October 24, 2019

By Email to OIRA_Submission@omb.eop.gov

RE: 2020 Census Post-Enumeration Survey Initial and Final Housing Unit Follow-up

To Whom It May Concern:

On behalf of Campaign Legal Center ("CLC"), we are pleased to offer this comment in response to the Federal Register notice published September 24, 2019, concerning the Census Bureau's plans for the 2020 Census Post-Enumeration Survey Initial Housing Unit Follow-up and Final Housing Unit Follow-up.[1]

CLC supports the collection of information for the Post-Enumeration Survey ("PES"), including the Housing Unit Follow-up ("HUFU") operations. The Bureau should follow through with its stated plans to use the PES to produce and publish estimates of net coverage and components of coverage[2] for the U.S. and for each state following the 2020 Census.

However, CLC urges the Bureau to go further and publish coverage estimates for large cities and counties, as the Bureau has done for recent decennial censuses. These estimates of sub-state census coverage are critical for guiding state policymaking, evaluating the fairness of statewide electoral processes, and facilitating efforts to improve future censuses.

To the extent the Bureau currently plans not to produce estimates of sub-state coverage for public consumption, the Bureau should reverse that plan and

---

[1] Submission for OMB Review; Comment Request, 84 Fed. Reg. 49983 (Sept. 24, 2019).

[2] "Net coverage" refers to the difference between the population recorded in official census figures and the estimated true population. The components of coverage are correct enumerations, erroneous enumerations, whole-person imputations, and omissions.

make any necessary adjustments to the PES operations, including the estimated workloads for Initial HUFU and Final HUFU.

I.   Background

In recent decades, the Census Bureau has used a technique known as "dual-system estimation" to evaluate the accuracy of each decennial census.[3] Dual-system estimation involves conducting a post-enumeration survey independent from the census, matching the responses with census data, and using the results to estimate how close the census came to capturing the true size of the population.[4]

For the 1990, 2000, and 2010 Censuses, the Bureau created and ultimately published estimates of the net undercount or overcount for the U.S. as a whole, for each state, and for certain smaller areas within states.[5]

For the 2020 Census, the Bureau plans to conduct the PES in a manner similar to 2010.[6] However, it appears that the Bureau has decided, at least tentatively, not to publish estimates of coverage for any geographic units within states. In a 2017 memorandum, Timothy Kennel, Assistant Division Chief for Statistical Methods, recommended that the Bureau not disseminate sub-state coverage estimates based on the results of the 2020 Census PES. Dr. Kennel stated that one of the "primary reasons" for this recommendation was "a perceived lack of demand" for sub-state coverage estimates among data users.

> We do not see a demand for place and county coverage estimates. Because there are no plans to use the post-enumeration survey to adjust the census, there is very little interest in sub-state estimates. We responded to no more than two requests, questions,

---

[3] *Decennial Census of Population and Housing – Coverage Measurement*, U.S. CENSUS BUREAU, https://www.census.gov/programs-surveys/decennial-census/about/coverage-measurement.2010.html (last visited Oct. 15, 2019).

[4] *Id.*

[5] Peter P. Davis & James Mulligan, *Census Coverage Measurement Estimation Report: Net Coverage for the Household Population in the United States* at 9, U.S. CENSUS BUREAU (May 12, 2012), available at https://www.census.gov/coverage_measurement/pdfs/g01.pdf (describing sub-state estimates produced for 2010 Census); Eric Schindler, *A.C.E. Revision II – Adjusted Data for States, Counties, and Places*, U.S. CENSUS BUREAU (April 9, 2003), available at https://www.census.gov/dmd/www/pdf/pp-60r.pdf (discussing estimates produced for 2000 Census); *1990 Net Undercount and Undercount Rate for Counties by State*, U.S. CENSUS BUREAU, https://www.census.gov/dmd/www/undercounty.html (last updated June 14, 2010).

[6] *2020 Census Operational Plan v4.0* at 152, U.S. CENSUS BUREAU (December 2018), available at https://www2.census.gov/programs-surveys/decennial/2020/program-management/planning-docs/2020-oper-plan4.pdf.

and comments on the sub-state coverage estimates after the release of the 2010 Census Coverage Measurement reports. The strong synthetic assumption required for the place and county estimates decreases their utility because they are not based solely on data specific to the individual areas. Furthermore, because mean squared error estimates were so large, there were no significant differences between the coverage estimates for the various counties and places.[7]

Dr. Kennel also stated that refraining from disseminating sub-state coverage estimates would reduce the risk of "schedule delays" and "data errors."[8]

On December 13, 2017, Dr. Kennel's recommendation was approved at a meeting of the 2020 Census Portfolio Management Governing Board. According to a document memorializing the meeting, "[i]t is planned that the [sub-state] estimates will be produced but not publicly disseminated."[9]

For convenience, Dr. Kennel's memorandum is attached to this comment as Exhibit A, and the document memorializing the December 13, 2017 meeting is attached as Exhibit B.

