STUDY SERIES
(Statistics #2021-01)

Empirical Study of Two Aspects of the TopDown Algorithm Output for Redistricting: Reliability & Variability

Tommy Wright,
Kyle Irimata

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

Report Issued: May 28, 2021

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

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 empirical results for ongoing research and development. The *TopDown Algorithm (TDA)* [1] is being planned for use to protect the confidentiality of respondent data collected during the 2020 Census. Following the 2010 Census, swapping 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 2021-04-28 version ($\epsilon = 10.3$, 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 minimum TOTAL between 550 and 599 people in a block group provides reliable characteristics of various demographic groups in a block group based on the *TDA*. A similar minimum TOTAL between 350 and 399 is observed for places and MCDs. No Congressional or state legislative district failed our test for reliability.

Part II is an update of our results reported in [5] where $\epsilon = 4.0$; whereas, throughout this study $\epsilon = 10.3$. The objective here is to assess the variability of data results from application of the 2021-04-28 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. Most noticeable, variability with the 2021-04-28 version of the *TDA* ($\epsilon = 10.3$) is less than what we reported in [5] with the 2019-10-31 version where $\epsilon = 4.0$.

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 colleagues Mary Mulry, Pat Cantwell, Eric Slud, and James Whitehorne for their reading of a draft of this study and for their comments and questions that have strengthened the paper's presentation and content. 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

We assume that a version of the *TopDown Algorithm (TDA)* [1] will be applied to the 2020 Census Edited File (CEF) and that 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 2021-04-28 version of the *TDA* ($\epsilon = 10.3$) 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, and places, we conclude that:

> *"for any block group with a TOTAL count between 550 and 599 people, and for MCDs and places between 350 and 399, 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 [5] where $\epsilon = 4.0$; throughout this updated study, $\epsilon = 10.3$. [It should be noted that the overall $\epsilon = 12.2$ (10.3 for the person file and 1.9 for the housing file).] The objective of this part of our study is to assess the variability of data results from application of the 2021-04-28 version of the *TDA* to the 2010 Census Edited File (2010 CEF) for Rhode Island and for three additional jurisdictions. *Given more development of the TDA, a larger ϵ, and additional focus on how to allocate this ϵ, we see less variability throughout.*

Our approach in Part II has two components: (1) report variability among the 25 runs and (2) report variability of the 25 runs relative to the official published results from the 2010 Census (i.e., *Public Law 94-171*).

The first component of these analyses is a follow-up to earlier analyses done 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 these analyses, 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 where we observe *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 where we observe 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 ($\epsilon = 10.3$). As we reported in [5], 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 very high level, *Figure 2 shows less relative variability using the 2021-04-28 version of the TDA than the 2019-10-31 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 [5] assessed the variability of data results from application of the 2019-10-31 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*. To 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 2021-04-28 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 $\frac{C_{SWA}(g)}{C_{SWA}}$ and the *TDA* ratio $\frac{C_{TDA}(g)}{C_{TDA}}$, 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 [5]. For definition of each demographic group, see APPENDIX A. For the demographic group g = ASIANNH18, $C_{SWA}(g) = 142$ and $C_{TDA}(g) = 140$ with difference of ratios $DR_g = 0.0027$. That is, the difference between the two ratios for demographic group g is 0.27 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 2021-04-28 version of the *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	1,587	[c]
TOTAL18	1,198	1,209	[c]
TOTALHISP	133	139	$\lvert \frac{133}{1,560} - \frac{139}{1,587} \rvert = 0.0023$
TOTALNH	1,427	1,448	$\lvert \frac{1,427}{1,560} - \frac{1,448}{1,587} \rvert = 0.0023$
WHITENH	1,169	1,185	$\lvert \frac{1,169}{1,560} - \frac{1,185}{1,587} \rvert = 0.0027$
BLACKNH	36	61	$\lvert \frac{36}{1,560} - \frac{61}{1,587} \rvert = 0.0154$
AIANNH	10	9	$\lvert \frac{10}{1,560} - \frac{9}{1,587} \rvert = 0.0007$
ASIANNH	187	182	$\lvert \frac{187}{1,560} - \frac{182}{1,587} \rvert = 0.0052$
HPINH	5	1	$\lvert \frac{5}{1,560} - \frac{1}{1,587} \rvert = 0.0026$
OTHERNH	11	1	$\lvert \frac{11}{1,560} - \frac{1}{1,587} \rvert = 0.0064$
MLTMNNH	9	9	$\lvert \frac{9}{1,560} - \frac{9}{1,587} \rvert = 0.0001$
HISP18	93	92	$\lvert \frac{93}{1,198} - \frac{92}{1,209} \rvert = 0.0015$
NONHISP18	1,105	1,117	$\lvert \frac{1,105}{1,198} - \frac{1,117}{1,209} \rvert = 0.0015$
WHITENH18	914	919	$\lvert \frac{914}{1,198} - \frac{919}{1,209} \rvert = 0.0028$
BLACKNH18	29	42	$\lvert \frac{29}{1,198} - \frac{42}{1,209} \rvert = 0.0105$
AIANNH18	8	9	$\lvert \frac{8}{1,198} - \frac{9}{1,209} \rvert = 0.0008$
ASIANNH18	142	140	$\lvert \frac{142}{1,198} - \frac{140}{1,209} \rvert = 0.0027$
HPINH18	2	1	$\lvert \frac{2}{1,198} - \frac{1}{1,209} \rvert = 0.0008$
OTHERNH18	6	1	$\lvert \frac{6}{1,198} - \frac{1}{1,209} \rvert = 0.0042$
MLTMNNH18	4	5	$\lvert \frac{4}{1,198} - \frac{5}{1,209} \rvert = 0.0008$

[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. For those who want 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, $DR_g = 0.0023$; the difference between the two ratios is $0.0023 \times 100\% = 0.23$ 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 $0.0080 \times 100\% = 0.80$ percentage points.

Table 1b: Block Group 110010047012 (1,709 HUs) Characteristics
($C_{TDA}(g)$ counts result from 2021-04-28 version of the *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	2,902	[c]
TOTAL18	2,261	2,280	[c]
TOTALHISP	92	116	0.0080
TOTALNH	2,783	2,786	0.0080
WHITENH	541	529	0.0059
BLACKNH	1,686	1,697	0.0017
AIANNH	12	3	0.0031
ASIANNH	515	522	0.0007
HPINH	1	1	0.0000
OTHERNH	3	6	0.0010
MLTMNNH	25	28	0.0010
HISP18	86	100	0.0058
NONHISP18	2,175	2,180	0.0058
WHITENH18	529	519	0.0063
BLACKNH18	1,151	1,167	0.0028
AIANNH18	12	3	0.0040
ASIANNH18	460	465	0.0005
HPINH18	1	1	0.0000
OTHERNH18	3	6	0.0013
MLTMNNH18	19	19	0.0001

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: {0.0086, 0.0215, 0.0096}; {0.0015, 0.0194, 0.0131 }; {0.0033, 0.0001, 0.0041}; and {0.0007, 0.0003, 0.0020}. Assume the *TDA* count is considered a reliable characteristic for the largest demographic group if its $DR_g \leq 0.0050$. None 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 2021-04-28 version of the *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	**77**	[c]	500	**520**	[c]	1,000	**1,001**	[c]	1,500	**1,542**	[c]
TOTAL18	73	**75**	[c]	386	**407**	[c]	745	**743**	[c]	1,035	**1,058**	[c]
TOTALHISP	18	**11**[2]	**0.0767**	18	**37**[2]	**0.0352**	30	**32**[3]	**0.0020**	1,237	**1,274**[1]	**0.0015**
TOTALNH	64	66	0.0767	482	483	0.00352	970	969	0.0020	263	268	0.0015
WHITENH	60	**57**[1]	**0.0086**	455	**462**[1]	**0.0215**	306	**309**[2]	**0.0027**	235	**233**[2]	**0.0056**
BLACKNH	0	0	0.0000	7	**12**[3]	**0.0091**	659	**650**[1]	**0.0096**	10	11	0.0005
AIANNH	4	0	0.0488	6	6	0.0005	4	1	0.0030	0	3	0.0019
ASIANNH	0	**2**[3]	**0.0260**	11	2	0.0182	0	8	0.0080	18	**15**[3]	**0.0023**
HPINH	0	0	0.0000	0	0	0.0000	0	0	0.0000	0	2	0.0013
OTHERNH	0	0	0.0000	1	1	0.0000	0	0	0.0000	0	1	0.0006
MLTMNNH	0	7	0.0909	2	0	0.0040	1	1	0.0000	0	3	0.0019
HISP18	14	9	0.0718	10	22	0.0281	21	22	0.0014	807	821	0.0037
NONHISP18	59	66	0.0718	376	385	0.0281	724	721	0.0014	228	237	0.0037
WHITENH18	55	57	0.0066	354	369	0.0105	255	255	0.0000	203	205	0.0024
BLACKNH18	0	0	0.0000	6	7	0.0017	464	461	0.0024	9	10	0.0008
AIANNH18	4	0	0.0548	5	6	0.0018	4	1	0.0040	0	2	0.0019
ASIANNH18	0	2	0.0267	9	2	0.0184	0	4	0.0054	16	15	0.0013
HPINH18	0	0	0.0000	0	0	0.0000	0	0	0.0000	0	2	0.0019
OTHERNH18	0	0	0.0000	0	1	0.0025	0	0	0.0000	0	1	0.0009
MLTMNNH18	0	7	0.0933	2	0	0.0052	1	0	0.0013	0	2	0.0019

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

Table 2: Characteristics of Twelve Block Groups (continued)
($C_{TDA}(g)$ counts result from 2021-04-28 version of the *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	**1,966**	[c]	3,000	**2,939**	[c]	5,001	**5,016**	[c]	10,000	**10,014**	[c]
TOTAL18	1,562	**1,567**	[c]	2,153	**2,112**	[c]	3,689	**3,697**	[c]	6,704	**6,742**	[c]
TOTALHISP	349	**336**[2]	**0.0036**	224	**204**[2]	**0.0053**	1,770	**1,806**[2]	**0.0061**	1,291	**1,286**[3]	**0.0007**
TOTALNH	1,651	1,630	0.0036	2,776	2,735	0.0053	3,231	3,210	0.0061	8,709	8,728	0.0007
WHITENH	1,308	**1,324**[1]	**0.0194**	2,580	**2,566**[1]	**0.0131**	2,891	**2,883**[1]	**0.0033**	3,565	**3,571**[2]	**0.0001**
BLACKNH	181	**164**[3]	**0.0071**	87	**73**[3]	**0.0042**	235	**234**[3]	**0.0003**	4,475	**4,482**[1]	**0.0001**
AIANNH	25	28	0.0017	65	57	0.0023	18	26	0.0016	30	46	0.0016
ASIANNH	106	90	0.0072	32	28	0.0011	59	58	0.0002	473	487	0.0013
HPINH	10	11	0.0006	1	3	0.0007	8	0	0.0016	2	4	0.0002
OTHERNH	3	6	0.0016	4	6	0.0007	7	7	0.0000	79	76	0.0003
MLTMNNH	18	7	0.0054	7	2	0.0017	13	2	0.0022	85	62	0.0023
HISP18	236	233	0.0024	110	96	0.0056	1,193	1,219	0.0063	783	800	0.0019
NONHISP18	1,326	1,334	0.0024	2,043	2,016	0.0056	2,496	2,478	0.0063	5,921	5,942	0.0019
WHITENH18	1,089	1,101	0.0054	1,931	1,920	0.0122	2,267	2,257	0.0040	2,630	2,638	0.0010
BLACKNH18	129	129	0.0003	40	32	0.0034	149	147	0.0006	2,868	2,869	0.0023
AIANNH18	20	24	0.0025	41	40	0.0001	14	21	0.0019	22	34	0.0018
ASIANNH18	72	64	0.0053	23	16	0.0031	50	45	0.0014	304	316	0.0015
HPINH18	4	3	0.0006	1	3	0.0010	4	0	0.0011	2	4	0.0003
OTHERNH18	2	6	0.0025	3	5	0.0010	5	6	0.0003	43	37	0.0009
MLTMNNH18	10	7	0.0019	4	0	0.0019	7	2	0.0014	52	44	0.0012

