004	0.004	0.004
WHITENH	16,986	16,981	2,431	2,419	2,098	2,096	4,027	4,030	5,252	5,250	3,177	3,186
	7	8	23	26	17	17	21	21	22	22	23	24
	0.000	0.000	0.010	0.011	0.008	0.008	0.005	0.005	0.004	0.004	0.007	0.008
BLACKNH	16,893	16,899	4,425	4,427	4,312	4,332	2,914	2,925	1,676	1,658	3,566	3,557
	7	9	19	19	26	32	16	19	16	24	15	17
	0.000	0.001	0.004	0.004	0.006	0.007	0.005	0.007	0.010	0.015	0.004	0.005
AIANNH	146	148	32	26	22	20	18	15	30	38	45	49
	8	8	5	7	4	5	4	5	8	11	6	7
	0.051	0.052	0.146	0.288	0.200	0.232	0.252	0.347	0.258	0.293	0.126	0.143
ASIANNH	86	89	7	8	6	7	8	5	12	17	54	52
	7	8	3	3	3	3	2	4	3	6	3	4
	0.083	0.086	0.439	0.405	0.463	0.420	0.305	0.704	0.262	0.336	0.060	0.069
HPINH	5	4	1	0	1	0	1	0	1	2	2	2
	4	5	1	1	1	2	2	2	2	2	1	2
	0.796	1.152	1.604	Inf	1.327	Inf	1.095	Inf	1.627	0.938	0.987	0.787
OTHERNH	15	19	4	7	3	5	2	1	4	3	3	3
	6	7	2	4	2	3	2	2	2	3	2	2
	0.397	0.380	0.584	0.538	0.604	0.540	0.952	1.637	0.681	0.854	0.702	0.611
MLTMNNH	73	73	21	21	11	14	12	13	18	17	12	8
	7	7	4	4	3	5	3	3	3	3	4	5
	0.094	0.095	0.201	0.200	0.290	0.330	0.275	0.266	0.176	0.196	0.301	0.668
HISP18	304	298	45	44	45	44	49	52	69	63	96	95
	10	11	7	7	6	6	6	7	7	9	7	7
	0.032	0.039	0.146	0.153	0.141	0.147	0.127	0.133	0.100	0.140	0.072	0.074
NONHISP18	25,065	25,065	5,171	5,170	4,681	4,688	5,123	5,119	5,284	5,282	4,806	4,806
	8	8	24	24	28	29	23	24	23	23	24	24
	0.000	0.000	0.005	0.005	0.006	0.006	0.005	0.005	0.004	0.004	0.005	0.005
WHITENH18	13,457	13,455	2,027	2,025	1,727	1,732	3,077	3,072	4,123	4,115	2,503	2,511
	6	6	21	21	19	19	16	17	16	17	19	21
	0.000	0.000	0.010	0.010	0.011	0.011	0.005	0.005	0.004	0.004	0.008	0.008
BLACKNH18	11,392	11,394	3,102	3,099	2,924	2,928	2,017	2,024	1,121	1,118	2,229	2,225
	6	7	15	15	21	21	17	19	15	15	12	13
	0.001	0.001	0.005	0.005	0.007	0.007	0.009	0.009	0.013	0.013	0.006	0.006
AIANNH18	113	115	23	21	18	16	14	11	23	29	35	38
	6	6	5	5	5	5	4	5	6	9	6	6
	0.052	0.055	0.207	0.248	0.273	0.318	0.278	0.441	0.257	0.298	0.161	0.165
ASIANNH18	53	54	5	8	4	5	5	2	9	12	31	27
	6	6	3	4	2	3	1	3	3	5	5	6
	0.113	0.112	0.540	0.490	0.586	0.506	0.302	1.597	0.342	0.376	0.152	0.219
HPINH18	4	2	0	0	1	0	1	0	1	1	1	1
	4	4	1	1	1	1	1	2	1	1	1	1
	0.998	1.992	1.827	Inf	1.344	Inf	1.449	Inf	2.206	1.311	1.179	1.428
OTHERNH18	4	5	1	1	1	0	1	1	1	2	1	1
	3	3	1	1	1	1	1	1	1	1	1	1
	0.617	0.556	1.173	1.039	1.297	Inf	1.389	0.959	0.975	0.714	1.232	1.183
MLTMNH18	42	40	13	16	7	7	8	9	8	5	6	3
	6	6	3	4	3	3	3	3	3	4	3	4
	0.139	0.160	0.227	0.272	0.378	0.410	0.346	0.330	0.381	0.840	0.529	1.442
$AVERV(\cdot)$	0.170	0.235	0.362	Inf	0.353	Inf	0.344	Inf	0.369	0.321	0.280	0.345
$MEDRV(\cdot)$	0.042	0.045	0.146	0.192	0.171	0.190	0.189	0.199	0.138	0.168	0.066	0.072

Source: Data from 25 Runs of the *TDA*, U. S. Bureau of the Census, Washington, D.C.

Table 11V. Counts & Measures of Variation for Tate County School Districts, MS Twenty-five Runs of the *TDA* for County Districts 01, 02, 03, 04, 05
($C_T(g)$ counts result from 2020 Census Redistricting Data Production Settings ($\epsilon = 17.14$ for persons) version of *TDA*.)