II.   Reasons to publish sub-state census coverage estimates

The Census Bureau appears to believe that data users outside the Bureau are uninterested in estimates of sub-state census coverage. That is not the case. Sub-state coverage estimates, even when imprecise, are useful for many important purposes and will be badly missed if the Bureau fails to publish them for the 2020 Census.

---

[7] Memorandum from Timothy Kennel to Patrick Cantwell, Recommendation to Not Disseminate Sub-State Estimates of Coverage from the Post-Enumeration Survey at 3, U.S. CENSUS BUREAU (2017), available at https://drive.google.com/drive/folders/1NMEkgqzLSzaonAArnU9nCFV18H7TN-X8 (select "NAACP Production," then "PMGB," then "PMGB 2017," then "12-13-17," then "Recommendation to Not Disseminate Sub-State Estimates of Coverage from the 2020 PES d3.pdf"). This document was published by the National Association for the Advancement of Colored People, which obtained it through a Freedom of Information Act request.

[8] *Id.*

[9] *2020 Census Portfolio Management Governing Board Meeting - December 13, 2017*, U.S. CENSUS BUREAU (Dec. 13, 2017), available at https://drive.google.com/drive/folders/1NMEkgqzLSzaonAArnU9nCFV18H7TN-X8 (select "NAACP Production," then "PMGB," then "PMGB 2017," then "12-13-17," then "Dec 13 2017 Meeting Information.pdf"). This document was published by the NAACP, which obtained it through its FOIA request.

A. Differential undercounts within states are relevant to state-level policymaking throughout the next decade.

One of the most important functions of the decennial census is to help states understand the distribution of population within their borders, so they can make informed policy decisions in areas ranging from education to economic development to redistricting. State policymakers and their constituents therefore have a strong interest in knowing whether the census undercounts some parts of the state more than others.

Differential undercounts within states are particularly relevant to policy debates about the allocation of state funding to local communities. Like the federal government, state governments use population data from the census to divide up vast sums of public money.[10] If some cities or counties within a state are disproportionately undercounted, census-based funding may produce an unfair result for those localities.

In recent decades, policymakers and stakeholders have actively discussed and debated how state funding policy should respond to intrastate differential undercounts. [11] Without sub-state coverage estimates, undercounted communities will lack critical evidence necessary to advocate for funding adjustments, and may not even be aware they are undercounted.

---

[10] *See* PRICEWATERHOUSECOOPERS, EFFECT OF CENSUS 2000 UNDERCOUNT ON FEDERAL FUNDING TO STATES AND SELECTED COUNTIES, 2002-2012 i (Aug. 7, 2001), available at http://govinfo.library.unt.edu/cmb/cmbp/reports/080601.pricewaterhouse/downloads/undercount_080601.pdf ("Many state-funded grant programs to localities also rely on census counts, compounding the misallocation of grant money" at the federal level.); Kravitz v. U.S. Dep't of Commerce, 336 F. Supp. 3d 545, 558 (D. Md. 2018) (quoting Glavin v. Clinton, 19 F. Supp. 2d 543, 550 (E.D. Va. 1998)) ("[I]t has been recognized 'that there is a direct correlation between decennial census population counts and federal and state funding allocations.'").

[11] *See, e.g.*, Jessica Hansen, *U.S. census missed about 37,000 state residents; 'Undercount' data may affect funding allocations*, MILWAUKEE JOURNAL SENTINEL (Dec. 7, 2002) (accessed via LexisNexis) (quoting Milwaukee mayor's policy director predicting that "efforts will be made" to adjust state funding for Milwaukee based on undercount estimated in Census Bureau data); Michele R. Marcucci, *Census errs by 83,000 in Bay Area, data reveals*, OAKLAND TRIBUNE (Dec. 7, 2002) (accessed via LexisNexis) ("Alameda County Supervisor Keith Carson said the adjusted numbers should be used [for intrastate funding], so that local governments can provide services that better meet the needs of communities that may not have been fully accounted for in the unadjusted numbers."); *cf.* Josh Wood, *North Dakota oil boom makes census count difficult*, ASSOCIATED PRESS (May 7, 2014), https://www.washingtontimes.com/news/2014/may/7/nds-oil-patch-presents-challenges-to-census/ (reporting on U.S. Senator Heidi Heitkamp's suggestion that census-based funding mechanisms should be altered in fairness to North Dakota communities whose rapid growth the census fails to capture).