Table 2: Characteristics of Twelve Block Groups (continued)
($C_{TDA}(g)$ counts result from 2021-04-28 version of the *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	**15,000**	[c]	19,506	**19,517**	[c]	29,677	**29,675**	[c]	37,452	**37,303**	[c]
TOTAL18	11,561	**11,545**	[c]	19,486	**19,454**	[c]	29,214	**29,198**	[c]	28,368	**28,284**	[c]
TOTALHISP	1,066	**1,026**[3]	**0.0022**	2,599	**2,581**[3]	**0.0010**	502	**501**[2]	**0.0000**	8,192	**8,091**[2]	**0.0018**
TOTALNH	14,023	13,974	0.0022	16,907	16,936	0.0010	29,175	29,174	0.0000	29,260	29,212	0.0018
WHITENH	7,901	**7,916**[1]	**0.0041**	10,579	**10,599**[1]	**0.0007**	28,555	**28,562**[1]	**0.0003**	23,326	**23,308**[1]	**0.0020**
BLACKNH	5,281	**5,273**[2]	**0.0015**	4,972	**4,975**[2]	**0.0000**	276	**275**[3]	**0.0000**	3,040	**3,040**[3]	**0.0003**
AIANNH	54	48	0.0004	275	286	0.0006	58	51	0.0002	601	610	0.0003
ASIANNH	643	629	0.0007	776	812	0.0018	246	238	0.0003	1,422	1,420	0.0001
HPINH	17	10	0.0005	80	75	0.0003	7	10	0.0001	340	346	0.0002
OTHERNH	42	32	0.0007	45	39	0.0003	15	10	0.0002	89	74	0.0004
MLTMNNH	85	66	0.0012	180	150	0.0015	18	28	0.0003	442	414	0.0007
HISP18	693	680	0.0010	2,597	2,567	0.0013	460	460	0.0000	5,506	5,449	0.0014
NONHISP18	10,868	10,865	0.0010	16,889	16,887	0.0013	28,754	28,738	0.0000	22,862	22,835	0.0014
WHITENH18	6,404	6,403	0.0007	10,562	10,572	0.0014	28,186	28,193	0.0008	18,751	18,741	0.0016
BLACKNH18	3,849	3,862	0.0016	4,971	4,971	0.0004	247	242	0.0002	2,118	2,107	0.0002
AIANNH18	46	46	0.0000	275	286	0.0006	58	51	0.0002	436	451	0.0006
ASIANNH18	494	486	0.0006	776	799	0.0012	227	213	0.0005	1,032	1,030	0.0000
HPINH18	9	10	0.0001	80	75	0.0003	7	8	0.0000	261	260	0.0000
OTHERNH18	22	19	0.0003	45	37	0.0004	14	10	0.0001	62	54	0.0003
MLTMNNH18	44	39	0.0004	180	147	0.0017	15	21	0.0002	202	192	0.0003

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.)

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

where the $C_{SWA(i)}$ counts are the counts for the TOTAL block group, 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 $DR_g \leq 0.01$), 0.4468 (or 44.68%) 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 0.4468

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 TOTAL-HISP, 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 3a 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
(Computations use $C_{TDA}(g)$ counts that result from 2021-04-28 version of the *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	0.1172	0.2812	0.4062
$100 \leq C_{SWA} \leq 149$	99	0.0909	0.3030	0.4646
$150 \leq C_{SWA} \leq 199$	124	0.1129	0.3710	0.5565
$200 \leq C_{SWA} \leq 249$	154	0.2143	0.4545	0.7143
$250 \leq C_{SWA} \leq 299$	209	0.2105	0.5167	0.7129
$300 \leq C_{SWA} \leq 349$	264	0.2121	0.5871	0.7803
$350 \leq C_{SWA} \leq 399$	407	0.2334	0.6757	0.8428
$400 \leq C_{SWA} \leq 449$	569	0.2900	0.7188	0.8963
$450 \leq C_{SWA} \leq 499$	915	0.3268	0.7628	0.9355
$500 \leq C_{SWA} \leq 549$	1,699	0.3431	0.7905	0.9370
$550 \leq C_{SWA} \leq 599$	3,238	0.3811	0.8295	0.9580
$600 \leq C_{SWA} \leq 649$	5,131	0.3962	0.8564	0.9723
$650 \leq C_{SWA} \leq 699$	6,683	0.4200	0.8692	0.9753
$700 \leq C_{SWA} \leq 749$	7,356	0.4468	0.8802	0.9826
$750 \leq C_{SWA} \leq 799$	8,170	0.4477	0.8973	0.9838
$800 \leq C_{SWA} \leq 849$	8,213	0.4785	0.9190	0.9907
$850 \leq C_{SWA} \leq 899$	8,441	0.4971	0.9231	0.9892
$900 \leq C_{SWA} \leq 949$	8,657	0.5021	0.9287	0.9928
$950 \leq C_{SWA} \leq 999$	8,723	0.5202	0.9411	0.9948
$1{,}000 \leq C_{SWA} \leq 1{,}049$	8,398	0.5460	0.9447	0.9936
$1{,}050 \leq C_{SWA} \leq 1{,}099$	8,345	0.5464	0.9575	0.9959
$1{,}100 \leq C_{SWA} \leq 1{,}149$	7,950	0.5552	0.9572	0.9969
$1{,}150 \leq C_{SWA} \leq 1{,}199$	7,860	0.5748	0.9626	0.9971
$1{,}200 \leq C_{SWA} \leq 1{,}249$	7,451	0.5770	0.9691	0.9977
$1{,}250 \leq C_{SWA} \leq 1{,}299$	7,124	0.6049	0.9698	0.9983
$1{,}300 \leq C_{SWA} \leq 1{,}349$	6,714	0.6151	0.9724	0.9993
$1{,}350 \leq C_{SWA} \leq 1{,}399$	6,507	0.6178	0.9743	0.9989
$1{,}400 \leq C_{SWA} \leq 1{,}449$	5,911	0.6287	0.9785	0.9980
$1{,}450 \leq C_{SWA} \leq 1{,}499$	5,617	0.6386	0.9810	0.9993
$1{,}500 \leq C_{SWA} \leq 1{,}549$	5,390	0.6471	0.9848	0.9996
$1{,}550 \leq C_{SWA} \leq 1{,}599$	4,856	0.6623	0.9841	0.9992
$1{,}600 \leq C_{SWA} \leq 1{,}649$	4,508	0.6528	0.9878	0.9998
$1{,}650 \leq C_{SWA} \leq 1{,}699$	4,325	0.6805	0.9864	0.9998
$1{,}700 \leq C_{SWA} \leq 1{,}749$	4,093	0.6895	0.9924	0.9993
$1{,}750 \leq C_{SWA} \leq 1{,}799$	3,689	0.6837	0.9883	0.9997
$1{,}800 \leq C_{SWA} \leq 1{,}849$	3,469	0.7094	0.9928	0.9997
$1{,}850 \leq C_{SWA} \leq 1{,}899$	3,252	0.7011	0.9889	1.0000
$1{,}900 \leq C_{SWA} \leq 1{,}949$	3,008	0.7048	0.9924	0.9997
$1{,}950 \leq C_{SWA} \leq 1{,}999$	2,832	0.7334	0.9926	0.9996
$2{,}000 \leq C_{SWA} \leq 2{,}049$	2,573	0.7178	0.9953	1.0000
$2{,}050 \leq C_{SWA} \leq 2{,}099$	2,356	0.7394	0.9949	1.0000
$2{,}100 \leq C_{SWA} \leq 2{,}149$	2,307	0.7391	0.9944	0.9991
$2{,}150 \leq C_{SWA} \leq 2{,}199$	2,033	0.7634	0.9970	1.0000
$2{,}200 \leq C_{SWA} \leq 2{,}249$	1,999	0.7564	0.9970	0.9995
$2{,}250 \leq C_{SWA} \leq 2{,}299$	1,892	0.7627	0.9963	1.0000
$2{,}300 \leq C_{SWA} \leq 2{,}349$	1,666	0.7533	0.9976	0.9994
$2{,}350 \leq C_{SWA} \leq 2{,}399$	1,622	0.7608	0.9957	1.0000
$2{,}400 \leq C_{SWA} \leq 2{,}449$	1,421	0.7643	0.9986	1.0000
$2{,}450 \leq C_{SWA} \leq 2{,}499$	1,350	0.7733	0.9970	0.9993
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 1,050 and 1,099. For block groups whose TOTAL C_{SWA} count is at least 1,099, 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 550 and 599. For block groups whose TOTAL C_{SWA} count is at least 599, 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 were released to the public (one run of the 2021-04-28 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 550 and 599 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".*

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

Table 4: For Each Run, the Stratum and Stratum Proportion When 0.9500 First Exceeded
(Proportion Computations use $C_{TDA}(g)$ counts that result from 2021-04-28 version of the *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	$550 \leq C_{SWA} \leq 599$	0.9589
2	$550 \leq C_{SWA} \leq 599$	0.9605
3	$550 \leq C_{SWA} \leq 599$	0.9623
4	$550 \leq C_{SWA} \leq 599$	0.9642
5	$550 \leq C_{SWA} \leq 599$	0.9608
6	$550 \leq C_{SWA} \leq 599$	0.9580
7	$550 \leq C_{SWA} \leq 599$	0.9592
8	$550 \leq C_{SWA} \leq 599$	0.9614
9	$550 \leq C_{SWA} \leq 599$	0.9595
10	$550 \leq C_{SWA} \leq 599$	0.9636
11	$550 \leq C_{SWA} \leq 599$	0.9592
12	$550 \leq C_{SWA} \leq 599$	0.9589
13	$550 \leq C_{SWA} \leq 599$	0.9592
14	$550 \leq C_{SWA} \leq 599$	0.9617
15	$550 \leq C_{SWA} \leq 599$	0.9589
16	$550 \leq C_{SWA} \leq 599$	0.9617
17	$550 \leq C_{SWA} \leq 599$	0.9617
18	$550 \leq C_{SWA} \leq 599$	0.9614
19	$550 \leq C_{SWA} \leq 599$	0.9592
20	$550 \leq C_{SWA} \leq 599$	0.9558
21	$550 \leq C_{SWA} \leq 599$	0.9592
22	$550 \leq C_{SWA} \leq 599$	0.9589
23	$550 \leq C_{SWA} \leq 599$	0.9580
24	$550 \leq C_{SWA} \leq 599$	0.9611
25	$550 \leq C_{SWA} \leq 599$	0.9568

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 MCDs and Places

As with the summary display in Table 4 for block groups, we present analagous intervals in Table 5a using results from the 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 for each run where we first observed that 0.9500 was exceeded is given in Table 5a for each run.

Table 5a: For Each Run, the Stratum and Stratum Proportion When 0.9500 First Exceeded
(Proportion Computations use $C_{TDA}(g)$ counts that result from 2021-04-28 version of the *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	$300 \leq C_{SWA} \leq 349$	0.9621
2	$250 \leq C_{SWA} \leq 299$	0.9580
3	$300 \leq C_{SWA} \leq 349$	0.9598
4	$250 \leq C_{SWA} \leq 299$	0.9580
5	$300 \leq C_{SWA} \leq 349$	0.9665
6	$300 \leq C_{SWA} \leq 349$	0.9688
7	$300 \leq C_{SWA} \leq 349$	0.9688
8	$300 \leq C_{SWA} \leq 349$	0.9621
9	$300 \leq C_{SWA} \leq 349$	0.9754
10	$300 \leq C_{SWA} \leq 349$	0.9576
11	$300 \leq C_{SWA} \leq 349$	0.9598
12	$300 \leq C_{SWA} \leq 349$	0.9777
13	$300 \leq C_{SWA} \leq 349$	0.9598
14	$300 \leq C_{SWA} \leq 349$	0.9688
15	$300 \leq C_{SWA} \leq 349$	0.9688
16	$300 \leq C_{SWA} \leq 349$	0.9643
17	$300 \leq C_{SWA} \leq 349$	0.9732
18	$300 \leq C_{SWA} \leq 349$	0.9665
19	$300 \leq C_{SWA} \leq 349$	0.9710
20	$300 \leq C_{SWA} \leq 349$	0.9621
21	$300 \leq C_{SWA} \leq 349$	0.9688
22	$350 \leq C_{SWA} \leq 399$	0.9520
23	$300 \leq C_{SWA} \leq 349$	0.9643
24	$300 \leq C_{SWA} \leq 349$	0.9598
25	$300 \leq C_{SWA} \leq 349$	0.9732

Using the data that were released to the public (one run of the 2021-04-28 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 300 and 349 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 & 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 4 for block groups and the summary display in Table 5b 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 for each run, where we first observed that 0.9500 was exceeded is given in Table 5b 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 were released to the public (one run of the 2021-04-28 version of *TDA*), we might say (as we did with block groups, also with MCDs and places) based on Table 5b, 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 at least 100% of the time".*

Table 5b: For Each Run, the Stratum and Stratum Proportion When 0.9500 First Exceeded
(Proportion Computations use $C_{TDA}(g)$ counts that result from 2021-04-28 version of the *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 5a, as well statements relating to Congressional and state legislative districts using Table 5b.