	(Measures of Variation)											
DIST-ID	Tate Schools		01	01	02	02	03	03	04	04	05	05
	$\bar{C}_T(g)$	$C_S(g)$	$\bar{C}_T(g)$	$C_S(g)$	$\bar{C}_T(g)$	$C_S(g)$	$\bar{C}_T(g)$	$C_S(g)$	$\bar{C}_T(g)$	$C_S(g)$	$\bar{C}_T(g)$	$C_S(g)$
	$\sqrt{V(1)_g}$	$\sqrt{V(2)_g}$	$\sqrt{V(1)_g}$	$\sqrt{V(2)_g}$	$\sqrt{V(1)_g}$	$\sqrt{V(2)_g}$	$\sqrt{V(1)_g}$	$\sqrt{V(2)_g}$	$\sqrt{V(1)_g}$	$\sqrt{V(2)_g}$	$\sqrt{V(1)_g}$	$\sqrt{V(2)_g}$
Demographic (g)	$RV(1)_g$	$RV(2)_g$	$RV(1)_g$	$RV(2)_g$	$RV(1)_g$	$RV(2)_g$	$RV(1)_g$	$RV(2)_g$	$RV(1)_g$	$RV(2)_g$	$RV(1)_g$	$RV(2)_g$
TOTAL	18,815	18,823	3,916	3,914	3,885	3,893	3,644	3,665	3,714	3,697	3,657	3,654
	18	20	22	22	21	23	20	30	26	31	16	16
	0.001	0.001	0.006	0.006	0.005	0.006	0.006	0.008	0.007	0.008	0.004	0.004
TOTAL18	13,892	13,893	2,776	2,780	2,833	2,826	2,789	2,799	2,766	2,755	2,728	2,733
	17	17	20	21	19	20	14	17	23	26	13	14
	0.001	0.001	0.007	0.007	0.007	0.007	0.005	0.006	0.008	0.009	0.005	0.005
TOTALHISP	423	399	95	87	64	63	106	110	51	32	106	107
	9	26	6	10	4	4	8	9	6	20	8	8
	0.021	0.064	0.066	0.118	0.063	0.066	0.073	0.078	0.119	0.631	0.072	0.071
TOTALNH	18,392	18,424	3,821	3,827	3,821	3,830	3,537	3,555	3,663	3,665	3,551	3,547
	18	37	22	23	21	23	19	26	24	24	18	18
	0.001	0.002	0.006	0.006	0.005	0.006	0.005	0.007	0.007	0.007	0.005	0.005
WHITENH	12,805	12,841	3,387	3,378	1,613	1,628	2,833	2,860	2,276	2,293	2,696	2,682
	13	39	14	17	14	21	14	30	20	26	12	19
	0.001	0.003	0.004	0.005	0.009	0.013	0.005	0.011	0.009	0.011	0.005	0.007
BLACKNH	5,394	5,389	373	400	2,158	2,139	678	666	1,363	1,349	822	835
	11	12	12	30	10	21	11	16	14	20	15	20
	0.002	0.002	0.033	0.074	0.004	0.010	0.016	0.024	0.010	0.015	0.018	0.024
AIANNH	101	103	35	32	23	26	15	19	10	11	17	15
	6	6	4	5	3	4	3	5	3	3	4	4
	0.056	0.059	0.110	0.160	0.142	0.167	0.181	0.265	0.294	0.284	0.207	0.274
ASIANNH	50	47	18	14	13	16	7	6	7	7	5	4
	4	5	4	6	3	4	2	2	2	2	2	3
	0.084	0.105	0.245	0.403	0.197	0.242	0.296	0.337	0.255	0.254	0.414	0.642
HPINH	3	3	1	2	1	0	0	0	0	0	1	1
	3	3	1	1	1	1	1	1	1	1	1	1
	0.764	0.841	0.952	0.714	1.894	Inf	2.114	Inf	1.710	Inf	1.362	0.980
OTHERNH	10	9	2	1	5	5	1	1	1	1	2	1
	5	5	2	2	2	2	1	1	1	1	2	2
	0.487	0.571	0.938	1.766	0.509	0.470	1.138	1.281	1.138	1.281	0.981	1.887
MLTMNNH	30	32	4	0	9	16	4	3	5	4	8	9
	7	7	2	5	3	8	2	2	3	3	3	3
	0.225	0.220	0.583	Inf	0.355	0.475	0.563	0.727	0.472	0.709	0.404	0.375
HISP18	227	215	48	47	39	34	59	63	24	16	57	55
	9	15	6	6	4	6	5	7	5	10	6	6
	0.041	0.070	0.117	0.123	0.113	0.191	0.091	0.110	0.197	0.603	0.105	0.112
NONHISP18	13,666	13,678	2,728	2,733	2,795	2,792	2,731	2,736	2,742	2,739	2,671	2,678
	14	19	20	21	18	19	12	13	22	22	15	17
	0.001	0.001	0.007	0.008	0.007	0.007	0.004	0.005	0.008	0.008	0.006	0.006
WHITENH18	9,735	9,747	2,444	2,438	1,274	1,278	2,204	2,219	1,747	1,755	2,066	2,057
	11	16	11	13	10	11	10	18	15	17	12	15
	0.001	0.002	0.005	0.005	0.008	0.009	0.004	0.008	0.009	0.010	0.006	0.007
BLACKNH18	3,793	3,790	244	261	1,482	1,471	506	498	979	965	582	595
	10	10	11	20	10	15	8	12	12	18	12	18
	0.003	0.003	0.045	0.078	0.006	0.010	0.016	0.024	0.013	0.019	0.021	0.030
AIANNH18	74	79	23	23	19	21	12	13	7	9	12	13
	5	7	5	5	3	4	3	3	3	3	3	3
	0.071	0.094	0.204	0.204	0.161	0.176	0.223	0.218	0.400	0.379	0.238	0.232
ASIANNH18	37	35	12	8	11	13	5	4	5	6	4	4
	4	4	4	5	3	4	2	2	2	2	2	2
	0.098	0.114	0.294	0.667	0.284	0.282	0.437	0.561	0.422	0.389	0.447	0.427
HPINH18	3	3	1	2	0	0	0	0	0	0	1	1
	3	3	1	2	1	1	1	1	1	1	1	1
	0.919	0.894	1.106	0.775	2.051	Inf	2.828	Inf	1.710	Inf	1.491	0.980
OTHERNH18	5	4	1	1	2	1	1	1	0	0	1	1
	3	3	1	1	2	2	1	1	1	1	1	1
	0.580	0.820	1.061	0.980	0.716	1.980	1.155	0.800	1.679	Inf	1.260	1.311
MLTMNH18	19	20	3	0	5	8	2	1	3	4	6	7
	6	6	3	4	3	4	2	2	2	2	3	3
	0.322	0.310	0.921	Inf	0.560	0.505	0.707	2.200	0.505	0.453	0.464	0.423
$AVERV(\cdot)$	0.184	0.209	0.335	Inf	0.355	Inf	0.493	Inf	0.449	Inf	0.376	0.390
$MEDRV(\cdot)$	0.049	0.067	0.113	0.141	0.128	0.171	0.136	0.164	0.226	0.332	0.156	0.172

Source: Data from 25 Runs of the *TDA*, U. S. Bureau of the Census, Washington, D.C.