Redistricting is another area where sub-state census coverage estimates are relevant to a legitimate policy debate. States typically use unadjusted census data as the population base for redistricting, but to perform this task in a fully informed and equitable manner, they need to understand the differential undercount within the state. In a typical redistricting process, there are many possible ways to draw a map that creates equally populous districts and complies with all other legal requirements. Decision-makers therefore must choose among multiple maps, each of which will have some tendency to advantage some communities and disadvantage others. In making this choice, map-drawers would be wise to avoid maps that further disempower undercounted communities, which are already politically disadvantaged by the undercount.[12]

As these examples illustrate, evidence of a differential undercount within a state is relevant to important state policy debates. Policymakers and activists engaged in these debates have appropriately cited the Bureau's sub-state coverage estimates in the past, and will do so in the future if the Bureau continues to publish the estimates.

B. Identifying undercounted communities helps stakeholders prepare to improve future censuses.

As the Bureau correctly noted in the Federal Register notice, a principal purpose of the PES is "to improve future censuses."[13] The PES promotes this goal in part because stakeholders outside the Bureau—including nonprofit organizations and state, local, and tribal governments—are prompted to take action when they learn that their communities have likely been undercounted in the past.

In many cases, stakeholders have responded to evidence of local undercounts by undertaking efforts to encourage census participation in their communities. For example, estimated undercounts from the 2010 Census have helped motivate many local governments to form Complete Count Committees, which work alongside the Bureau to improve coverage for the 2020 Census.[14]

---

[12] Moreover, to the extent that states wish to consider adjusting the redistricting population base to counteract the differential undercount and better represent the state's total population, the states need sub-state coverage data to do so. These adjustments, if done rigorously, could be compatible with the Constitution's one-person, one-vote requirements. *See* Fletcher v. Lamone, 831 F. Supp. 2d 887, 894-95 (D. Md. 2011) (three-judge court) (citing Karcher v. Daggett, 462 U.S. 725, 738 (1983)).

[13] Submission for OMB Review; Comment Request, 84 Fed. Reg. at 49983-84.

[14] *See, e.g.*, Danielle E. Gaines, *Data Will Help Guide Officials Pushing for Complete Count in 2020 Census*, MARYLAND MATTERS (Aug. 15, 2019), https://www.marylandmatters.org/2019/08/15/data-will-help-guide-officials-pushing-for-complete-count-in-2020-census/ ("In the 2010 Census, an estimated 19,900 people were

Similarly, census advocates in the private sector have promoted get-out-the-count campaigns for the express purpose of avoiding a repeat of local undercounts that the Bureau estimated for 2010.[15] Journalists writing about the 2020 Census have also pointed to sub-state coverage estimates from the 2010 Census to put the importance of census participation in local context.[16]

Stakeholders have also pressed the Bureau to increase its own efforts to count local communities that were estimated to be undercounted in past censuses. This advocacy directed at the Bureau sometimes takes the form of public statements,[17] and other times takes the form of litigation.[18]

By encouraging their neighbors to participate in the census and pushing the Bureau to improve its methods, stakeholders outside the federal government play an important role in promoting an accurate census. This role will be

---

undercounted in Prince George's County alone, resulting in an estimated federal funding loss of more than $363 million, Harrison said. So Maryland planning officials are mobilizing 'Complete Count' committees around the state and making use of historical response data like never before."); Hayley Munguia, *Long Beach kicks off effort to count every single resident in 2020 Census*, PRESS-TELEGRAM (Aug. 9, 2019), https://www.presstelegram.com/2019/08/09/long-beach-kicks-off-effort-to-count-every-single-resident-in-2020-census/ (reporting that mayor of Long Beach, California cited local 2010 undercount in explaining importance of Long Beach's Complete Count Committee for the 2020 Census); *Edinburg Holds Kickoff 2020 Census Complete Count Committee Meeting*, CITY OF EDINBURG, TEX. (Oct 10, 2018), http://mail.cityofedinburg.com/newsevents.php?news_id=2239 ("The City of Edinburg and all of Hidalgo County were undercounted in the previous census back in 2010. As a result, the City has been working with the U.S. Census Bureau since last year.").

[15] *See, e.g.*, Stuart Schrader, Mary Elizabeth Hughes & Mac McComas, *Accurate census count critical for Baltimore kids*, BALTIMORE SUN (Sept. 5, 2019), https://www.baltimoresun.com/opinion/op-ed/bs-ed-op-0906-census-children-20190905-4vu7yj6jrja2xlsdjexg66rwt4-story.html ("Census Bureau data suggest that the Maryland undercount [in the 2010 Census] was clustered in the counties near Baltimore and DC."); Kevin Frazier, *2020 census is politically crucial for redistricting*, DAILY CALIFORNIAN (March 1, 2019), https://www.dailycal.org/2019/03/01/2020-census-is-politically-crucial-for-redistricting/ (citing estimated 2010 undercount for Berkeley, California).