Remark 2: Much of our focus in Part I has been in the context of the total population chracteristics for block groups, MCDs and places, and Congressional and state legislative districts. In Table 3a 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 450 and 499 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.

Part II

VARIABILITY ASSESSMENT OF DATA TREATED BY THE TOPDOWN ALGORITHM

II.1. INTRODUCTION

Part II is actually an update of our results in [5] where $\epsilon = 4.0$ and the 2019-10-31 version of *TDA* was used; whereas, throughout this study, $\epsilon = 10.3$ and advances have been made resulting in the 2021-04-28 version of *TDA*. The careful reader will note that we reuse wording from [5] 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 [5], 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 ($\epsilon = 10.3$). 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 [5]) 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) $\times$ 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 $\epsilon = 10.3$, 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 $\frac{1,052,567}{38} = 27,699.1$. 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 ($\epsilon = 10.3$)

		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	526,449	526,118	526,173	526,394	525,872	526,695
DEV		-0.5	0.5	165.5	-165.5	-110.5	110.5	-411.5	411.5
DEVP		0.00	0.00	0.03	-0.03	-0.02	0.02	-0.08	0.08
TOTAL18	828,611	412,778	415,833	412,736	415,826	412,776	415,807	412,512	416,054
TOTALHISP	130,655	76,100	54,555	76,248	54,402	76,230	54,402	76,153	54,539
TOTALHISPP	12.41	14.46	10.37	14.48	10.34	14.49	10.33	14.48	10.35
TOTALNH	921,912	450,183	471,729	450.201	471,716	449,943	471,992	449,719	472,156
TOTALNHP	87.59	85.54	89.63	85.52	89.66	85.51	89.67	85.52	89.65
WHITENH	803,685	377,109	426,576	377,022	426,658	376,955	426,735	377,012	426,677
WHITENHP	76.35	71.66	81.05	71.62	81.10	71.64	81.07	71.69	81.01
BLACKNH	57,927	37,627	20,300	37,704	20,219	37,705	20,247	37,517	20,406
BLACKNHP	5.50	7.15	3.86	7.16	3.84	7.17	3.85	7.13	3.87
AIANNH	6,839	3,142	3,697	3,201	3,672	3,126	3,717	3,141	3,735
AIANNHP	0.65	0.60	0.70	0.61	0.70	0.59	0.71	0.60	0.71
ASIANNH	34,194	17,705	16,489	17,692	16,505	17,684	16,496	17,723	16,478
ASIANNHP	3.25	3.36	3.13	3.36	3.14	3.36	3.13	3.37	3.13
HPINH	655	383	272	427	242	400	263	355	293
HPINHP	0.06	0.07	0.05	0.08	0.05	0.08	0.05	0.07	0.06
OTHERNH	10,296	8,492	1,804	8,443	1,845	8,454	1,845	8,457	1,829
OTHERNHP	0.98	1.61	0.34	1.60	0.35	1.61	0.35	1.61	0.35
MLTMNNH	8,316	5,725	2,591	5,712	2,575	5,619	2,689	5,514	2,738
MLTMNNHP	0.79	1.09	0.49	1.09	0.49	1.07	0.51	1.05	0.52
HISP18	84,715	49,303	35,412	49,333	35,349	49,428	35,253	49,331	35,368
HISP18P	10.22	11.94	8.52	11.95	8.50	11.97	8.48	11.96	8.50
NONHISP18	743,896	363,475	380,421	363,403	380,477	363,348	380,554	363,181	380,686
NONHISP18P	89.78	88.06	91.48	88.05	91.50	88.03	91.52	88.04	91.50
WHITENH18	660,823	312,240	348,583	312,178	348,640	312,163	348,684	312,232	348,589
WHITENH18P	79.75	75.64	83.83	75.64	83.84	75.63	83.86	75.69	83.78
BLACKNH18	39,485	25,402	14,083	25,414	14,060	25,425	14,068	25,326	14,153
BLACKNH18P	4.77	6.15	3.39	6.16	3.38	6.16	3.38	6.14	3.40
AIANNH18	4,963	2,332	2,631	2,326	2,645	2,291	2,666	2,317	2,670
AIANNH18P	0.60	0.56	0.63	0.56	0.64	0.56	0.64	0.56	0.64
ASIANNH18	25,333	13,276	12,057	13,229	12,106	13,282	12,035	13,326	12,008
ASIANNH18P	3.06	3.22	2.90	3.21	2.91	3.22	2.89	3.23	2.89
HPINH18	500	307	193	334	175	313	195	275	221
HPINH18P	0.06	0.07	0.05	0.08	0.04	0.08	0.05	0.07	0.05
OTHERNH18	7,290	6,061	1,229	6,059	1,224	6,067	1,214	6,008	1,271
OTHERNH18P	0.88	1.47	0.30	1.47	0.29	1.47	0.29	1.46	0.31
MLTMNH18	5,502	3,857	1,645	3,863	1,627	3,807	1,692	3,697	1,774
MLTMNH18P	0.66	0.93	0.40	0.94	0.39	0.92	0.41	0.90	0.43

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)
($\epsilon = 10.3$)

	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	27,836	27,823	28,716	28,201
DEV	461.9	379.9	698.9	501.9	136.9	123.9	1,016.9	501.9
DEVP	1.64	1.35	2.46	1.78	0.49	0.45	3.54	1.78
TOTAL18	20,914	19,846	25,361	23,599	20,746	19,706	25,506	23,592
TOTALHISP	10,282	16,288	1,409	3,217	10,142	16,134	1,525	3,192
TOTALHISPP	36.51	58.01	4.96	11.41	36.43	57.99	5.31	11.32
TOTALNH	17,879	11,791	26,989	24,984	17,694	11,689	27,191	25,009
TOTALNHP	63.49	41.99	95.04	88.59	63.57	42.01	94.69	88.68
WHITENH	10,222	3,553	22,028	21,210	10,216	3,531	22,030	21,305
WHITENHP	36.30	12.65	77.57	75.21	36.70	12.69	76.72	75.55
BLACKNH	4,862	4,332	1,124	2,348	4,814	4,309	1,164	2,318
BLACKNHP	17.27	15.43	3.96	8.33	17.29	15.49	4.05	8.22
AIANNH	283	216	135	172	254	186	170	170
AIANNHP	1.00	0.77	0.48	0.61	0.91	0.67	0.59	0.60
ASIANNH	1,526	3,032	3,262	826	1,587	3,051	5,253	781
ASIANNHP	5.42	10.80	11.49	2.93	5.70	10.97	11.33	2.77
HPINH	25	11	16	14	18	6	27	9
HPINHP	0.09	0.04	0.06	0.05	0.06	0.02	0.09	0.03
OTHERNH	457	189	224	241	438	196	253	220
OTHERNHP	1.62	0.67	0.79	0.85	1.57	0.70	0.88	0.78
MLTMNNH	504	458	200	173	367	410	294	206
MLTMNNHP	1.79	1.63	0.70	0.61	1.32	1.47	1.02	0.73
HISP18	6,458	11,014	1,241	2,097	6,369	10,919	1,262	2,088
HISP18P	30.88	55.50	4.89	8.89	30.70	55.41	4.95	8.85
NONHISP18	14,456	8,832	24,120	21,502	14,377	8,787	24,244	21,504
NONHISP18P	69.12	44.50	95.11	91.11	69.30	44.59	95.05	91.15
WHITENH18	9,131	3,062	19,682	18,839	9,134	3,049	19,703	18,919
WHITENH18P	43.66	15.43	77.61	79.83	44.03	15.47	77.25	80.19
BLACKNH18	3,309	3,027	973	1,599	3,279	3,006	990	1,585
BLACKNH18P	15.82	15.25	3.84	6.78	15.81	15.25	3.88	6.72
AIANNH18	197	154	110	136	186	140	123	123
AIANNH18P	0.94	0.78	0.43	0.58	0.90	0.71	0.48	0.52
ASIANNH18	1,170	2,135	2,989	611	1,197	2,160	2,980	577
ASIANNH18P	5.59	10.76	11.79	2.59	5.77	10.96	11.68	2.45
HPINH18	20	11	14	13	11	5	21	5
HPINH18P	0.10	0.06	0.06	0.06	0.05	0.03	0.08	0.02
OTHERNH18	326	125	186	178	325	125	201	170
OTHERNH18P	1.56	0.63	0.73	0.75	1.57	0.63	0.79	0.72
MLTMNH18	303	318	166	126	245	302	226	125
MLTMNH18P	1.45	1.60	0.65	0.53	1.18	1.53	0.89	0.53

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 HIS18P 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)
($\epsilon = 10.3$)

	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	14,072	13,707	13,714	13,660
DEV	-153.2	-213.2	-85.2	-321.2	37.8	-327.2	-320.2	-374.2
DEVP	-1.10	-1.54	-0.61	-2.34	0.27	-2.39	-2.34	-2.74
TOTAL18	12,835	12,800	9,607	11,205	12,899	12,699	9,523	11,166
TOTALHISP	1,002	1,768	5,905	1,049	1,086	1,692	5,826	1,033
TOTALHISPP	7.22	12.79	42.33	7.65	7.72	12.34	42.48	7.56
TOTALNH	12,879	12,053	8,044	12,664	12,986	12,015	7,888	12,627
TOTALNHP	92.78	87.21	57.67	92.35	92.28	87.66	57.52	92.44
WHITENH	9,922	8,714	3,465	9,539	9,899	8,697	3,464	9,547
WHITENHP	71.48	63.05	24.84	69.56	70.35	63.45	25.26	69.89
BLACKNH	581	1,125	3,015	1,495	605	1,128	2,969	1,509
BLACKNHP	4.19	8.14	21.61	10.90	4.30	8.23	21.65	11.05
AIANNH	46	104	189	126	66	123	152	99
AIANNHP	0.33	0.75	1.35	0.92	0.47	0.90	1.11	0.72
ASIANNH	2,175	1,776	794	792	2,167	1,753	823	803
ASIANNHP	15.67	12.85	5.69	5.78	15.40	12.79	6.00	5.88
HPINH	12	16	12	1	25	11	6	9
HPINHP	0.09	0.12	0.09	0.01	0.18	0.08	0.04	0.07
OTHERNH	57	148	257	396	85	130	240	392
OTHERNHP	0.41	1.07	1.84	2.89	0.60	0.95	1.75	2.87
MLTMNNH	86	170	312	315	139	173	234	268
MLTMNNHP	0.62	1.23	2.24	2.30	0.99	1.26	1.71	1.96
HISP18	951	1,475	3,518	693	977	1,398	3,498	666
HISP18P	7.41	11.52	36.62	6.18	7.57	11.01	36.73	5.96
NONHISP18	11,884	11,325	6,089	10,512	11,922	11,301	6,025	10,500
NONHISP18P	92.59	88.48	63.38	93.82	92.43	88.99	63.27	94.04
WHITENH18	9,081	8,339	3,040	8,119	9,068	8,338	3,038	8,137
WHITENH18P	70.75	65.15	31.64	72.46	70.30	65.66	31.90	72.87
BLACKNH18	560	972	1,971	1,144	557	976	1,945	1,163
BLACKNH18P	4.36	7.59	20.52	10.21	4.32	7.69	20.42	10.42
AIANNH18	45	82	129	101	50	99	110	85
AIANNH18P	0.35	0.64	1.34	0.90	0.39	0.78	1.16	0.76
ASIANNH18	2,052	1,655	575	635	2,037	1,633	589	644
ASIANNH18P	15.99	12.93	5.99	5.67	15.79	12.86	6.19	5.77
HPINH18	10	14	11	1	22	8	2	3
HPINH18P	0.08	0.11	0.11	0.01	0.17	0.06	0.02	0.03
OTHERNH18	51	126	190	280	69	110	181	281
OTHERNH18P	0.40	0.98	1.98	2.50	0.53	0.87	1.90	2.52
MLTMNH18	85	137	173	232	119	137	160	187
MLTMNH18P	0.66	1.07	1.80	2.07	0.92	1.08	1.68	1.67

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
($\epsilon = 10.3$)

$$\text{2010 Census IDEAL POPULATION} = \frac{34,707}{5} = 6,941.4 \qquad \textit{TDA} \text{ IDEAL POPULATION} = \frac{34,702}{5} = 6,940.4$$