Table 12V. Counts & Measures of Variation for Tylertown (Walthall County), MS Twenty-five Runs of the *TDA* for County Districts 01, 02, 03, 04
($C_T(g)$ counts result from 2020 Census Production Redistricting Data Settings ($\epsilon = 17.14$ for persons) version of *TDA*.)

	(Measures of Variation)									
DIST-ID	Tylertown		01	01	02	02	03	03	04	04
	$\bar{C}_T(g)$	$C_S(g)$	$\bar{C}_T(g)$	$C_S(g)$	$\bar{C}_T(g)$	$C_S(g)$	$\bar{C}_T(g)$	$C_S(g)$	$\bar{C}_T(g)$	$C_S(g)$
	$\sqrt{V(1)_g}$	$\sqrt{V(2)_g}$	$\sqrt{V(1)_g}$	$\sqrt{V(2)_g}$	$\sqrt{V(1)_g}$	$\sqrt{V(2)_g}$	$\sqrt{V(1)_g}$	$\sqrt{V(2)_g}$	$\sqrt{V(1)_g}$	$\sqrt{V(2)_g}$
Demographic (g)	$RV(1)_g$	$RV(2)_g$	$RV(1)_g$	$RV(2)_g$	$RV(1)_g$	$RV(2)_g$	$RV(1)_g$	$RV(2)_g$	$RV(1)_g$	$RV(2)_g$
TOTAL	1,611	1,609	402	405	396	399	417	391	397	414
	2	3	14	14	13	13	16	30	8	19
	0.001	0.002	0.035	0.036	0.033	0.033	0.037	0.077	0.021	0.047
TOTAL18	1,239	1,233	327	327	320	320	326	313	266	273
	6	9	10	10	12	12	15	20	10	12
	0.005	0.007	0.031	0.031	0.038	0.038	0.045	0.062	0.036	0.043
TOTALHISP	42	42	9	12	10	7	18	9	5	14
	5	5	3	4	3	4	5	10	3	10
	0.124	0.123	0.336	0.346	0.314	0.594	0.303	1.138	0.560	0.682
TOTALNH	1,570	1,567	393	393	386	392	399	382	392	400
	5	6	13	13	12	13	15	23	9	12
	0.003	0.004	0.034	0.034	0.031	0.034	0.039	0.060	0.022	0.030
WHITENH	848	860	360	371	205	215	249	246	33	28
	3	13	10	15	9	14	12	13	7	9
	0.003	0.015	0.028	0.039	0.045	0.063	0.050	0.052	0.216	0.307
BLACKNH	681	679	20	17	174	174	132	119	356	369
	3	4	6	7	7	7	7	15	8	15
	0.005	0.006	0.320	0.417	0.041	0.041	0.056	0.123	0.022	0.042
AIANNH	15	14	5	5	4	3	5	3	1	3
	4	4	3	3	2	2	2	3	1	2
	0.257	0.293	0.510	0.514	0.545	0.673	0.439	1.056	1.041	0.721
ASIANNH	15	12	5	0	1	0	8	12	1	0
	3	4	3	6	1	2	3	5	1	1
	0.201	0.348	0.543	Inf	1.025	Inf	0.314	0.383	1.344	Inf
HPINH	1	0	0	0	0	0	0	0	0	0
	1	1	0	0	0	1	1	1	0	0
	1.061	Inf	2.708	Inf	1.458	Inf	1.827	Inf	4.899	Inf
OTHERNH	2	0	1	0	1	0	1	0	0	0
	2	3	1	1	1	1	1	2	0	0
	0.731	Inf	1.550	Inf	1.033	Inf	1.406	Inf	2.291	Inf
MLTMNNH	7	2	2	0	2	0	3	2	1	0
	4	7	1	2	2	3	2	2	1	1
	0.534	3.367	0.695	Inf	1.016	Inf	0.699	1.140	1.356	Inf
HISP18	28	27	6	7	6	4	13	8	3	8
	5	5	2	3	3	3	5	8	2	5
	0.185	0.200	0.397	0.369	0.547	0.846	0.399	0.940	0.557	0.629
NONHISP18	1,210	1,206	321	320	314	316	313	305	263	265
	7	8	9	9	11	11	14	16	9	9
	0.006	0.007	0.029	0.029	0.035	0.036	0.044	0.051	0.035	0.036
WHITENH18	716	723	297	302	183	188	211	210	25	23
	4	8	8	10	8	10	11	11	7	7
	0.006	0.012	0.028	0.032	0.045	0.052	0.051	0.052	0.267	0.304
BLACKNH18	463	462	13	14	126	127	88	81	236	240
	4	5	4	4	6	6	6	9	6	8
	0.009	0.010	0.289	0.279	0.051	0.051	0.068	0.114	0.027	0.032
AIANNH18	11	10	4	4	3	1	4	3	1	2
	4	4	2	2	2	3	2	2	1	2
	0.365	0.439	0.517	0.522	0.734	2.683	0.658	0.792	1.169	0.781
ASIANNH18	13	10	4	0	1	0	7	10	0	0
	3	4	2	5	1	1	3	4	1	1
	0.259	0.430	0.547	Inf	1.541	Inf	0.390	0.387	1.705	Inf
HPINH18	1	0	0	0	0	0	0	0	0	0
	1	1	0	0	0	1	1	1	0	0
	1.228	Inf	3.391	Inf	1.604	Inf	1.895	Inf	4.899	Inf
OTHERNH18	1	0	0	0	0	0	1	0	0	0
	2	2	1	1	1	1	1	1	0	0
	1.195	Inf	2.134	Inf	1.333	Inf	2.067	Inf	2.708	Inf
MLTMNH18	5	1	1	0	1	0	2	1	0	0
	3	5	1	2	1	2	2	2	1	1
	0.580	4.712	0.842	Inf	0.976	Inf	1.126	2.236	2.121	Inf
$AVERV(\cdot)$	0.338	Inf	0.748	Inf	0.622	Inf	0.596	Inf	1.265	Inf
$MEDRV(\cdot)$	0.193	0.246	0.453	0.466	0.546	0.759	0.352	0.589	0.800	0.702

Source: Data from 25 Runs of the *TDA*, U. S. Bureau of the Census, Washington, D.C.