[16] *See, e.g.*, Alexandra Seltzer, *PBC sends more than 700,000 addresses to census ahead of 2020*, PALM BEACH POST (Aug. 17, 2018), https://www.palmbeachpost.com/news/local/pbc-sends-more-than-700-000-addresses-census-ahead-2020/GGQUqyxl3heyfUZRjy9IAL/ (citing estimated 2010 Census undercount for Palm Beach County, Florida).

[17] *See, e.g.*, Robin Fields, *State Census Sampling Shows Huge Undercount*, L.A. TIMES (Dec. 7, 2002) (accessed via LexisNexis) (quoting a UCLA urban planning professor urging the Bureau to "utilize this information [on the estimated differential undercount] to identify areas that are hard to count" and adjust its methods accordingly).

[18] *See* NAACP v. Bureau of the Census, 382 F. Supp. 3d 349, 356, 377 (D. Md. 2019) (granting in part and denying in part motion to dismiss lawsuit where plaintiffs sought, *inter alia*, to prevent the Bureau from repeating or worsening its historical undercount of Prince George's County, Maryland).

seriously diminished if external stakeholders no longer have access to estimates of undercounts and overcounts within states.

C. Sub-state census coverage is relevant for understanding whether the state's political process is equally open to all citizens, regardless of race or ethnicity.

An accurate census is indispensable for an equitable distribution of political power within a state.[19] When some parts of the state are disproportionately undercounted, demographic groups that are concentrated in those areas are deprived of an equal opportunity to exercise power and elect their preferred candidates to statewide office.

Because sub-state differential undercounts thwart equality of political opportunity, they are relevant to enforcement of the federal Voting Rights Act ("VRA"). The VRA bans state and local voting practices that "result[] in a denial or abridgement of the right of any citizen of the United States to vote on account of race or color."[20] To decide whether a violation of the VRA has occurred, a court must consider, "based on the totality of circumstances," whether members of a racial or ethnic group "have less opportunity than other members of the electorate to participate in the political process and to elect representatives of their choice."[21]

Evidence of a disproportionate undercount of a minority community by the census is part of the totality of circumstances to be considered in deciding liability under the VRA.[22] Similarly, evidence of a local undercount can help determine whether a minority community is "sufficiently large and geographically compact to constitute a majority in a single-member district," which is a precondition for liability in certain VRA cases.[23]

Apart from the utility of sub-state coverage estimates in VRA litigation, there is intrinsic value in providing transparency about how differential undercounts have (or have not) diluted the voting strength of minority groups within a state.

---

[19] *See* Dep't of Commerce v. U.S. House of Reps., 525 U.S. 316, 332-33 (1999) (dilution of voting strength within states is a cognizable injury in litigation challenging federal census procedure).

[20] 52 U.S.C. § 10301(a).

[21] 52 U.S.C. § 10301(b).

[22] *See* Ward v. Columbus Cty., 782 F. Supp. 1097, 1104 (E.D.N.C. 1991); Perez v. Texas, No. 11–CA–360, 2011 WL 9160142, at *11-*12 (W.D. Tex. Sept. 2, 2011) (allowing plaintiffs to proceed with VRA challenge using differential census undercount as "one of many factual allegations used to support their [VRA] vote dilution claim").

[23] Thornburg v. Gingles, 478 U.S. 30, 50 (1986).

Americans whose voices in government have been muffled due to unequal undercounts deserve to know they have suffered this injury. And advocates promoting reforms to the electoral process deserve the chance to show, based on the best possible evidence, that the current playing field is not level due to differential undercounts and other injustices.

> ### D. Statistical uncertainty in sub-state coverage estimates does not prevent those estimates from being useful.

In recommending that the Bureau not publish sub-state coverage estimates for the 2020 Census, Dr. Kennel noted that the coverage estimates for the 2010 Census did not find statistically significant variation in census coverage across different counties and cities.[24] But cancelling the 2020 sub-state coverage estimates because of the statistical insignificance of the 2010 results would be a mistake, for at least two reasons: (1) because statistical significance or insignificance is itself a useful piece of evidence about the distributive accuracy of the census, and (2) because, even assuming they lack statistical significance, sub-state coverage estimates can provide a useful impression of how census coverage varies within a state.

First, the Bureau's failure to find statistically significant differences in estimated sub-state coverage is itself an important finding about the 2010 Census. Although this result certainly does not prove that the 2010 Census was equally accurate throughout the U.S., it does provide some support for an inference that the true inequalities in census coverage from place to place were not as extreme as might otherwise have been feared.[25]

If the Bureau produces sub-state coverage estimates for the 2020 Census, it may find once again that the estimated net coverage errors are statistically insignificant. On the other hand, the outcome may be different for 2020 than it was for 2010. If the Bureau were to calculate statistically significant

---

[24] Ex. A at 3.