	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	34,702	7,044	6,571	7,033	7,066	6,988
DEV		32.6	-392.4	132.6	163.6	63.6		103.6	-369.4	92.6	125.6	47.6
DEVP		0.47	-5.99	1.87	2.30	0.91		1.47	-5.62	1.32	1.78	0.68
TOTAL18	25,363	5,214	4,732	5,171	5,345	4,901	25,384	5,267	4,730	5,171	5,313	4,903
TOTALHISP	494	66	75	85	120	148	521	98	80	80	104	159
TOTALHISPP	1.42	0.95	1.15	1.20	1.69	2.11	1.50	1.39	1.22	1.14	1.47	2.28
TOTALNH	34,213	6,908	6,474	6,989	6,985	6,857	34,181	6,946	6,491	6,953	6,962	6,829
TOTALNHP	98.58	99.05	98.85	98.80	98.31	97.89	98.50	98.61	98.78	98.86	98.53	97.72
WHITENH	16,981	2,419	2,096	4,030	5,250	3,186	16,989	2,455	2,084	4,020	5,249	3,181
WHITENHP	48.93	34.69	32.00	56.97	73.89	45.48	48.96	34.85	31.72	57.16	74.29	45.52
BLACKNH	16,899	4,427	4,332	2,925	1,658	3,557	16,870	4,421	4,345	2,893	1,660	3,551
BLACKNHP	48.69	63.48	66.15	41.35	23.34	50.78	48.61	62.76	66.12	41.13	23.49	50.82
AIANNH	148	26	20	15	38	49	143	28	24	21	34	36
AIANNHP	0.43	0.37	0.31	0.21	0.53	0.70	0.41	0.40	0.37	0.30	0.48	0.52
ASIANNH	89	8	7	5	17	52	100	14	20	8	9	49
ASIANNHP	0.26	0.11	0.11	0.07	0.24	0.74	0.29	0.20	0.30	0.11	0.13	0.70
HPINH	4	0	0	0	2	2	0	0	0	0	0	0
HPINHP	0.01	0.00	0.00	0.00	0.03	0.03	0.00	0.00	0.00	0.00	0.00	0.00
OTHERNH	19	7	5	1	3	3	4	2	2	0	0	0
OTHERNHP	0.05	0.10	0.08	0.01	0.04	0.04	0.01	0.03	0.03	0.00	0.00	0.00
MLTMNNH	73	21	14	13	17	8	75	26	16	11	10	12
MLTMNNHP	0.21	0.30	0.21	0.18	0.24	0.11	0.22	0.37	0.24	0.16	0.14	0.17
HISP18	298	44	44	52	63	95	320	71	57	43	61	88
HISP18P	1.17	0.84	0.93	1.01	1.18	1.94	1.26	1.35	1.21	0.83	1.15	1.79
NONHISP18	25,065	5,170	4,688	5,119	5,282	4,806	25,064	5,196	4,673	5,128	5,252	4,815
NONHISP18P	98.83	99.16	99.07	98.99	98.82	98.06	98.74	98.65	98.79	99.17	98.85	98.21
WHITENH18	13,455	2,025	1,732	3,072	4,115	2,511	13,464	2.044	1,697	3,097	4,112	2,514
WHITENH18P	53.05	38.84	36.60	59.41	76.99	51.23	53.04	38.81	35.88	59.89	77.40	51.27
BLACKNH18	11,394	3,099	2,928	2,024	1,118	2,225	11,386	3,110	2,937	2,004	1,107	2,228
BLACKNH18P	44.92	59.44	61.88	39.14	20.92	45.40	44.86	59.05	62.09	38.75	20.84	45.44
AIANNH18	115	21	16	11	29	38	116	22	18	17	23	36
AIANNH18P	0.45	0.40	0.34	0.21	0.54	0.78	0.46	0.42	0.38	0.33	0.43	0.73
ASIANNH18	54	8	5	2	12	27	60	7	13	4	4	32
ASIANNH18P	0.21	0.15	0.11	0.04	0.22	0.55	0.24	0.13	0.27	0.08	0.08	0.65
HPINH18	2	0	0	0	1	1	0	0	0	0	0	0
HPINH18P	0.01	0.00	0.00	0.00	0.02	0.02	0.00	0.00	0.00	0.00	0.00	0.00
OTHERNH18	5	1	0	1	2	1	0	0	0	0	0	0
OTHERNH18P	0.02	0.02	0.00	0.02	0.04	0.02	0.00	0.00	0.00	0.00	0.00	0.00
MLTMNH18	40	16	7	9	5	3	38	13	8	6	6	5
MLTMNH18P	0.16	0.31	0.15	0.17	0.09	0.06	0.15	0.25	0.17	0.12	0.11	0.10

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 = 48.96% and BLACKNHP = 48.61% 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 ($\epsilon = 10.3$)

$$\text{2010 Census IDEAL POPULATION} = \frac{18,823}{5} = 3,764.6 \qquad \textit{TDA} \text{ IDEAL POPULATION} = \frac{18,831}{5} = 3,766.2$$

	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	18,831	3,919	3,886	3,654	3,750	3,622
DEV		149.4	128.4	-99.6	-67.6	-110.6		152.8	119.8	-112.2	-16.2	-144.2
DEVP		3.82	3.30	-2.72	-1.83	-3.03		3.90	3.08	-3.07	-0.43	-3.98
TOTAL18	13,893	2,780	2,826	2,799	2,755	2,733	13,909	2,788	2,833	2,796	2,773	2,719
TOTALHISP	399	87	63	110	32	107	388	87	70	102	57	72
TOTALHISPP	2.12	2.22	1.62	3.00	0.87	2.93	2.06	2.22	1.80	2.79	1.52	1.99
TOTALNH	18,424	3,827	3,830	3,555	3,665	3,547	18,443	3,832	3,816	3,552	3,693	3,550
TOTALNHP	97.88	97.78	98.38	97.00	99.13	97.07	97.94	97.78	98.20	97.21	98.48	98.01
WHITENH	12,841	3,378	1,628	2,860	2,293	2,682	12,827	3,401	1,610	2,850	2,267	2,699
WHITENHP	68.22	86.31	41.82	78.04	62.02	73.40	68.12	86.78	41.43	78.00	60.45	74.52
BLACKNH	5,389	400	2,139	666	1,349	835	5,420	388	2,152	676	1,380	824
BLACKNHP	28.63	10.22	54.94	18.17	36.49	22.85	28.78	9.90	55.38	18.50	36.80	22.75
AIANNH	103	32	26	19	11	15	112	26	27	16	26	17
AIANNHP	0.55	0.82	0.67	0.52	0.30	0.41	0.59	0.66	0.69	0.44	0.69	0.47
ASIANNH	47	14	16	6	7	4	51	11	18	5	15	2
ASIANNHP	0.25	0.36	0.41	0.16	0.19	0.11	0.27	0.28	0.46	0.14	0.40	0.06
HPINH	3	2	0	0	0	1	0	0	0	0	0	0
HPINHP	0.02	0.05	0.00	0.00	0.00	0.03	0.00	0.00	0.00	0.00	0.00	0.00
OTHERNH	9	1	5	1	1	1	18	3	5	2	3	5
OTHERNHP	0.05	0.03	0.13	0.03	0.03	0.03	0.10	0.08	0.13	0.05	0.08	0.14
MLTMNNH	32	0	16	3	4	9	15	3	4	3	2	3
MLTMNNHP	0.17	0.00	0.41	0.08	0.11	0.25	0.08	0.08	0.10	0.08	0.05	0.08
HISP18	215	47	34	63	16	55	226	53	46	62	29	36
HISP18P	1.55	1.69	1.20	2.25	0.58	2.01	1.62	1.90	1.62	2.22	1.05	1.32
NONHISP18	13,678	2,733	2,792	2,736	2,739	2,678	13,683	2,735	2,787	2,734	2,744	2,683
NONHISP18P	98.45	98.31	98.80	97.75	99.42	97.99	98.38	98.10	98.38	97.78	98.95	98.68
WHITENH18	9,747	2,438	1,278	2,219	1,755	2,057	9,738	2,456	1,265	2,207	1,734	2,076
WHITENH18P	70.16	87.70	45.22	79.28	63.70	75.27	70.01	88.09	44.65	78.93	62.53	76.35
BLACKNH18	3,790	261	1,471	498	965	595	3,800	248	1,485	504	977	586
BLACKNH18P	27.28	9.39	52.05	17.79	35.03	21.77	27.32	8.90	52.42	18.03	35.23	21.55
AIANNH18	79	23	21	13	9	13	82	22	18	13	16	13
AIANNH18P	0.57	0.83	0.74	0.46	0.33	0.48	0.59	0.79	0.64	0.46	0.58	0.48
ASIANNH18	35	8	13	4	6	4	36	4	12	5	13	2
ASIANNH18P	0.25	0.29	0.46	0.14	0.22	0.15	0.26	0.14	0.42	0.18	0.47	0.07
HPINH18	3	2	0	0	0	1	0	0	0	0	0	0
HPINH18P	0.02	0.07	0.00	0.00	0.00	0.04	0.00	0.00	0.00	0.00	0.00	0.00
OTHERNH18	4	1	1	1	0	1	14	3	3	2	3	3
OTHERNH18P	0.03	0.04	0.04	0.04	0.00	0.04	0.10	0.11	0.11	0.07	0.11	0.11
MLTMNH18	20	0	8	1	4	7	13	2	4	3	1	3
MLTMNH18P	0.14	0.00	0.28	0.04	0.15	0.26	0.09	0.07	0.14	0.11	0.04	0.11

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 = 68.12% and BLACKNHP = 28.78% 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
($\epsilon = 10.3$)

2010 Census IDEAL POPULATION $= \frac{1,609}{4} = 402.25$ *TDA* IDEAL POPULATION $= \frac{1,617}{4} = 404.25$

	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	1,617	398	411	401	407
DEV		2.8	-3.2	-11.2	11.8		-6.2	6.8	-3.2	2.8
DEVP		0.68	-0.81	-2.88	2.84		-1.57	1.64	-0.81	0.68
TOTAL18	1,233	327	320	313	273	1,244	323	335	312	274
TOTALHISP	42	12	7	9	14	45	12	11	18	4
TOTALHISPP	2.61	2.96	1.75	2.30	3.38	2.78	3.02	2.68	4.49	0.98
TOTALNH	1,567	393	392	382	400	1,572	386	400	383	403
TOTALNHP	97.39	97.04	98.25	97.70	96.62	97.22	96.98	97.32	95.51	99.02
WHITENH	860	371	215	246	28	850	368	207	244	31
WHITENHP	53.45	91.60	53.88	62.92	6.76	56.57	92.46	50.36	60.85	7.62
BLACKNH	679	17	174	119	369	676	14	171	122	369
BLACKNHP	42.20	4.20	43.61	30.43	89.13	41.81	3.52	41.61	30.42	90.66
AIANNH	14	5	3	3	3	19	0	12	5	2
AIANNHP	0.87	1.23	0.75	0.77	0.72	1.18	0.00	2.92	1.25	0.49
ASIANNH	12	0	0	12	0	14	2	6	6	0
ASIANNHP	0.75	0.00	0.00	3.07	0.00	0.87	0.50	1.46	1.50	0.00
HPINH	0	0	0	0	0	0	0	0	0	0
HPINHP	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
OTHERNH	0	0	0	0	0	5	2	1	1	1
OTHERNHP	0.00	0.00	0.00	0.00	0.00	0.31	0.50	0.24	0.25	0.25
MLTMNNH	2	0	0	2	0	8	0	3	5	0
MLTMNNHP	0.12	0.00	0.00	0.51	0.00	0.49	0.00	0.73	1.25	0.00
HISP18	27	7	4	8	8	26	9	5	8	4
HISP18P	2.19	2.14	1.25	2.56	2.93	2.09	2.79	1.49	2.56	1.46
NONHISP18	1,206	320	316	305	265	1,218	314	330	304	270
NONHISP18P	97.81	97.86	98.75	97.44	97.07	97.91	97.21	98.51	97.44	98.54
WHITENH18	723	302	188	210	23	717	301	183	208	25
WHITENH18P	58.64	92.35	58.75	67.09	8.42	57.64	93.19	54.63	66.67	9.12
BLACKNH18	462	14	127	81	240	464	9	132	81	242
BLACKNH18P	37.47	4.28	39.69	25.88	87.91	37.30	2.79	39.40	25.96	88.32
AIANNH18	10	4	1	3	2	11	0	6	3	2
AIANNH18P	0.81	1.22	0.31	0.96	0.73	0.88	0.00	1.79	0.96	0.73
ASIANNH18	10	0	0	10	0	14	2	6	6	0
ASIANNH18P	0.81	0.00	0.00	3.19	0.00	1.13	0.62	1.79	1.92	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	5	2	1	1	1
OTHERNH18P	0.00	0.00	0.00	0.00	0.00	0.40	0.62	0.30	0.32	0.36
MLTMNH18	1	0	0	1	0	7	0	2	5	0
MLTMNH18P	0.08	0.00	0.00	0.32	0.00	0.56	0.00	0.60	1.60	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 = 52.57% and BLACKNHP = 41.81% 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 = 50.36% 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