II.8. CONCLUDING REMARKS FOR PART II

For completeness, our first general remark is copied from our earlier report [5].

General Remark 1: Differential Privacy, *TDA*, and ϵ

Our objective in Part II of this study has been to report on the level of variability in results from the *TDA* and to reveal any effects on variability given advances with the *TDA* and an increased ϵ to 17.14. Our intent has not been to discuss how the *TDA* is constructed or how it operates. However, we feel compelled to offer a few such comments in this general remark, though our knowledge and understanding about the *TDA* is limited [1], [2].

The objective of the *TDA* is to bring privacy protection to respondent data. There are three things to consider: (*i*) a database (i.e., the 2010 CEF); (*ii*) a query made to the database (e.g., the number of people with certain characteristics in the database); and (*iii*) a randomized data protection mechanism that gives differential privacy (i.e., a probability distribution which is a part of the *TDA*). As Dwork (2014) [2] notes, "On an intuitive level, the goal of *differential privacy* is to obscure the presence or absence of any individual (in a database), or small group of individuals, while at the same time preserving statistical utility."

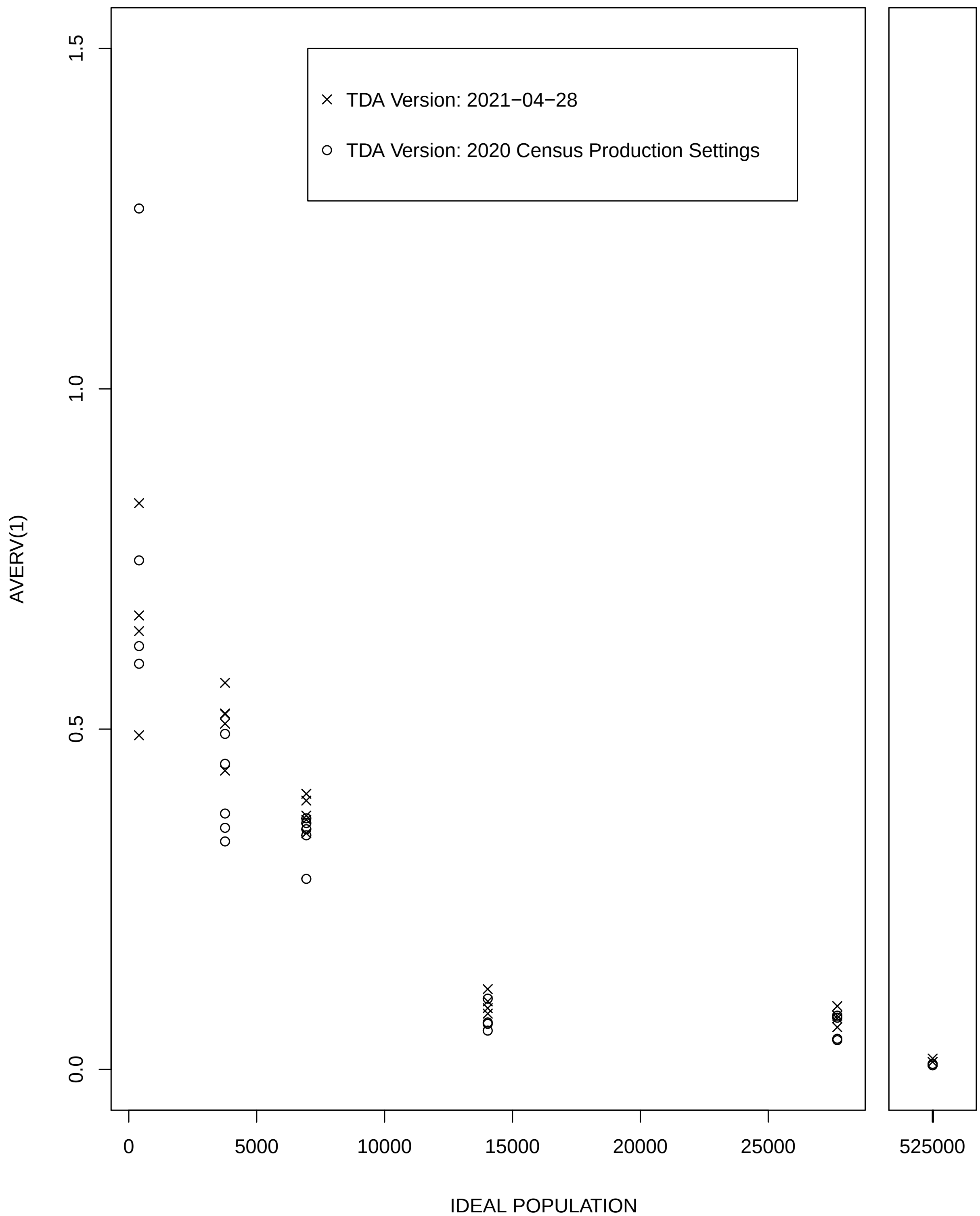
With differential privacy, the degree of privacy protection is reported by a positive quantity ϵ. Consider two different values of ϵ, ϵ_1 and ϵ_2. If $\epsilon_1 < \epsilon_2$, more privacy is offered with ϵ_1 than with ϵ_2. While details of the *TDA* and its foundation based on principles of *differential privacy* [4] are out-of-scope for this study (whose focus is only observing variability of output from the *TDA*), we note that the *TDA* has two components; and we share a little of our limited understanding. For simplicity, assume that an investigator is interested in knowing the count of persons in the 2010 CEF data with certain very specific characteristics. Thus a query is made of the 2010 CEF data (the answer sought should be a nonnegative integer). In the first component (noisy measurement) of the implementation of the *TDA*, random noise is generated and added to the answer from our query of the 2010 CEF data. The source of the random noise is a probability distribution (differentially private random mechanism) with positive probability at each of the integers ...-3, -2, -1, 0, 1, 2, 3,... Thus the "noisy answer" that is to be returned to the investigator submitting the query is

$$\text{"noisy answer"} = (\text{the query's answer using 2010 CEF data}) + (\text{random noise which is an integer}).$$

However, if the random noise is a negative integer whose absolute value is greater than the query's answer using the 2010 CEF data, then our noisy answer would be a negative noisy answer, which is not feasible. Thus, action is needed. This is the purpose of component two (post-processing) of the *TDA*, to ensure that our "final noisy answer" to the query is a nonnegative integer. So some more work is needed before the investigator eventually gets a "final noisy answer" to the original query.