[25] This is true even though the Bureau produces sub-state coverage estimates synthetically, rather than through direct measurement of the city or county's real population. For the 2010 Census, the Bureau calculated its sub-state coverage estimates by creating logistic regression models, accounting for multiple variables (including, for example, race and Hispanic origin) that correlate with undercoverage or overcoverage in the nationwide PES sample. *See* Andrew Keller & Tyler Fox, *Census Coverage Measurement Estimation Report: Components of Census Coverage for the Household Population in the United States* at 6, U.S. CENSUS BUREAU (May 22, 2012), available at https://www.census.gov/coverage_measurement/pdfs/g04.pdf; Ex. A at 2. A city or county with a very large real undercount or overcount would probably have population characteristics that are statistically correlated with undercounting or overcounting, and therefore would be relatively likely to have a synthetic estimate of net coverage error statistically different from zero.

estimates of local undercounts or overcounts for the 2020 Census, that finding would provide evidence that the distributive accuracy of the census decreased from 2010 to 2020.

Second, even if every individual city or county's estimated net coverage error is statistically insignificant, careful data users may still gain insight into the differential undercount by studying the estimates. The mere lack of statistical significance does not mean that the Bureau's estimates should be discarded as random statistical noise.

On the contrary, it appears that variations in the Bureau's sub-state coverage estimates do bear at least some relation to true differences in the accuracy of the census from place to place. Anecdotally, it is easy to notice patterns in the sub-state coverage estimates that align with known realities, such as the persistent undercount of people of color in decennial censuses.

For example, Alabama's population in 2010 was 67 percent non-Hispanic White, [26] while the state's three most populous cities—Birmingham, Montgomery, and Mobile—all had Black majorities.[27] The estimated statewide net undercount for Alabama in the 2010 Census was only 0.13 percent, but the estimated net undercounts for Birmingham, Montgomery, and Mobile were 1.63 percent, 1.99 percent, and 0.67 percent, respectively.[28] Similarly, other predominantly Black cities in predominantly non-Hispanic White states were estimated to have higher net undercounts than those states. Examples of this phenomenon include Atlanta, Georgia; Savannah, Georgia; Baton Rouge, Louisiana; Shreveport, Louisiana; Baltimore, Maryland; Detroit, Michigan; Jackson, Mississippi; and Memphis, Tennessee.[29]

Even though none of these cities in isolation had a statistically significant estimated undercount, the fact that so many of them fit this pattern supports

---

[26] American FactFinder, Community Facts, U.S. CENSUS BUREAU, https://factfinder.census.gov/faces/nav/jsf/pages/community_facts.xhtml (enter "Alabama" in the search box) (last visited Oct. 9, 2019).

[27] *See id.* (enter "Birmingham city, Alabama" in the search box); *id.* (enter "Montgomery city, Alabama"); *id.* (enter "Mobile city, Alabama"); *id.* (enter "Alabama," then select "Compare Cities and Towns for Population, Housing, Area, and Density").

[28] Davis & Mulligan, *supra* note 5, at 29.

[29] *See id.* at 29-35; American FactFinder, Community Facts, U.S. CENSUS BUREAU, https://factfinder.census.gov/faces/nav/jsf/pages/community_facts.xhtml (enter "Georgia") (last visited Oct. 9, 2019); *id.* (enter "Atlanta city, Georgia"); *id.* (enter "Savannah city, Georgia"); *id.* (enter "Louisiana"); *id.* (enter "Baton Rouge city, Louisiana"); *id.* (enter "Shreveport city, Louisiana"); *id.* (enter "Maryland"); *id.* (enter "Baltimore city, Maryland"); *id.* (enter "Michigan"); *id.* (enter "Detroit city, Michigan"); *id.* (enter "Mississippi"); *id.* (enter "Jackson city, Mississippi"); *id.* (enter "Tennessee"); *id.* (enter "Memphis city, Tennessee").

at least a reasonable inference—if not a scientifically bulletproof conclusion—that the variations in sub-state estimates correspond to real inequalities in census coverage. Policymakers and community advocates should not be required to ignore this collective evidence of injustice merely because the evidence is inconclusive on its own.

Given the choice between publishing imperfect data and suppressing a potentially important source of insight into the differential undercount within states, the Bureau should err on the side of transparency.

III.   Implications for the PES and HUFU operations

The proper design of a statistical survey depends on its intended uses. Consistent with this general principle, the Census Bureau should make sure that its plans for the 2020 Census PES—including the estimated workloads for the Initial HUFU and Final HUFU operations—are up to the task of producing sub-state coverage estimates as well as state and national estimates.