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

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: $\bar{C}_T(g) = 17,685$ persons; $\sqrt{V(1)_g} = 33$ persons; and $RV(1)_g = 0.002$. We also observe: $C_S(g) = 17,705$ persons; $\sqrt{V(2)_g} = 39$ 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 ($\epsilon = 10.3$):

> *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	0.011
Rhode Island	CD-02	526,283.50	0.016
Rhode Island	SLDU-01	27,699.10	0.062
Rhode Island	SLDU-02	27,699.10	0.093
Rhode Island	SLDU-03	27,699.10	0.079
Rhode Island	SLDU-04	27,699.10	0.075
Rhode Island	SLDL-01	14,034.2	0.118
Rhode Island	SLDL-02	14,034.20	0.082
Rhode Island	SLDL-03	14,034.20	0.090
Rhode Island	SLDL-04	14,034.20	0.100
Panola County, MS	D-01	6,941.40	0.373
Panola County, MS	D-02	6,941.40	0.405
Panola County, MS	D-03	6,941.40	0.347
Panola County, MS	D-04	6,941.40	0.395
Panola County, MS	D-05	6,941.40	0.367
Tate County Schools, MS	D-01	3,764.60	0.439
Tate County Schools, MS	D-02	3,764.60	0.508
Tate County Schools, MS	D-03	3,764.60	0.522
Tate County Schools, MS	D-04	3,764.60	0.523
Tate County Schools, MS	D-05	3,764.60	0.568
Tylertown, MS	D-01	402.25	0.667
Tylertown, MS	D-02	402.25	0.644
Tylertown, MS	D-03	402.25	0.491
Tylertown, MS	D-04	402.25	0.832

Plot of AVERV(1) for IDEAL POPULATION Values Noted Above

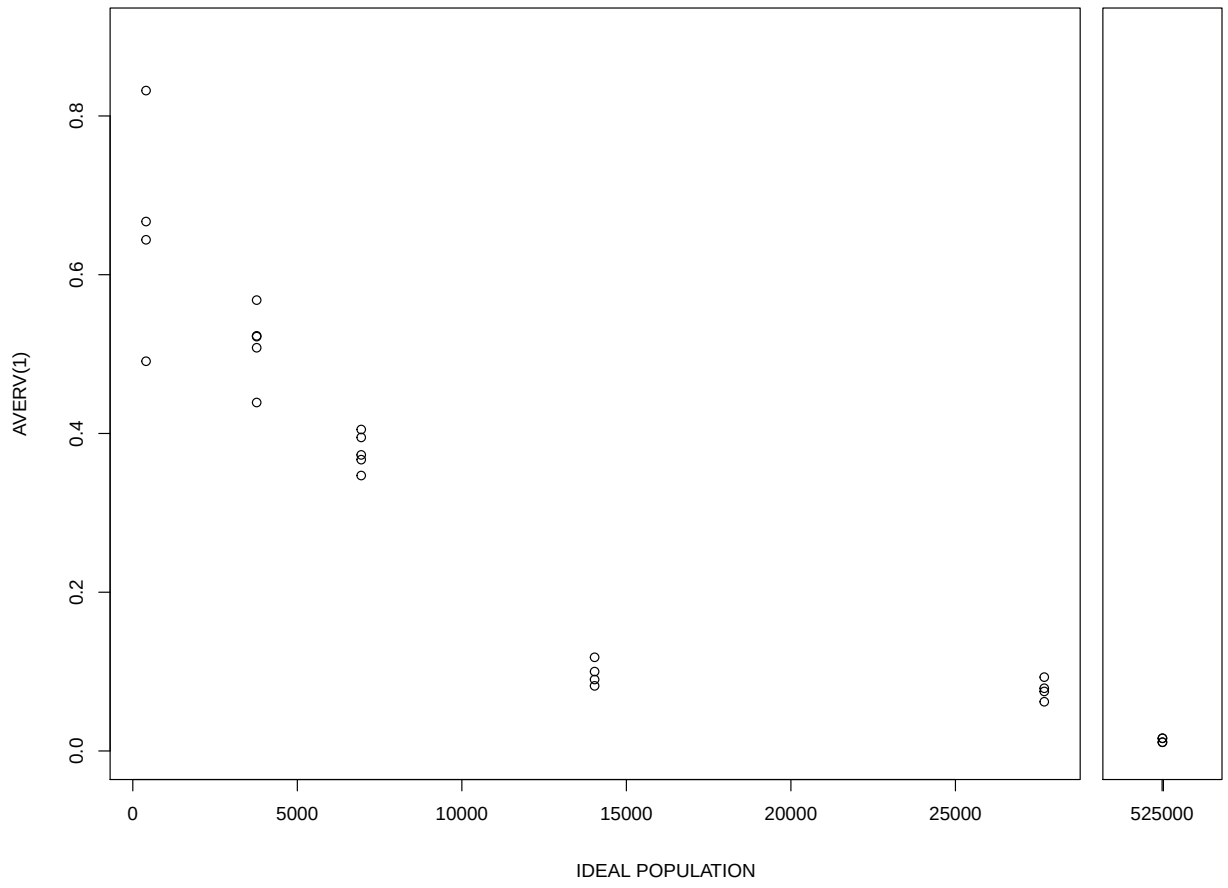


Table 7V. Counts & Measures of Variation for Rhode Island Twenty-five Runs of the *TDA* for Congressional Districts (CD) 01, and 02 (2013)
($\epsilon = 10.3$)

	(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	526,138	526,283	526,429	526,284
	0	0	162	217	162	217
	0.000	0.000	0.000	0.000	0.000	0.000
TOTAL18	828,610	828,611	412,694	412,778	415,916	415,833
	28	28	106	135	107	135
	0.000	0.000	0.000	0.000	0.000	0.000
TOTALHISP	130,666	130,655	76,180	76,100	54,477	54,555
	27	23	88	125	93	121
	0.000	0.000	0.001	0.002	0.002	0.002
TOTALNH	921,901	921,912	449,949	450,183	471,953	471,729
	21	23	125	265	123	255
	0.000	0.000	0.000	0.001	0.000	0.001
WHITENH	803,691	803,685	377,029	377,109	426,662	426,576
	12	14	37	88	39	94
	0.000	0.000	0.000	0.000	0.000	0.000
BLACKNH	57,919	57,927	37,645	37,627	20,274	20,300
	13	15	62	65	60	66
	0.000	0.000	0.002	0.002	0.003	0.003
AIANNH	6,851	6,839	3,129	3,142	3,722	3,697
	16	20	37	40	33	42
	0.002	0.003	0.012	0.013	0.009	0.011
ASIANNH	34,192	34,194	17,685	17,705	16,507	16,489
	15	15	33	39	30	35
	0.000	0.000	0.002	0.002	0.002	0.002
HPINH	660	655	376	383	284	272
	11	12	28	29	24	27
	0.017	0.018	0.075	0.076	0.085	0.098
OTHERNH	10,291	10,296	8,470	8,492	1,821	1,804
	14	15	41	47	45	48
	0.001	0.001	0.005	0.006	0.025	0.026
MLTMNNH	8,298	8,316	5,615	5,725	2,682	2,591
	27	32	71	131	64	112
	0.003	0.004	0.013	0.023	0.024	0.043
HISP18	84,723	84,715	49,352	49,303	35,372	35,412
	23	24	72	87	72	83
	0.000	0.000	0.001	0.002	0.002	0.002
NONHISP18	743,887	743,896	363,342	363,475	380,544	380,421
	27	28	66	148	67	140
	0.000	0.000	0.000	0.000	0.000	0.000
WHITENH18	660,826	660,823	312,199	312,240	348,627	348,583
	12	12	26	48	27	52
	0.000	0.000	0.000	0.000	0.000	0.000
BLACKNH18	39,482	39,485	25,403	25,402	14,079	14,083
	14	14	39	39	42	42
	0.000	0.000	0.002	0.002	0.003	0.003
AIANNH18	4,968	4,963	2,315	2,332	2,653	2,631
	11	12	29	34	26	34
	0.002	0.002	0.013	0.015	0.010	0.013
ASIANNH18	25,332	25,333	13,267	13,276	12,064	12,057
	12	12	25	26	26	27
	0.000	0.000	0.002	0.002	0.002	0.002
HPINH18	502	500	299	307	203	193
	10	10	22	23	19	22
	0.020	0.020	0.074	0.076	0.096	0.113
OTHERNH18	7,284	7,290	6,060	6,061	1,223	1,229
	10	11	35	35	36	37
	0.001	0.002	0.006	0.006	0.030	0.030
MLTMNH18	5,493	5,502	3,798	3,857	1,695	1,645
	24	26	48	76	43	66
	0.004	0.005	0.013	0.020	0.025	0.040
$AVERV(\cdot)$	0.003	0.003	0.011	0.012	0.016	0.020
$MEDRV(\cdot)$	0.000	0.000	0.002	0.002	0.003	0.003

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) ($\epsilon = 10.3$)

	(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	27,875	28,161	27,783	28,079	28,620	28,398	28,221	28,201
	59	292	54	301	72	233	43	48
	0.002	0.010	0.002	0.011	0.003	0.008	0.002	0.002
TOTAL18	20,781	20,914	19,677	19,846	25,421	25,361	23,597	23,599
	52	143	47	176	62	86	36	37
	0.002	0.007	0.002	0.009	0.002	0.003	0.002	0.002
TOTALHISP	10,105	10,282	16,082	16,288	1,527	1,409	3,159	3,217
	46	182	38	209	45	126	32	66
	0.005	0.018	0.002	0.013	0.029	0.090	0.010	0.021
TOTALNH	17,769	17,879	11,701	11,791	27,093	26,989	25,062	24,984
	46	119	42	99	45	113	45	90
	0.003	0.007	0.004	0.008	0.002	0.004	0.002	0.004
WHITENH	10,207	10,222	3,518	3,553	22,026	22,028	21,280	21,210
	18	23	16	38	18	18	17	72
	0.002	0.002	0.005	0.011	0.001	0.001	0.001	0.003
BLACKNH	4,861	4,862	1,318	4,332	1,156	1,124	2,313	2,348
	26	26	25	28	28	13	22	23
	0.005	0.005	0.006	0.007	0.024	0.038	0.010	0.010
AIANNH	267	283	197	216	149	135	183	172
	19	25	13	23	19	23	16	19
	0.072	0.087	0.068	0.106	0.129	0.174	0.085	0.112
ASIANNH	1,551	1,526	3,053	3,032	3,244	3,262	795	826
	27	37	16	26	20	27	22	38
	0.017	0.024	0.005	0.009	0.006	0.008	0.028	0.046
HPINH	22	25	13	11	19	16	13	14
	8	8	9	9	8	8	5	5
	0.358	0.338	0.656	0.822	0.425	0.525	0.404	0.381
OTHERNH	452	457	202	189	244	224	238	241
	19	20	12	18	20	28	20	21
	0.042	0.043	0.061	0.097	0.081	0.124	0.086	0.086
MLTMNNH	409	504	398	458	256	200	210	173
	27	99	28	66	22	60	26	46
	0.067	0.196	0.070	0.143	0.086	0.300	0.125	0.264
HISP18	6,364	6,458	10,894	11,014	1,268	1,241	2,058	2,097
	38	102	27	123	38	47	27	47
	0.006	0.016	0.002	0.011	0.030	0.038	0.013	0.022
NONHISP18	14174	14,456	8,783	8,832	24,153	24,120	21,538	21,502
	31	50	34	60	42	53	31	48
	0.002	0.003	0.004	0.007	0.002	0.002	0.001	0.002
WHITENH18	9,133	9,131	3,041	3,062	19,683	19,682	18,882	18,839
	13	13	13	24	17	17	16	46
	0.001	0.001	0.004	0.008	0.001	0.001	0.001	0.002
BLACKNH18	3,302	3,309	3,021	3,027	986	973	1,599	1,599
	22	23	17	18	23	27	18	18
	0.007	0.007	0.006	0.006	0.023	0.027	0.011	0.011
AIANNH18	190	197	147	154	113	110	137	136
	15	17	11	13	15	16	13	13
	0.080	0.085	0.074	0.085	0.130	0.144	0.097	0.098
ASIANNH18	1,187	1,170	2,148	2,135	2,984	2,989	595	611
	19	26	15	20	15	16	18	24
	0.016	0.022	0.007	0.009	0.005	0.005	0.030	0.039
HPINH18	17	20	10	11	15	14	9	13
	7	8	7	7	6	7	4	5
	0.435	0.399	0.708	0.669	0.433	0.470	0.379	0.388
OTHERNH18	325	326	135	125	189	186	175	178
	16	16	12	15	14	14	16	16
	0.048	0.048	0.086	0.120	0.073	0.076	0.089	0.090
MLTMNH18	263	303	281	318	181	166	142	126
	18	44	23	44	18	23	19	25
	0.068	0.144	0.081	0.137	0.101	0.141	0.133	0.196
$AVERV(\cdot)$	0.062	0.073	0.093	0.114	0.079	0.109	0.075	0.089
$MEDRV(\cdot)$	0.011	0.020	0.006	0.011	0.027	0.038	0.020	0.030