Statistical theory permits deep explicit understanding of the variability caused by generation of the random noise in the first component. In particular, if $\epsilon_1 < \epsilon_2$, the variability in the noise addition with ϵ_1 is more than the variability in the noise with ϵ_2. The variability and uncertainty due to the activity of the second component is less well understood by us, and we believe it currently contributes more variability and uncertainty than the first component for some queries. We believe that the empirical variability reported in this study is an overall combination of variability and uncertainty from the two components.

Figure 2



General Remark 2: Effects on Variability Due to Advances with TDA and Higher ϵ.

We observed reductions in variability from the 2021-04-28 version of the TDA with $\epsilon = 10.3$ to the 2020 Census redistricting data production settings version of the TDA with $\epsilon = 17.14$. We see this visually by comparing Tables 7V, 8V, 9V, 10V, 11V, and 12V of this study with Tables 7V, 8V, 9V, 10V, 11V, and 12V of our earlier study [6]. At a high level, Figure 2 shows $AVERV(1)$

values for each of the districts as shown in Figure 1 using the 2021-04-28 version and the 2020 Census redistricting data production settings version of the *TDA*. The $AVERV(1)$ values for the 2020 Census redistricting data production settings version tend to be lower than for the 2021-04-28 version.

General Remark 3: Repeat of Some Earlier Specific Remarks [5, 6]

In this remark, we repeat two specific remarks (slightly edited) made in our earlier study [6]:

Need for Better Understanding of the *TDA:* The output of the version of the *TDA* studied in this paper infuses noise via differentially private mechanisms with a total privacy-loss budget of $\epsilon = 17.14$. It then post-processes those noisy estimates into fully consistent non-negative, integer-valued data with the same schema as was produced in 2010. The observation that $RV(2)_g > RV(1)_g$ (also $\sqrt{V(2)_g} > \sqrt{V(1)_g}$) in the majority of the variation tables may be a reflection of some phenomenon like a bias caused by post-processing. If there is something like bias, it is relative to the official (swapping) counts from the 2010 Census and not necessarily relative to the unknown true counts. A stronger understanding of the cumulative effects of the noise infusion and post-processing, as they affect jurisdictions with smaller populations, would be beneficial. This is a topic for further study.

Study Limitation: This study is limited in that new data (*TDA*) was retrofitted into existing redistricting plans developed using similar, but different data (2010 Census) treated by swapping. In practice, redistricting plans would be drawn using one set of data to satisfy desired parameters. In Congressional redistricting, for instance, DEV would not exceed 1 for any district, by design.

REFERENCES

[1] Abowd, J., Ashmead, R., Garfinkel, S., Kifer, D., Leclerc, P., Machanavajjhala, A., Moran, B., Sexton, W., and Zhuravlev, P. (October 2019 Draft) U.S. Bureau of the Census, Washington, D.C. < https://github.com/uscensusbureau/census2020-das-2010ddp/blob/master/doc/20191020_1843_Consistency_for_Large_Scale_Differentially_Private_Histograms.pdf >

[2] Dwork, C. (2014). "Differential Privacy: A Cryptographic Approach to Private Data Analysis," in *Privacy, Big data, and the Public Good*, (Editors: J. Lane, V. Stodden, S. Bender, and H. Nissenbaum), New York, NY: Cambridge University Press, 296-322.

[3] **Table P2** HISPANIC OR LATINO, AND NOT HISPANIC OR LATINO BY RACE, Universe: Total population, 2010 Census Redistricting Data (Public Law 94-171) Summary File *Also* **Table P4** HISPANIC OR LATINO, AND NOT HISPANIC OR LATINO BY RACE FOR THE POPULATION 18 YEARS AND OVER, Universe: Total population 18 years and over, 2010 Census Redistricting Data (Public Law 94-171) Summary File, American FactFinder, U. S. Bureau of the Census, Washington, D.C.

[4] **Table P9** HISPANIC OR LATINO, AND NOT HISPANIC OR LATINO BY RACE, Universe: Total population, 2010 Census Congressional District Summary File (113^{th} Congress) *Also* **Table P11** HISPANIC OR LATINO, AND NOT HISPANIC OR LATINO BY RACE FOR THE POPULATION 18 YEARS AND OVER, Universe: Population 18 years and over, 2010 Census Congressional District Summary File (113^{th} Congress), American FactFinder, U. S. Bureau of the Census, Washington, D.C.

[5] Wright, T. and Irimata, K. (2020). "Variability Assessment of Data Treated by the TopDown Algorithm for Redistricting," *Study Series (Statistics #2020-02)*, Center for Statistical Research and Methodology, U.S. Bureau of the Census, Washington, D.C.

[6] Wright, T. and Irimata, K. (2021). "Empirical Study of Two Aspects of The TopDown Algorithm Output for Redistricting: Reliability & Variability", *Study Series (Statistics #2021-01)*, Center for Statistical Research and Methodology, U.S. Bureau of the Census, Washington, D.C.

[7] Zayatz, L. (2007). "Disclosure Avoidance Practices and Research at the U.S. Census Bureau: An Update", *Journal of Official Statistics, Vol 21, No. 2*, 253-265.

[8] **Voting Rights Act of 1965**, 110^{th} Public law, 89^{th} Congress, *U.S. Statutes at Large, Vol 79*, starts p 437.