The HUFU operations support the PES by seeking to resolve discrepancies between the address list created for the census and the separate, independent address list created for the PES. In the Initial HUFU operation, a staff of "listers" will "collect additional information that might allow a resolution" of addresses that could not be matched between the PES address list and the preliminary census address list.[30] Similarly, the Final HUFU operation will collect information in an effort to answer questions that arise in the Bureau's matching of addresses between the PES address list and the final census address list.[31]

The Bureau has estimated that Initial HUFU will involve collecting information from 180,000 housing units, and that Final HUFU will involve collecting information from 8,000 housing units. The Bureau has not explained the precise method it used to calculate these estimates. However, it stands to reason that these estimates factor in, among other assumptions, the total number of blocks or addresses the Bureau plans to canvass in the independent listing for the PES.

In turn, those plans regarding the scope of the independent listing operation may have been premised on certain assumptions about the types of data products that would be produced using the PES. If the Bureau assumed that it would not use PES data to calculate estimates of sub-state census coverage for public consumption, then the Bureau might have designed the PES on a scale

---

[30] Submission for OMB Review; Comment Request, 84 Fed. Reg. at 49984.

[31] *Id.*

that is too small to produce data of suitable quality for use in deriving sub-state coverage estimates.

CLC takes no position on whether the Bureau's operational plans for the PES, including its assumptions about HUFU workload, are sufficient. However, the Bureau should revisit those plans, make sure they are compatible with the production of sub-state coverage estimates, and revise them if they are not.

Above all, the Bureau should commit to publishing net coverage estimates for large cities and counties after the 2020 Census. These estimates, however imperfect, are essential for judging the fairness of this census and planning to make the next one better.

<div align="center">*   *   *</div>

Respectfully submitted,

/s/ Paul M. Smith

Paul M. Smith
Vice President, Litigation & Strategy
psmith@campaignlegalcenter.org

Danielle Lang
Co-Director, Voting Rights & Redistricting
dlang@campaignlegalcenter.org

Jeff Zalesin
Law Clerk*
jzalesin@campaignlegalcenter.org
*Not admitted to the practice of law

Campaign Legal Center
1101 14th St. NW, Suite 400
Washington, DC 20005
(202) 736-2200

Exhibit A

UNITED STATES DEPARTMENT OF COMMERCE
Economics and Statistics Administration
U.S. Census Bureau
Washington, DC 20233-0001

DSSD 2020 POST-ENUMERATION SURVEY MEMORANDUM SERIES #A-

MEMORANDUM FOR   Patrick J. Cantwell
                               Chief, Decennial Statistical Studies Division

From: Timothy Kennel
                               Assistant Division Chief for Statistical Methods

Subject: Recommendation to Not Disseminate Sub-State Estimates of
                               Coverage from the Post-Enumeration Survey

This memorandum documents the motivation and some implications of not disseminating sub-state estimates from the Post-Enumeration Survey.

For any further information, contact Timothy Kennel at 301.763.6795 or Timothy.L.Kennel@census.gov.

Attachments

1. Recommendation

For the 2010 Census, net coverage estimates of persons and components of coverage estimates for persons were produced for counties and places. To account for synthetic error in the net coverage estimates for these small areas, the mean squared error of county and place estimates was estimated. Given declining utility and interest in the sub-state estimates as well as operational and technical risks associated with producing sub-state estimates, we propose to drop sub-state net and component coverage estimates of persons and housing units form the 2020 PES reports. Our recommendation is to *not publish* the following sub-state estimates in official reports:

- Point and variance estimates of components of coverage for places
- Point and variance estimates of components of coverage for counties
- Point and mean squared error estimates of net coverage for places
- Point and mean squared error estimates of net coverage for counties
- Mean squared error estimates for net coverage of states

Instead of publishing mean squared error estimates for net coverage of states, we recommend producing estimates of sampling variance. This reduction in scope will reduce the cost and complexity of the PES-Estimation system.

Note that the categories for correct enumerations are not dealt with in this memo, even though correct enumerations for persons were differentiated based on geographic level. Furthermore, we recommend producing all of the sub-state estimates described for internal research.

2. Background
2.1. Recent History of County and Place Undercount Estimates

Original plans for the post-enumeration survey in Census 2000 included producing dual system estimates for all blocks in the United States. Because those plans included using the post-enumeration survey to adjust the census counts, we anticipated considerable interest in sub-state coverage. With the landmark Supreme Court decision against sampling in the census that ruled out the planned coverage adjustment of Census 2000 counts for apportionment, much of the interest in and need for place, county and even block group and block coverage estimates diminished. The demand for dual system estimates was further diminished as the U.S. Census Bureau has demonstrated no intention to adjust the census counts for other purposes. Nevertheless, synthetic sub-state net coverage estimates were disseminated as part of Census 2000.

For the 2010 Census, sub-state estimates of net coverage and components of coverage were also disseminated. As an innovation for the 2010 Census Coverage Measurement, the mean squared

errors of place and county estimates of net coverage were estimated and disseminated. This included the contribution to mean squared error of synthetic error in the estimates.