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) ($\epsilon = 10.3$)

	(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	14,040	13,881	13,725	13,821	13,679	13,949	13,611	13,713
	49	166	45	106	44	273	46	112
	0.003	0.012	0.003	0.008	0.003	0.020	0.003	0.008
TOTAL18	12,887	12,835	12,665	12,800	9,512	9,607	11,146	11,205
	43	68	36	139	33	100	37	70
	0.003	0.005	0.003	0.011	0.004	0.010	0.003	0.006
TOTALHISP	1,089	1,002	1,720	1,768	5,787	5,905	1,028	1,049
	24	90	30	57	33	123	31	37
	0.022	0.090	0.018	0.032	0.006	0.021	0.030	0.036
TOTALNH	12,951	12,879	12,006	12,053	7,892	8,044	12,583	12,664
	40	83	37	60	37	156	32	87
	0.003	0.006	0.003	0.005	0.005	0.019	0.003	0.007
WHITENH	9,915	9,922	8,696	8,714	3,446	3,465	9,548	9,539
	18	20	24	30	13	23	22	24
	0.002	0.002	0.003	0.003	0.004	0.007	0.002	0.002
BLACKNH	614	581	1,123	1,125	2,976	3,015	1,476	1,495
	18	37	17	18	21	44	19	27
	0.029	0.061	0.016	0.016	0.007	0.015	0.013	0.018
AIANNH	61	46	102	104	164	189	112	126
	13	19	14	14	16	30	11	18
	0.209	0.419	0.132	0.131	0.096	0.157	0.103	0.144
ASIANNH	2,156	2,175	1,773	1,776	801	794	795	792
	21	29	17	17	18	19	20	20
	0.010	0.013	0.010	0.010	0.022	0.024	0.025	0.026
HPINH	11	12	13	16	10	12	6	1
	6	6	5	6	5	6	4	7
	0.568	0.518	0.376	0.357	0.521	0.462	0.705	0.597
OTHERNH	75	57	143	148	250	257	380	396
	13	22	16	16	16	17	14	21
	0.172	0.391	0.110	0.111	0.063	0.067	0.037	0.054
MLTMNNH	121	86	155	170	246	312	266	315
	16	39	18	23	23	70	18	52
	0.135	0.453	0.114	0.135	0.092	0.224	0.066	0.166
HISP18	981	951	1,404	1,475	3,484	3,518	675	693
	20	36	28	76	24	42	24	31
	0.021	0.038	0.020	0.051	0.007	0.012	0.036	0.044
NONHISP18	11,907	11,884	11,261	11,325	6,029	6,089	10,471	10,512
	34	41	33	72	25	65	25	48
	0.003	0.003	0.003	0.006	0.004	0.011	0.002	0.005
WHITENH18	9,080	9,081	8,324	8,339	3,037	3,040	8,127	8,119
	16	16	19	24	11	11	15	17
	0.002	0.002	0.002	0.003	0.004	0.004	0.002	0.002
BLACKNH18	574	560	963	972	1,949	1,971	1,139	1,144
	14	20	16	19	16	27	14	15
	0.025	0.036	0.017	0.019	0.008	0.014	0.013	0.013
AIANNH18	52	45	80	82	114	129	95	101
	11	13	12	12	12	19	9	11
	0.212	0.286	0.152	0.150	0.105	0.149	0.094	0.108
ASIANNH18	2,040	2,052	1,650	1,655	583	575	642	635
	18	21	10	11	16	18	13	15
	0.009	0.010	0.006	0.007	0.027	0.031	0.020	0.023
HPINH18	9	10	11	14	7	11	4	1
	6	6	5	6	5	6	3	4
	0.630	0.583	0.440	0.412	0.653	0.549	0.712	3.924
OTHERNH18	59	51	117	126	183	190	270	280
	10	13	11	14	13	14	10	14
	0.168	0.251	0.097	0.113	0.069	0.076	0.037	0.051
MLTMNH18	93	85	116	137	154	173	195	232
	13	15	14	25	16	25	13	39
	0.139	0.177	0.119	0.185	0.104	0.143	0.068	0.170
$AVERV(\cdot)$	0.118	0.168	0.082	0.088	0.090	0.101	0.100	0.570
$MEDRV(\cdot)$	0.023	0.051	0.017	0.026	0.015	0.022	0.028	0.031

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
($\epsilon = 10.3$)

	(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	34,706	34,707	7,023	6,974	6,565	6,549	7,064	7,074	7,073	7,105	6,982	7,005
	5	5	23	54	25	30	24	26	27	42	25	34
	0.000	0.000	0.003	0.008	0.004	0.005	0.003	0.004	0.004	0.006	0.004	0.005
TOTAL18	25,368	25,363	5,235	5,214	4,737	4,732	5,164	5,171	5,337	5,345	4,894	4,901
	15	16	19	28	21	22	17	19	22	23	21	22
	0.001	0.001	0.004	0.005	0.004	0.005	0.003	0.004	0.004	0.004	0.004	0.005
TOTALHISP	502	494	91	66	74	75	82	85	104	120	150	148
	12	15	11	27	9	9	7	7	10	10	11	11
	0.025	0.030	0.124	0.415	0.117	0.117	0.081	0.086	0.097	0.156	0.075	0.077
TOTALNH	34,204	34,213	6,932	6,908	6,491	6,474	6,981	6,989	6,969	6,985	6,832	6,857
	14	16	22	33	24	29	22	23	29	34	27	37
	0.000	0.000	0.003	0.005	0.004	0.005	0.003	0.003	0.004	0.005	0.004	0.005
WHITENH	16,987	16,981	2,453	2,419	2,084	2,096		4,0194,030	5,251	5,250	3,180	3,186
	12	14	12	36	18	21	17	20	19	19	17	18
	0.001	0.001	0.005	0.015	0.008	0.010	0.004	0.005	0.004	0.004	0.005	0.006
BLACKNH	16,890	16,899	4,409	4,427	4,357	4,332	2,920	2,925	1,661	1,658	3,544	3,557
	12	15	19	26	18	31	17	17	15	15	14	19
	0.001	0.001	0.004	0.006	0.004	0.007	0.006	0.006	0.009	0.009	0.004	0.005
AIANNH	141	148	31	26	23	20	17	15	27	38	43	49
	7	10	6	7	5	6	5	5	6	13	7	9
	0.051	0.068	0.184	0.280	0.214	0.300	0.268	0.337	0.222	0.335	0.166	0.189
ASIANNH	87	89	11	8	9	7	8	5	11	17	47	52
	12	12	4	5	5	5	4	5	5	8	7	8
	0.133	0.132	0.0398	0.668	0.492	0.739	0.524	1.023	0.416	0.44	0.141	0.157
HPINH	4	4	1	0	1	0	1	0	1	2	0	2
	3	3	1	2	1	2	1	2	1	2	1	2
	0.651	0.693	1.544	Inf	1.337	Inf	1.187	Inf	1.634	0.877	1.58	0.854
OTHERNH	18	19	5	7	4	5	3	1	3	3	4	3
	7	7	3	3	3	3	2	2	2	2	3	3
	0.371	0.356	0.565	0.494	0.834	0.660	0.611	2.392	0.690	0.739	0.856	1.033
MLTMNNH	76	73	22	21	13	14	13	13	16	17	13	8
	14	15	6	6	5	5	5	5	5	5	5	7
	0.184	0.199	0.291	0.303	0.376	0.353	0.361	0.366	0.320	0.305	0.374	0.906
HISP18	304	298	55	44	48	44	50	52	58	63	93	95
	8	10	10	15	6	7	7	7	6	8	8	8
	0.027	0.034	0.179	0.340	0.119	0.162	0.136	0.137	0.107	0.126	0.086	0.087
NONHISP18	25,064	25,065	5,180	5,170	4,689	4,688	5,114	5,119	5,279	5,282	4,802	4,806
	16	16	18	20	19	19	16	17	22	22	20	21
	0.001	0.001	0.003	0.004	0.004	0.004	0.003	0.003	0.004	0.004	0.004	0.004
WHITENH18	13,461	13,455	2,043	2,025	1,711	1,732	3,073	3,072	4,124	4,115	2,510	2,511
	10	12	12	21	13	25	15	15	13	16	12	13
	0.001	0.001	0.006	0.011	0.008	0.015	0.005	0.00	0.003	0.004	0.005	0.005
BLACKNH18	11,390	11,394	3,093	3,099	2,945	2,928	2,104	2,024	1,119	1,118	2,219	2,225
	9	10	13	15	14	22	11	15	13	13	11	13
	0.001	0.001	0.004	0.005	0.005	0.008	0.005	0.007	0.011	0.011	0.005	0.006
AIANNH18	110	115	23	21	18	16	13	11	19	29	37	38
	6	8	4	5	4	5	4	5	5	11	6	6
	0.057	0.070	0.189	0.218	0.244	0.309	0.345	0.444	0.260	0.376	0.157	0.155
ASIANNH18	54	54	8	8	7	5	4	2	7	12	28	27
	8	8	4	4	4	4	3	3	4	7	4	4
	0.150	0.149	0.481	0.474	0.547	0.842	0.656	1.718	0.604	0.551	0.144	0.152
HPINH18	4	2	1	0	1	0	1	0	1	1	0	1
	3	3	1	1	1	1	1	1	1	1	1	1
	0.716	1.517	1.581	Inf	1.356	Inf	1.196	Inf	1.736	1.281	1.736	0.894
OTHERNH18	5	5	1	1	1	0	1	1	1	2	1	1
	4	4	2	2	1	2	1	1	2	2	1	1
	0.788	0.788	1.501	1.876	1.985	Inf	1.061	0.980	1.372	0.900	1.516	1.456
MLTMNH18	41	40	11	16	7	7	8	9	8	5	7	3
	9	9	4	6	3	3	4	4	3	4	3	6
	0.207	0.214	0.388	0.399	0.440	0.421	0.478	0.442	0.398	0.783	0.463	1.860
$AVERV(\cdot)$	0.168	0.213	0.373	Inf	0.405	Inf	0.347	Inf	0.395	0.346	0.367	0.393
$MEDRV(\cdot)$	0.039	0.051	0.182	0.291	0.167	0.231	0.202	0.237	0.165	0.230	0.113	0.119

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
($\epsilon = 10.3$)