[9] **Thornburg v Gingles (1986)**, U.S. Supreme Court, *Vol 478, U.S. 30.*

APPENDIX A. Data Dictionary for Demographic Groups

DIST-ID:	Identification for geographical area: e.g., congressional or state legislative, county, or state
TOTAL:	Total population
DEV:	Deviation from Ideal = TOTAL - (IDEAL POPULATION)
DEVP:	Percent deviation from Ideal = [DEV/(IDEAL POPULATION)]× 100%
TOTAL18	All individuals 18 years of age or older
TOTALHISP:	All individuals of any race and who chose Hispanic
TOTALHISPP:	[TOTALHISP/TOTAL]×100%
TOTALNH:	All individuals of any race and who chose Not Hispanic
TOTALNHP:	[TOTALNH/TOTAL]×100%
WHITENH:	All individuals who chose White and Not Hispanic
WHITENHP:	[WHITENH/TOTAL]×100%
BLACKNH:	All individuals who chose Black either singly or in combination with White and chose Not Hispanic
BLACKNHP:	[BLACKNH/TOTAL]×100%
AIANNH:	All individuals who chose AIAN either singly or in combination with White and chose Not Hispanic
AIANNHP:	[AIANNH/TOTAL]×100%
ASIANNH:	All individuals who chose Asian either singly or in combination with White and chose Not Hispanic
ASIANNHP:	[ASIANNH/TOTAL]×100%
HPINH:	All individuals who chose Hawaiian or Other Pacific Islander either singly or in combination with White and chose Not Hispanic
HPINHP:	[HPINH/TOTAL]×100%
OTHERNH:	All individuals who chose Some other race either singly or in combination with White and chose Not Hispanic
OTHERNHP:	[OTHERNH/TOTAL]×100%
MLTMNNH:	All individuals who chose two or more minority groups and may or may not have chosen White but did not select Hispanic *(We believe this definition needs to be clarified. We believe that the counts for White "and" each of the 5 other race categories should be subtracted from the two or more raves counts to obtain the counts for MLTMNNH.)*
MLTMNNHP:	[MLTMNNH/TOTAL]×100%
HISP18:	All individuals 18 years of age or older of any race who chose Hispanic
HISP18P:	[HISP18/TOTAL18]×100%
NONHISP18:	All individuals 18 years of age or older of any race who chose Not Hispanic
NONHISP18P:	[NONHISP18/TOTAL18]×100%
WHITENH18:	All individuals 18 years of age or older who chose White and Not Hispanic
WHITENH18P:	[WHITENH18/TOTAL18]×100%
BLACKNH18:	All individuals 18 years of age or older who chose Black either singly or in combination with White and chose Not Hispanic
BLACKNH18P:	[BLACKNH18/TOTAL18]×100%
AIANNH18:	All individuals 18 years of age or older who chose AIAN either singly or in combination with White and chose Not Hispanic
AIANNH18P:	[AIANNH18/TOTAL18]×100%
ASIANNH18:	All individuals 18 years of age or older who chose Asian either singly or in combination with White and chose Not Hispanic
ASIANNH18P:	[ASIANNH18/TOTAL18]×100%
HPINH18:	All individuals 18 years of age or older who chose Hawaiian or Other Pacific Islander either singly or in combination with White and chose Not Hispanic
HPINH18P:	[HPINH18/TOTAL18]×100%
OTHERNH18:	All individuals 18 years of age or older who chose some other race either singly or in combination with White and chose Not Hispanic
OTHERNH18P:	[OTHERNH18/TOTAL18]×100%
MLTMNNH18:	All individuals 18 years of age or older who chose two or more minority races and chose Not Hispanic *(See note above for MLTMNNH.)*
MLTMNNH18P:	[MLTMNNH18/TOTAL18]×100%

APPENDIX B. Computation Illustration for Measures of Variation in Table 7V

For the demographic group $g = ASIANNH$ of CD-01 in Table 7V, we illustrate the computations for $\bar{C}_T(g)$, $\sqrt{V(1)_g}$, $RV(1)_g$, $C_S(g)$, $\sqrt{V(2)_g}$, and $RV(2)_g$ which are all defined in Section II.7 of this report. The same details follow for all other demographic groups as well as all entries in Tables 7V; 8V; 9V; 10V; 11V; and 12V. From the 2010 Census (swapping), Table 7 gives $C_S(g) = 17,705$ There are 25 TDA runs, and the details for the i^{th} run are given on row i of the table below for $i = 1, 2, ..., 25$.