In both the 2000 and 2010 post-enumeration surveys, the synthetic place and county estimates of net coverage were based on models that effectively predicted net coverage from summary characteristics of the counties and places, such as the age-race-sex composition of their census counts. Thus, these estimates were best interpreted as our prediction of how census coverage for counties and places likely varied according to area characteristics, rather than providing direct information on the coverage of each individual county or place.

2.2. Background on Mean Squared Errors

Estimating mean squared errors is important for sub-state estimates because there can be considerable bias for synthetic estimates of small areas, such as counties and places. Yet, the calculation of mean squared errors is computationally challenging, requires knowledge of advanced statistics and statistical programming, and is based on strong and dubious assumptions.

The process for estimating mean squared errors for the 2010 CCM involved creating direct post-stratified estimates of net coverage for small areas that were taken as unbiased, and then modeling the variances of the differences between these estimates and the production sysnthetic estimates. From these results, estimates of expected squared bias could be derived as a variance component. This variance modeling used a Bayesian approach via Markov Chain Monte Carlo (MCMC) algorithms as well as smoothing of the variance estimates through generalized variance functions with random effects. The technical expertise was so demanding for these calculations that mathematical statisticians developed and ran the production software with input from a Senior Mathematical Statistician at the U.S. Census Bureau. The complexity of the technical requirements and the high programming skills needed make the calculation of mean squared error one of the highest risk activities of the PES-Estimation system.

Furthermore, the theoretical foundation of the mean squared error estimates rests on the assumed unbiasedness of the direct post-stratified estimates of net coverage for the small areas. Due to limited sample within small areas, however, the post-stratification used is rather simple, so considerable heterogeneity bias may exist in the post-stratified direct estimates. Unfortunately the amount of heterogeneity bias is not knowable, nor are there benchmarks for PES point and mean squared error estimates. Additionally, since the sample size is rather small for many counties and places, the sampling variances of the direct coverage estimates for most counties and places are large and the corresponding variance estimates are unstable (based on few degrees of freedom). In conclusion, estimating mean squared errors for synthetic sub-state estimates of net coverage relies on dubious assumptions, complex statistical methods, and the expenditure of significant resources.

3. Justification for Dropping Sub-State Coverage Estimates

Dropping sub-state estimates from the 2020 PES reports will result in obvious changes to public-facing data products as well as to PES systems and budgets. This section discusses the main reasons to drop sub-state estimates from the 2020 PES reports. The primary reasons to not disseminate mean squared error estimates and sub-state coverage estimates include a perceived lack of demand for them, a reduction in the risk of schedule delays, and a reduction in the risk of data errors.

We do not see a demand for place and county coverage estimates. Because there are no plans to use the post-enumeration survey to adjust the census, there is very little interest in sub-state estimates. We responded to no more than two requests, questions, and comments on the sub-state coverage estimates after the release of the 2010 Census Coverage Measurement reports. The strong synthetic assumption required for the place and county estimates decreases their utility because they are not based solely on data specific to the individual areas. Furthermore, because mean squared error estimates were so large, there were no significant differences between the coverage estimates for the various counties and places.

Dropping the sub-state estimates from the 2020 PES external reports will reduce the risk of schedule delays. Because there are thousands of sub-state estimates, it takes considerable time to review the estimates to make sure that they make sense. Given that we only have three weeks in the current schedule for reviewing all estimates, there is considerable risk of not meeting deadlines to complete reviews for sub-state estimates. Furthermore, given the short production timeframe of two weeks and the complexity and computational demands involved in producing mean squared error estimates, there is considerable uncertainty about the ability to produce the sub-state estimates and mean squared error estimates in the scheduled time frame.

The statistical and software development of the mean squared error estimates was the most technically challenging and complex aspect of the 2010 Census Coverage Measurement (CCM) estimation system. Not publishing mean squared error estimates will reduce the risk of disseminating data errors. Furthermore, because no IT area at the Census Bureau has experience with the mean squared error estimation methods required by the PES, the risk of bugs in the software developed to calculate mean squared errors is especially high.

Not publishing county and place estimates of coverage should mitigate the risk of disseminating the PES reports late and reduce the risk of publishing errors. Thus, there are substantial operational advantages to not producing county and place estimates of coverage.

4. Disadvantages

Despite the operational advantages to dropping the sub-state estimates, including their mean squared error estimates, there are some undesirable consequences of not disseminating these estimates to the public. Dropping sub-state estimates will result in less information for data users. Relying on design-based variance estimates for the state coverage estimates will likely lead to some underestimation of their error variances.