	(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	18,812	18,823	8,915	3,914	3,878	3,893	3,641	3,665	3,749	3,697	3,629	3,654
	21	24	23	23	17	23	18	30	22	57	17	30
	0.001	0.001	0.006	0.006	0.004	0.006	0.005	0.008	0.006	0.015	0.005	0.008
TOTAL18	13,895	13,893	2,783	2,780	2,822	2,826	2,786	2,799	2,785	2,755	2,719	2,733
	21	21	16	16	15	15	14	19	13	13	15	21
	0.001	0.001	0.006	0.006	0.005	0.005	0.005	0.007	0.005	0.012	0.006	0.008
TOTALHISP	418	399	93	87	63	63	103	110	70	32	90	107
	12	22	8	10	8	8	5	9	9	39	9	19
	0.028	0.056	0.091	0.117	0.124	0.125	0.052	0.081	0.131	1.206	0.104	0.182
TOTALNH	18,394	18,424	3,822	3,827	3,815	3,830	3,538	3,555	3,680	3,665	3,540	3,547
	23	38	19	20	15	21	16	23	20	24	14	16
	0.001	0.002	0.005	0.005	0.004	0.006	0.005	0.007	0.005	0.007	0.004	0.005
WHITENH	12,815	12,841	3,395	3,378	1,610	1,628	2,839	2,860	2,282	2,293	2,689	2,682
	12	29	11	20	9	20	12	24	11	16	10	12
	0.001	0.002	0.003	0.006	0.005	0.012	0.004	0.008	0.005	0.007	0.004	0.005
BLACKNH	5,394	5,389	371	400	2,159	2,139	674	666	1,364	1,349	825	835
	15	16	13	31	7	22	10	13	16	22	10	14
	0.003	0.003	0.035	0.078	0.003	0.010	0.014	0.019	0.012	0.016	0.013	0.017
AIANNH	96	103	32	32	23	26	13	19	16	11	12	15
	8	11	7	7	6	7	5	8	4	7	4	5
	0.083	0.102	0.216	0.219	0.270	0.264	0.370	0.408	0.268	0.595	0.378	0.366
ASIANNH	49	47	15	14	12	16	7	6	9	7	6	4
	9	10	5	5	6	7	3	3	4	4	3	4
	0.191	0.204	0.298	0.333	0.516	0.461	0.516	0.572	0.386	0.616	0.570	0.990
HPINH	2	3	0	2	1	0	1	0	0	0	0	1
	3	3	1	2	1	1	1	1	1	1	1	1
	1.334	1.056	1.871	0.883	1.827	Inf	1.895	Inf	2.061	inf	2.578	1.265
OTHERNH	11	9	3	1	3	5	1	1	2	1	2	1
	5	6	2	3	2	3	1	1	2	2	2	2
	0.494	0.619	0.744	2.898	0.835	0.649	1.359	1.470	0.799	2.227	1.046	2.154
MLTMNNH	28	32	5	0	7	16	4	3	6	4	6	9
	10	10	4	6	5	10	2	3	3	4	4	5
	0.346	0.328	0.694	Inf	0.645	0.612	0.701	0.841	0.552	0.958	0.621	0.537
HISP18	225	215	48	47	36	34	61	63	33	16	48	55
	10	15	7	7	7	8	4	4	6	18	7	10
	0.46	0.068	0.147	0.153	0.207	0.222	0.63	0.071	0.183	1.122	0.140	0.180
NONHISP18	13,669	13,678	2,734	2,733	2,786	2,792	2,726	2,736	2,752	2,739	2,671	2,678
	17	19	12	12	11	12	13	17	14	19	14	15
	0.001	0.001	0.005	0.005	0.004	0.004	0.005	0.006	0.005	0.007	0.005	0.006
WHITENH18	9,743	9,747	2,453	2,438	1,270	1,278	2,203	2,219	1,752	1,755	2,066	2,057
	9	10	8	17	7	10	8	18	11	11	9	12
	0.001	0.001	0.003	0.007	0.005	0.008	0.003	0.008	0.006	0.006	0.004	0.006
BLACKNH18	3,793	3,790	244	261	1,483	1,471	504	498	976	965	586	595
	8	9	6	18	4	12	8	9	9	14	8	12
	0.002	0.002	0.024	0.067	0.003	0.008	0.015	0.019	0.009	0.015	0.014	0.021
AIANNH18	73	79	22	23	18	21	11	13	12	9	10	13
	8	10	6	6	5	6	4	5	4	5	4	5
	0.103	0.123	0.256	0.251	0.297	0.288	0.377	0.356	0.318	0.501	0.370	0.374
ASIANNH18	37	35	10	8	10	13	5	4	8	6	5	4
	6	6	4	4	5	6	3	4	3	4	3	3
	0.169	0.183	0.402	0.522	0.518	0.464	0.689	0.915	0.432	0.614	0.552	0.684
HPINH18	2	3	0	2	0	0	0	0	0	0	0	1
	2	3	1	2	1	1	1	1	0	0	1	1
	1.462	0.919	2.291	0.917	2.236	Inf	2.134	Inf	2.894	inf	2.579	1.020
OTHERNH18	5	4	2	1	1	1	1	1	1	0	1	1
	4	4	2	2	1	1	1	1	2	2	2	2
	0.804	1.109	0.784	1.811	1.700	1.095	1.414	0.938	1.643	Inf	1.673	1.673
MLTMNH18	17	20	3	0	5	8	2	1	4	4	4	7
	8	8	3	4	4	6	2	2	3	3	3	4
	0.462	0.422	0.909	Inf	0.949	0.691	0.811	2.200	0.742	0.700	0.704	0.592
$AVERV(\cdot)$	0.277	0.260	0.439	Inf	0.508	Inf	0.522	Inf	0.523	Inf	0.568	0.505
$MEDRV(\cdot)$	0.065	0.085	0.181	0.186	0.238	0.243	0.216	0.219	0.225	0.605	0.255	0.274

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
($\epsilon = 10.3$)

	(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	1,624	1,609	409	405	402	399	410	391	404	414
	11	19	7	8	7	8	9	21	7	12
	0.007	0.012	0.017	0.019	0.018	0.020	0.022	0.053	0.017	0.030
TOTAL18	1,245	1,233	331	327	325	320	319	313	270	273
	11	17	5	6	6	8	9	11	5	6
	0.009	0.013	0.014	0.019	0.019	0.026	0.028	0.035	0.020	0.023
TOTALHISP	47	42	13	12	10	7	18	9	6	14
	8	9	3	3	4	5	4	10	3	9
	0.173	0.225	0.236	0.255	0.358	0.720	0.235	1.127	0.491	0.634
TOTALNH	1,577	1,567	396	393	392	392	392	382	398	400
	11	15	6	6	6	6	8	13	6	6
	0.007	0.009	0.014	0.016	0.016	0.016	0.21	0.033	0.014	0.015
WHITENH	850	860	364	371	208	215	246	246	32	28
	9	13	5	9	7	10	8	8	4	6
	0.011	0.016	0.014	0.023	0.033	0.044	0.034	0.034	0.123	0.211
BLACKNH	683	679	21	17	172	174	129	119	361	369
	9	10	4	6	7	7	5	12	4	9
	0.013	0.015	0.204	0.349	0.038	0.040	0.041	0.098	0.011	0.024
AIANNH	16	14	5	5	4	3	4	3	2	3
	8	8	4	4	4	4	3	3	2	2
	0.516	0.588	0.753	0.753	0.791	1.281	0.639	0.980	0.872	0.667
ASIANNH	13	12	3	0	2	0	7	12	1	0
	5	5	2	3	1	3	3	6	1	2
	0.410	0.450	0.777	Inf	0.601	inf	0.469	0.521	0.948	Inf
HPINH	3	0	1	0	1	0	1	0	0	0
	3	4	1	1	2	2	1	1	1	1
	0.962	Inf	1.563	Inf	1.549	inf	1.216	inf	2.438	Inf
OTHERNH	4	0	1	0	1	0	1	0	0	0
	4	5	1	2	2	2	2	2	1	1
	0.924	Inf	1.406	Inf	1.317	inf	1.195	Inf	1.705	Inf
MLTMNNH	9	2	2	0	2	0	4	2	1	0
	4	8	1	2	2	3	2	3	1	1
	0.508	3.959	0.933	Inf	0.752	Inf	0.669	1.513	1.061	Inf
HISP18	30	27	8	7	6	4	13	8	3	8
	7	8	3	3	3	4	4	7	2	5
	0.233	0.292	0.332	0.391	0.500	0.941	0.306	0.818	0.676	0.652
NONHISP18	1,215	1,206	323	320	319	316	306	305	266	265
	11	14	5	6	7	7	7	7	5	5
	0.009	0.012	0.015	0.018	0.021	0.024	0.024	0.024	0.017	0.018
WHITENH18	719	723	301	302	184	188	208	210	25	23
	8	9	4	4	6	7	7	7	4	4
	0.011	0.012	0.014	0.014	0.035	0.039	0.033	0.034	0.143	0.187
BLACKNH18	463	462	14	14	127	127	85	81	237	240
	8	8	3	3	5	5	4	6	2	4
	0.017	0.017	0.228	0.225	0.040	0.040	0.046	0.072	0.010	0.015
AIANNH18	12	10	4	4	3	1	3	3	2	2
	6	7	3	3	2	3	2	2	1	1
	0.528	0.665	0.743	0.750	0.758	3.225	0.737	0.800	0.832	0.686
ASIANNH18	11	10	2	0	2	0	6	10	1	0
	5	5	2	3	2	2	3	5	1	2
	0.434	0.491	0.825	Inf	0.770	Inf	0.461	0.505	1.027	Inf
HPINH18	1	0	0	0	0	0	0	0	0	0
	2	2	1	1	1	1	1	1	0	0
	1.212	Inf	2.373	Inf	1.700	Inf	1.585	Inf	2.708	Inf
OTHERNH18	2	0	0	0	0	0	1	0 0		0
	2	3	1	1	1	1	1	1	0	0
	1.333	inf	1.920	inf	2.808	Inf	1.406	inf	2.291	Inf
MLTMNH18	6	1	1	0	2	0	3	1	1	0
	3	6	1	1	1	2	2	2	1	1
	0.444	5.621	0.968	Inf	0.754	Inf	0.654	2.289	1.228	Inf
$AVERV(\cdot)$	0.388	Inf	0.667	Inf	0.644	Inf	0.491	Inf	0.832	Inf
$MEDRV(\cdot)$	0.322	0.371	0.537	0.570	0.551	1.111	0.384	0.660	0.754	0.659

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 [1].

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 10.3. Our intent has not been to discuss how the *TDA* is constructed nor 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 (noise processing) 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 "noised answer' that is to be returned to the investigator submitting the query is

"noised answer" = (the query's answer using 2010 CEF data) + (random noise which is an integer).

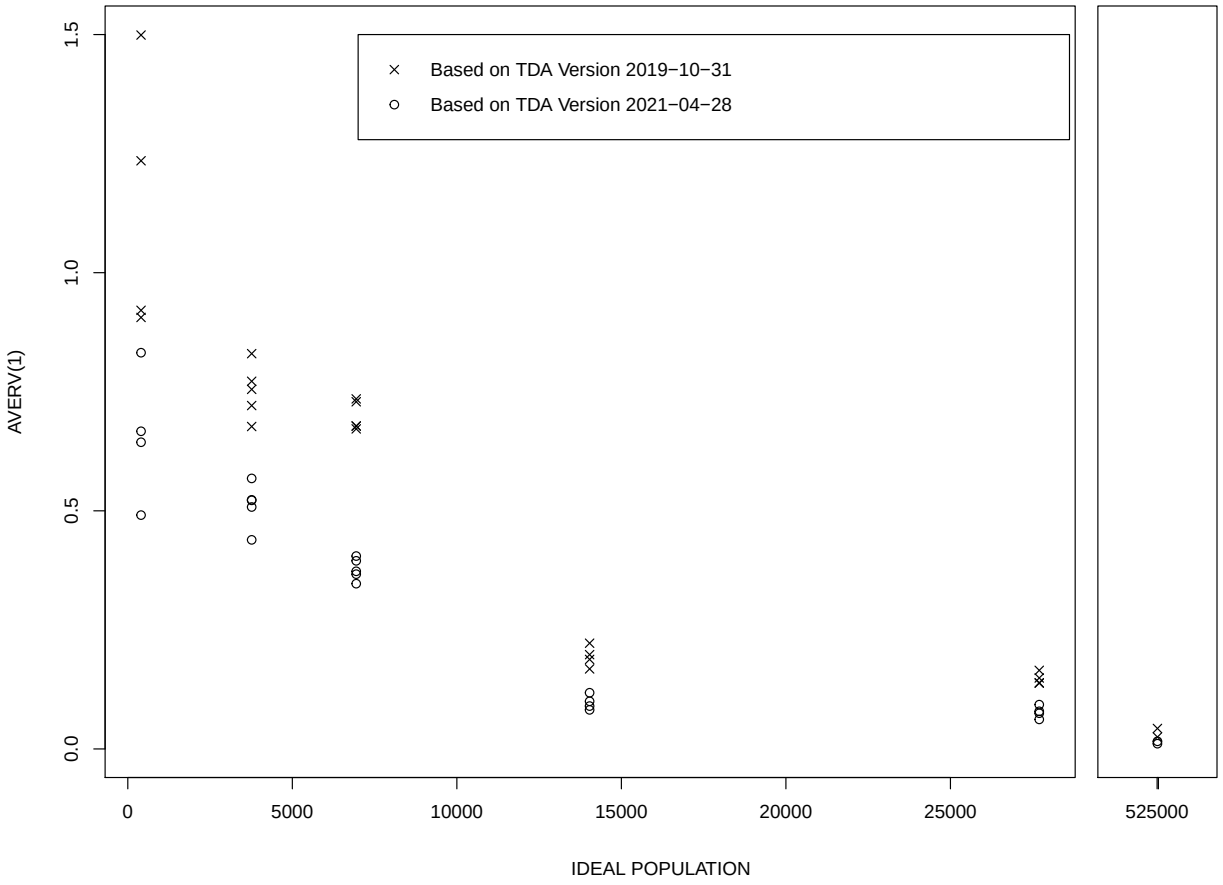
However, if the random noise is a negative integer whose absolute value is greater than the query's answer using the 2010 CEF data, then our noised answer would be a negative noised answer, which is not feasible. Thus, action is needed. This is the purpose of component two (post-processing) of the *TDA*, to ensure that our "final noised answer" to the query is a nonnegative integer. So some more work is needed before the investigator eventually gets a "final noised 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. We believe that the empirical variability reported in this study is an overall combination of variability and uncertainty from the two components.