Run i	$C_{Ti}(g)$	$(C_{Ti}(g) - \bar{C}_T(g))^2$	$(C_{Ti}(g) - C_S(g))^2$
1.	17,709	$(17,709 - 17,680.24)^2 = 827.14$	$(17,709 - 17,705)^2 = 16$
2.	17,680	$(17,680 - 17,680.24)^2 = 0.06$	$(17,680 - 17,705)^2 = 625$
3.	17,668	$(17,668 - 17,680.24)^2 = 149.82$	$(17,668 - 17,705)^2 = 1,369$
4.	17,678	$(17,678 - 17,680.24)^2 = 5.02$	$(17,678 - 17,705)^2 = 729$
5.	17,693	$(17,693 - 17,680.24)^2 = 162.82$	$(17,693 - 17,705)^2 = 144$
6.	17,671	$(17,671 - 17,680.24)^2 = 85.38$	$(17,671 - 17,705)^2 = 1,156$
7.	17,664	$(17,664 - 17,680.24)^2 = 263.74$	$(17,664 - 17,705)^2 = 1,681$
8.	17,662	$(17,662 - 17,680.24)^2 = 332.70$	$(17,662 - 17,705)^2 = 1,849$
9.	17,692	$(17,692 - 17,680.24)^2 = 138.30$	$(17,692 - 17,705)^2 = 169$
10.	17,690	$(17,690 - 17,680.24)^2 = 95.26$	$(17,690 - 17,705)^2 = 225$
11.	17,646	$(17,646 - 17,680.24)^2 = 1,172.38$	$(17,646 - 17,705)^2 = 3,481$
12.	17,707	$(17,707 - 17,680.24)^2 = 716.10$	$(17,707 - 17,705)^2 = 4$
13.	17,642	$(17,642 - 17,680.24)^2 = 1,462.30$	$(17,642 - 17,705)^2 = 3,969$
14.	17,666	$(17,666 - 17,680.24)^2 = 202.78$	$(17,666 - 17,705)^2 = 1,521$
15.	17,705	$(17,705 - 17,680.24)^2 = 613.06$	$(17,705 - 17,705)^2 = 0$
16.	17,681	$(17,681 - 17,680.24)^2 = 0.58$	$(17,681 - 17,705)^2 = 576$
17.	17,704	$(17,704 - 17,680.24)^2 = 564.54$	$(17,704 - 17,705)^2 = 1$
18.	17,676	$(17,676 - 17,680.24)^2 = 17.98$	$(17,676 - 17,705)^2 = 841$
19.	17,667	$(17,667 - 17,680.24)^2 = 175.30$	$(17,667 - 17,705)^2 = 1,444$
20.	17,690	$(17,690 - 17,680.24)^2 = 95.26$	$(17,690 - 17,705)^2 = 225$
21.	17,645	$(17,645 - 17,680.24)^2 = 1,241.86$	$(17,645 - 17,705)^2 = 3,600$
22.	17,689	$(17,689 - 17,680.24)^2 = 76.74$	$(17,689 - 17,705)^2 = 256$
23.	17,682	$(17,682 - 17,680.24)^2 = 3.10$	$(17,682 - 17,705)^2 = 529$
24.	17,722	$(17,722 - 17,680.24)^2 = 1,743.90$	$(17,722 - 17,705)^2 = 289$
25.	17,677	$(17,677 - 17,680.24)^2 = 10.50$	$(17,677 - 17,705)^2 = 784$
Totals	442,006	10,156.56	25,483

Thus we have (compare with corresponding entries of Table 7V):

$\bar{C}_T(g) = \dfrac{442,006}{25} = 17,680.24 \approx \mathbf{17,680}$	$C_S(g) = \mathbf{17,705}$
$\sqrt{V(1)_g} = \sqrt{\dfrac{10,156.56}{25}} = 20.16 \approx \mathbf{20}$	$\sqrt{V(2)_g} = \sqrt{\dfrac{25,483}{25}} = 31.93 \approx \mathbf{32}$
$RV(1)_g = \dfrac{\sqrt{V(1)_g}}{\bar{C}_T(g)} = 0.00114 \approx \mathbf{0.001}$	$RV(2)_g = \dfrac{\sqrt{V(2)_g}}{C_S(g)} = 0.00180 \approx \mathbf{0.002}$

APPENDIX C. Determination of C^*_{SWA} Using 18 and Over Characteristics

As an alternative to the results in Table 3, Table 3′ below reveals an empirical answer to our question where we use TOTAL18 demographic groups in place of TOTAL demographic groups. More specifically, we use TOTAL18, HISP18, WHITENH18, BLACKNH18, AIANNH18, ASIANNH18, and HPINH18 in place of TOTAL, HISP18, WHITENH, BLACKNH, AIANNH, ASIANNH, and HPINH, respectively.

Table 3′: Proportion of Block Groups in Each Stratum for Three Criteria

(Proportion computations result from 2020 Census Redistricting Data Production Settings ($\epsilon = 17.14$ for persons) version of *TDA*.)

Population: United States (50 States & DC)

		Reliable Characteristics Criteria		
Stratum for Block Groups Using C_{SWA} for TOTAL	Number of Block Groups	Criterion I LDG $DR_g \le 0.01$	Criterion II LDG $DR_g \le 0.03$	Criterion III LDG $DR_g \le 0.05$
$50 \le C_{SWA} \le 99$	128	0.1406	0.3984	0.5469
$100 \le C_{SWA} \le 149$	99	0.2020	0.4747	0.6768
$150 \le C_{SWA} \le 199$	124	0.2177	0.5081	0.7177
$200 \le C_{SWA} \le 249$	154	0.2792	0.5974	0.8052
$250 \le C_{SWA} \le 299$	209	0.2775	0.6316	0.8325
$300 \le C_{SWA} \le 349$	264	0.3485	0.7652	0.9205
$350 \le C_{SWA} \le 399$	407	0.3587	0.7838	0.9189
$400 \le C_{SWA} \le 449$	569	0.4130	0.7926	0.9279
$450 \le C_{SWA} \le 499$	915	0.3934	0.8372	0.9486
$500 \le C_{SWA} \le 549$	1,699	0.4473	0.8723	0.9670
$550 \le C_{SWA} \le 599$	3,238	0.4682	0.8799	0.9710
$600 \le C_{SWA} \le 649$	5,131	0.4835	0.8953	0.9745
$650 \le C_{SWA} \le 699$	6,683	0.4839	0.9051	0.9791
$700 \le C_{SWA} \le 749$	7,356	0.5103	0.9226	0.9844
$750 \le C_{SWA} \le 799$	8,170	0.5274	0.9263	0.9836
$800 \le C_{SWA} \le 849$	8,213	0.5418	0.9364	0.9864
$850 \le C_{SWA} \le 899$	8,441	0.5664	0.9494	0.9884
$900 \le C_{SWA} \le 949$	8,657	0.5734	0.9485	0.9886
$950 \le C_{SWA} \le 999$	8,723	0.5866	0.9555	0.9883
$1{,}000 \le C_{SWA} \le 1{,}049$	8,398	0.6162	0.9640	0.9887
$1{,}050 \le C_{SWA} \le 1{,}099$	8,345	0.6258	0.9685	0.9887
$1{,}100 \le C_{SWA} \le 1{,}149$	7,950	0.6360	0.9707	0.9896
$1{,}150 \le C_{SWA} \le 1{,}199$	7,860	0.6529	0.9726	0.9898
$1{,}200 \le C_{SWA} \le 1{,}249$	7,451	0.6642	0.9764	0.9906
$1{,}250 \le C_{SWA} \le 1{,}299$	7,124	0.6683	0.9746	0.9883
$1{,}300 \le C_{SWA} \le 1{,}349$	6,714	0.6845	0.9768	0.9885
$1{,}350 \le C_{SWA} \le 1{,}399$	6,507	0.6968	0.9825	0.9900
$1{,}400 \le C_{SWA} \le 1{,}449$	5,911	0.7195	0.9831	0.9888
$1{,}450 \le C_{SWA} \le 1{,}499$	5,617	0.7216	0.9838	0.9909
$1{,}500 \le C_{SWA} \le 1{,}549$	5,390	0.7343	0.9818	0.9866
$1{,}550 \le C_{SWA} \le 1{,}599$	4,856	0.7381	0.9811	0.9889
$1{,}600 \le C_{SWA} \le 1{,}649$	4,508	0.7540	0.9843	0.9891
$1{,}650 \le C_{SWA} \le 1{,}699$	4,325	0.7612	0.9817	0.9875
$1{,}700 \le C_{SWA} \le 1{,}749$	4,093	0.7794	0.9863	0.9895
$1{,}750 \le C_{SWA} \le 1{,}799$	3,689	0.7742	0.9883	0.9905
$1{,}800 \le C_{SWA} \le 1{,}849$	3,469	0.7737	0.9853	0.9888
$1{,}850 \le C_{SWA} \le 1{,}899$	3,252	0.7949	0.9865	0.9883
$1{,}900 \le C_{SWA} \le 1{,}949$	3,008	0.7896	0.9860	0.9880
$1{,}950 \le C_{SWA} \le 1{,}999$	2,832	0.8030	0.9887	0.9894
$2{,}000 \le C_{SWA} \le 2{,}049$	2,573	0.8232	0.9852	0.9872
$2{,}050 \le C_{SWA} \le 2{,}099$	2,356	0.8226	0.9881	0.9907
$2{,}100 \le C_{SWA} \le 2{,}149$	2,307	0.8279	0.9896	0.9918
$2{,}150 \le C_{SWA} \le 2{,}199$	2,033	0.8224	0.9803	0.9813
$2{,}200 \le C_{SWA} \le 2{,}249$	1,999	0.8369	0.9865	0.9870
$2{,}250 \le C_{SWA} \le 2{,}299$	1,892	0.8451	0.9884	0.9884
$2{,}300 \le C_{SWA} \le 2{,}349$	1,666	0.8487	0.9862	0.9862
$2{,}350 \le C_{SWA} \le 2{,}399$	1,622	0.8539	0.9846	0.9846
$2{,}400 \le C_{SWA} \le 2{,}449$	1,421	0.8656	0.9880	0.9887
$2{,}450 \le C_{SWA} \le 2{,}499$	1,350	0.8793	0.9926	0.9926
Total	199,698			

Using Criterion II and searching from top to bottom for the first stratum whose proportion is at least 0.9500: From Table 3′, take C^*_{SWA} to be between 950 and 999. For block groups whose TOTAL C_{SWA} count is at least 999, the difference of ratios between the C_{TDA} and C_{SWA} ratios for the LDG will tend to be less than or equal to 3% (using our data).

Using Criterion III and searching from top to bottom for the first stratum whose proportion is at least 0.9500: From Table 3′, take C^*_{SWA} to be between 500 and 549. For block groups whose TOTAL C_{SWA} count is at least 549, the difference of ratios between the C_{TDA} and C_{SWA} ratios for the LDG will tend to be less than or equal to 5% (using our data).

Using the data that will be released to the public (one run of the 2020 Census redistricting data production settings version of *TDA*), we might say, empirically based on the data for the block groups used in our study, that

> *"for any block group with a TOTAL count between 500 and 549 people, the difference between the TDA ratio of the largest demographic group (LDG) and the corresponding SWA ratio for the LDG among the 18 years and over pouplation is less than or equal to 5 percentage points at least 95% of the time".*

We applied the same version of the *TDA* to the same underlying CEF data 25 independent times, i.e., for 25 additional runs focusing on the 18 years and over population. The stratum (using C_{SWA} for TOTAL) for each run, where we first observed that 0.9500 was exceeded is given in Table 3a′ for each run is between 450 and 499 people in 23 of the 25 runs.

Table 3a′: For Each Run, the Stratum and Stratum Proportion When 0.9500 First Exceeded
(Proportion computations result from 2020 Census Redistricting Data Production Settings ($\epsilon = 17.14$ for persons) version of *TDA*.)
Population: United States (50 States & DC)

	Criterion III LDG $DR_g \leq 0.05$	
TDA Run	Stratum for Block Groups	Proportion When 0.9500 First Exceeded
1	$450 \leq C_{SWA} \leq 499$	0.9552
2	$450 \leq C_{SWA} \leq 499$	0.9617
3	$450 \leq C_{SWA} \leq 499$	0.9607
4	$450 \leq C_{SWA} \leq 499$	0.9650
5	$450 \leq C_{SWA} \leq 499$	0.9541
6	$450 \leq C_{SWA} \leq 499$	0.9519
7	$450 \leq C_{SWA} \leq 499$	0.9639
8	$450 \leq C_{SWA} \leq 499$	0.9574
9	$450 \leq C_{SWA} \leq 499$	0.9683
10	$400 \leq C_{SWA} \leq 449$	0.9525
11	$450 \leq C_{SWA} \leq 499$	0.9563
12	$450 \leq C_{SWA} \leq 499$	0.9639
13	$450 \leq C_{SWA} \leq 499$	0.9650
14	$450 \leq C_{SWA} \leq 499$	0.9661
15	$450 \leq C_{SWA} \leq 499$	0.9596
16	$450 \leq C_{SWA} \leq 499$	0.9519
17	$450 \leq C_{SWA} \leq 499$	0.9650
18	$400 \leq C_{SWA} \leq 449$	0.9543
19	$450 \leq C_{SWA} \leq 499$	0.9672
20	$450 \leq C_{SWA} \leq 499$	0.9552
21	$450 \leq C_{SWA} \leq 499$	0.9552
22	$450 \leq C_{SWA} \leq 499$	0.9661
23	$450 \leq C_{SWA} \leq 499$	0.9639
24	$450 \leq C_{SWA} \leq 499$	0.9585
25	$450 \leq C_{SWA} \leq 499$	0.9639