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

We have observed reductions in variability between the 2019-10-31 version of the *TDA* with $\epsilon = 4.0$ and the 2021-04-28 version of the *TDA* with $\epsilon = 10.3$. One can see this visually by comparing Tables 7V, 8V, 9V, 10V, 11V, and 12V of this study with corresponding Tables IVa; Va; VIa; VII.1.a; VII.2.a; and VII.3.a of our earlier study [5], respectively. At a very high level, Figure 2 shows $AVERV(1)$ values for each of the districts as shown in Figure 1 using the 2019-10-31 version and the 2021-04-28 version of the *TDA*. In every case, the $AVERV(1)$ values for the 2021-04-28 version are lower than for the 2019-10-31 version.

Figure 2



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

In this remark, we repeat two specific remarks (slightly edited) made in our earlier study [5]:

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 $\epsilon = 10.3$. 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] 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.

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

[8] **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.	17,628	$(17,628 - 17,684.80)^2 = 3,226.24$	$(17,628 - 17,705)^2 = 5,929$
2.	17,685	$(17,685 - 17,684.80)^2 = 0.04$	$(17,685 - 17,705)^2 = 400$
3.	17,671	$(17,671 - 17,684.80)^2 = 190.44$	$(17,671 - 17,705)^2 = 1,156$
4.	17,669	$(17,669 - 17,684.80)^2 = 249.64$	$(17,669 - 17,705)^2 = 1,296$
5.	17,713	$(17,713 - 17,684.80)^2 = 795.24$	$(17,713 - 17,705)^2 = 64$
6.	17,692	$(17,692 - 17,684.80)^2 = 51.84$	$(17,692 - 17,705)^2 = 169$
7.	17,692	$(17,692 - 17,684.80)^2 = 51.84$	$(17,692 - 17,705)^2 = 169$
8.	17,640	$(17,640 - 17,684.80)^2 = 2,007.04$	$(17,640 - 17,705)^2 = 4,225$
9.	17,715	$(17,715 - 17,684.80)^2 = 912.04$	$(17,715 - 17,705)^2 = 100$
10.	17,625	$(17,625 - 17,684.80)^2 = 3,576.04$	$(17,625 - 17,705)^2 = 6,400$
11.	17,718	$(17,718 - 17,684.80)^2 = 1,102.24$	$(17,718 - 17,705)^2 = 169$
12.	17,707	$(17,707 - 17,684.80)^2 = 492.84$	$(17,707 - 17,705)^2 = 4$
13.	17,703	$(17,703 - 17,684.80)^2 = 331.24$	$(17,703 - 17,705)^2 = 4$
14.	17,649	$(17,649 - 17,684.80)^2 = 1,281.64$	$(17,649 - 17,705)^2 = 3,136$
15.	17,692	$(17,692 - 17,684.80)^2 = 51.84$	$(17,692 - 17,705)^2 = 169$
16.	17,736	$(17,736 - 17,684.80)^2 = 2,621.44$	$(17,736 - 17,705)^2 = 961$
17.	17,654	$(17,654 - 17,684.80)^2 = 948.64$	$(17,654 - 17,705)^2 = 2,601$
18.	17,684	$(17,684 - 17,684.80)^2 = 0.64$	$(17,684 - 17,705)^2 = 441$
19.	17,750	$(17,750 - 17,684.80)^2 = 4,251.04$	$(17,750 - 17,705)^2 = 2,025$
20.	17,678	$(17,678 - 17,684.80)^2 = 46.24$	$(17,678 - 17,705)^2 = 729$
21.	17,633	$(17,633 - 17,684.80)^2 = 2,683.24$	$(17,633 - 17,705)^2 = 5,184$
22.	17,720	$(17,720 - 17,684.80)^2 = 1,239.04$	$(17,720 - 17,705)^2 = 225$
23.	17,669	$(17,669 - 17,684.80)^2 = 249.64$	$(17,669 - 17,705)^2 = 1,296$
24.	17,723	$(17,723 - 17,684.80)^2 = 1,459.24$	$(17,723 - 17,705)^2 = 324$
25.	17,674	$(17,674 - 17,684.80)^2 = 116.64$	$(17,674 - 17,705)^2 = 961$
Totals	442,120	27,936.00	38,137.00

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

$\bar{C}_T(g) = \dfrac{442,120}{25} = 17,684.80 \approx \mathbf{17,685}$	$C_S(g) = \mathbf{17,705}$
$\sqrt{V(1)_g} = \sqrt{\dfrac{27,936}{25}} = 33.43 \approx \mathbf{33}$	$\sqrt{V(2)_g} = \sqrt{\dfrac{38,137}{25}} = 39.06 \approx \mathbf{39}$
$RV(1)_g = \dfrac{\sqrt{V(1)_g}}{\bar{C}_T(g)} = 0.00189 \approx \mathbf{0.002}$	$RV(2)_g = \dfrac{\sqrt{V(2)_g}}{C_S(g)} = 0.00221 \approx \mathbf{0.002}$

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

As an alternative to the results in Table 3, Table 3a 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 3a: Proportion of Block Groups in Each Stratum for Three Criteria
(Computations use $C_{TDA}(g)$ counts that result from 2021-04-28 version of the *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	0.1172	0.2734	0.44453
$100 \leq C_{SWA} \leq 149$	99	0.1010	0.3737	0.5253
$150 \leq C_{SWA} \leq 199$	124	0.2339	0.4516	0.6048
$200 \leq C_{SWA} \leq 249$	154	0.2208	0.6104	0.7987
$250 \leq C_{SWA} \leq 299$	209	0.2392	0.5837	0.7943
$300 \leq C_{SWA} \leq 349$	264	0.2803	0.6477	0.8750
$350 \leq C_{SWA} \leq 399$	407	0.2948	0.7297	0.8968
$400 \leq C_{SWA} \leq 449$	569	0.3199	0.7680	0.9420
$450 \leq C_{SWA} \leq 499$	915	0.3749	0.8131	0.9574
$500 \leq C_{SWA} \leq 549$	1,699	0.4097	0.8434	0.9623
$550 \leq C_{SWA} \leq 599$	3,238	0.4271	0.8786	0.9784
$600 \leq C_{SWA} \leq 649$	5,131	0.4461	0.9039	0.9827
$650 \leq C_{SWA} \leq 699$	6,683	0.4667	0.9078	0.9843
$700 \leq C_{SWA} \leq 749$	7,356	0.4927	0.9250	0.9882
$750 \leq C_{SWA} \leq 799$	8,170	0.5093	0.9300	0.9867
$800 \leq C_{SWA} \leq 849$	8,213	0.5264	0.9456	0.9911
$850 \leq C_{SWA} \leq 899$	8,441	0.5473	0.9451	0.9887
$900 \leq C_{SWA} \leq 949$	8,657	0.5563	0.9541	0.9903
$950 \leq C_{SWA} \leq 999$	8,723	0.5665	0.9631	0.9922
$1{,}000 \leq C_{SWA} \leq 1{,}049$	8,398	0.5910	0.9609	0.9894
$1{,}050 \leq C_{SWA} \leq 1{,}099$	8,345	0.6001	0.9681	0.9901
$1{,}100 \leq C_{SWA} \leq 1{,}149$	7,950	0.6057	0.9670	0.9889
$1{,}150 \leq C_{SWA} \leq 1{,}199$	7,860	0.6220	0.9738	0.9907
$1{,}200 \leq C_{SWA} \leq 1{,}249$	7,451	0.6247	0.9749	0.9886
$1{,}250 \leq C_{SWA} \leq 1{,}299$	7,124	0.6446	0.9752	0.9903
$1{,}300 \leq C_{SWA} \leq 1{,}349$	6,714	0.6555	0.9791	0.9899
$1{,}350 \leq C_{SWA} \leq 1{,}399$	6,507	0.6634	0.9793	0.9914
$1{,}400 \leq C_{SWA} \leq 1{,}449$	5,911	0.6686	0.9794	0.9895
$1{,}450 \leq C_{SWA} \leq 1{,}499$	5,617	0.6913	0.9843	0.9931
$1{,}500 \leq C_{SWA} \leq 1{,}549$	5,390	0.6970	0.9833	0.9902
$1{,}550 \leq C_{SWA} \leq 1{,}599$	4,856	0.7039	0.9827	0.9881
$1{,}600 \leq C_{SWA} \leq 1{,}649$	4,508	0.7209	0.9847	0.9889
$1{,}650 \leq C_{SWA} \leq 1{,}699$	4,325	0.7309	0.9864	0.9917
$1{,}700 \leq C_{SWA} \leq 1{,}749$	4,093	0.7393	0.9871	0.9892
$1{,}750 \leq C_{SWA} \leq 1{,}799$	3,689	0.7346	0.9878	0.9905
$1{,}800 \leq C_{SWA} \leq 1{,}849$	3,469	0.7521	0.9873	0.9902
$1{,}850 \leq C_{SWA} \leq 1{,}899$	3,252	0.7494	0.9852	0.9889
$1{,}900 \leq C_{SWA} \leq 1{,}949$	3,008	0.7643	0.9904	0.9924
$1{,}950 \leq C_{SWA} \leq 1{,}999$	2,832	0.7662	0.9866	0.9887
$2{,}000 \leq C_{SWA} \leq 2{,}049$	2,573	0.7781	0.9868	0.9891
$2{,}050 \leq C_{SWA} \leq 2{,}099$	2,356	0.7742	0.9877	0.9898
$2{,}100 \leq C_{SWA} \leq 2{,}149$	2,307	0.7807	0.9887	0.9905
$2{,}150 \leq C_{SWA} \leq 2{,}199$	2,033	0.7919	0.9843	0.9852
$2{,}200 \leq C_{SWA} \leq 2{,}249$	1,999	0.8044	0.9885	0.9900
$2{,}250 \leq C_{SWA} \leq 2{,}299$	1,892	0.8018	0.9884	0.9900
$2{,}300 \leq C_{SWA} \leq 2{,}349$	1,666	0.7995	0.9904	0.9922
$2{,}350 \leq C_{SWA} \leq 2{,}399$	1,622	0.8089	0.9883	0.9901
$2{,}400 \leq C_{SWA} \leq 2{,}449$	1,421	0.8100	0.9859	0.9859
$2{,}450 \leq C_{SWA} \leq 2{,}499$	1,350	0.8096	0.9852	0.9874
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 3a, take C^*_{SWA} to be between 900 and 949. For block groups whose TOTAL C_{SWA} count is at least 949, 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 3a, 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 the data that were released to the public (one run of the 2021-04-28 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 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 for each run, where we first observed that 0.9500 was exceeded is given in Table 4a for each run is between 450 and 499 people in 23 of the 25 runs.

Table 4a: For Each Run, the Stratum and Stratum Proportion When 0.9500 First Exceeded
(Proportion Computations use $C_{TDA}(g)$ counts that result from 2021-04-28 version of the *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	$500 \leq C_{SWA} \leq 549$	0.9659
2	$450 \leq C_{SWA} \leq 499$	0.9607
3	$450 \leq C_{SWA} \leq 499$	0.9552
4	$450 \leq C_{SWA} \leq 499$	0.9574
5	$450 \leq C_{SWA} \leq 499$	0.9574
6	$450 \leq C_{SWA} \leq 499$	0.9552
7	$450 \leq C_{SWA} \leq 499$	0.9563
8	$450 \leq C_{SWA} \leq 499$	0.9563
9	$450 \leq C_{SWA} \leq 499$	0.9628
10	$450 \leq C_{SWA} \leq 499$	0.9541
11	$450 \leq C_{SWA} \leq 499$	0.9716
12	$450 \leq C_{SWA} \leq 499$	0.9628
13	$450 \leq C_{SWA} \leq 499$	0.9574
14	$450 \leq C_{SWA} \leq 499$	0.9585
15	$450 \leq C_{SWA} \leq 499$	0.9650
16	$450 \leq C_{SWA} \leq 499$	0.9574
17	$450 \leq C_{SWA} \leq 499$	0.9607
18	$450 \leq C_{SWA} \leq 499$	0.9705
19	$500 \leq C_{SWA} \leq 549$	0.9670
20	$450 \leq C_{SWA} \leq 499$	0.9628
21	$450 \leq C_{SWA} \leq 499$	0.9683
22	$450 \leq C_{SWA} \leq 499$	0.9650
23	$450 \leq C_{SWA} \leq 499$	0.9541
24	$450 \leq C_{SWA} \leq 499$	0.9607
25	$450 \leq C_{SWA} \leq 499$	0.9672