Obviously, dropping sub-state estimates will mean data users will not have access to county- and place-level tables showing estimates of net coverage and components of coverage. In the 2010 CCM reports, county and place estimates were at the end of reports in an appendix and not part of the core estimates in the main report body. There is very little evidence that these sub-state estimates were used inside or outside the Census Bureau. The lack of statistically significant differences in coverage for states, counties, and places, as a result of large mean squared error estimates, may have decreased the interest in the estimates. Since we plan to produce the sub-state estimates for internal Census consumption, we could disseminate some sub-state estimates, if requested.

Estimates of the measures of error for synthetic totals that do not account for synthetic error will tend to underestimate the mean squared error. Since the bias can be substantial for sub-state synthetic estimates, it is a good practice to publish mean squared errors for synthetic estimates of small areas. For large areas, the variance may be close to the mean squared error, making it less less important to estimate a bias term for large areas. In 2000, only estimates of variance were reported; in 2010, state, place, and county uncertainty estimates for net coverage included a squared bias term. Because the squared bias was forced to be positive in 2010, there is reason to believe that the mean squared error estimates from 2010 may have been overestimated. In fact, the estimated mean squared errors were so large for sub-state estimates that none of the net coverage estimates were statistically different from each other. In 2020, we plan to use successive difference replication to measure the variance with respect to the sample design, rather than to an underlying superpopulation model. This will represent a change in the estimation of uncertainty from previous years. The exact impact of this on the coverage estimates will need further research. As part of developing models for synthetic estimates, we will consider methods to reduce the bias of synthetic estimates, such as including state-level indicators as fixed effects in the models.

There are disadvantages associated with not publishing sub-state coverage estimates and mean squared errors to the public. However, the negative impacts of not producing county and place estimates are mitigated by an anticipated lack of interest in sub-state coverage estimates and the dominance of the variance in the mean squared error for large areas.

5. Recommendation

The CMDE IPT recommends not including coverage estimates for counties and places in public-facing documents. This means dropping tables for the county and place net coverage estimates as well as components of coverage for individual places and counties. Furthermore, we also recommend producing successive difference replication variance estimates for net coverage estimates and reducing the bias of synthetic estimates through enhancements to the modeling. We recommend the PES-Estimation system produce the county, place, and mean squared error estimates as alternative estimates that will not be published in the main 2020 PES reports.

To research estimating mean squared errors for the 2030 PES, we recommend that the variables needed for mean squared error estimation and sub-state estimates be collected and retained on the sample and estimation files.

Exhibit B

United States Census 2020

Governance ▸ 2020 Project Portfolio: 2020 Census Portfolio Management Governing Board Meeting - December 13, 2017

Meeting Date	2020 Census Portfolio Management Governing Board Meeting - December 13, 2017
Agenda	1. Status Update: 2020 Census Program Update, Deb Stempowski (15 minutes), 1-1:15pm 2. Status Update: Architecture Design Document, Quyen Nguyen (20 minutes), 1:15-1:35pm 3. Status Update: Transition Plan, Quyen Nguyen (20 minutes), 1:35-1:55pm 4. Decision Topic: GQAC Study Plan, Diane Barrett (15 minutes), 1:55-2:10pm 5. Decision Topic: 2020 Post-Enumeration Survey, Tim Kennel (15 minutes), 2:10-2:25pm
Notes	INTERNAL CENSUS BUREAU USE ONLY – PREDECISIONAL 1. 2020 Census Program Update •Recruiting for 2018 End-To-End Test. We are little behind the recruiting goal for peak operations. Currently around 2700 applicants, 1118 qualified with a target for 5300 recruited and 1049 deployed during peak operations. Additional efforts have been added to increase the number of applicants. A media campaign is planned for January. •LUCA now covers over 87 percent of the population. LUCA registration ends December 15 except for disaster affected areas which ends January 31, 2018. The extension for disaster areas is publicized on the website. Other governments may request an extension to January 31 but that is not publicized on the website. 33 states including DC participating. •The 2020 Integrated Master Schedule will be baselined on December 15. 2. Architecture Design Document and Transition Plan The Architecture document is the IT solution to support 2020. This documentation, similar to the 2020 Operational Plan, is the 2020 IT Road Map. Key aspects of the plan are scalability and security. The output of the Transition Plan is the Integrated Implementation Plan (IIP). Action: Quyen will ensure that the 2018 End-to-End Test placemat is accurate and will provide it to executives for the briefing books. Action: Quyen will add a date to the placemat for version control purposes until it is baselined. Then he will add the baseline date. Action: Quyen will schedule to brief at IPT Leads to get feedback on the placement. Action: Quyen will review charts and colors for clarity and revise, as needed. Action: Quyen will meet with Melissa Therrin, the DCMD lead for disaster, to discuss the IT Disaster Recovery plan. 3. GQAC Study Plan Statement from GQAC SME's added December 18, 2017. "At the December 13th PMGB GQAC presentation, we inadvertently mentioned that GQAC updates would not be made to the MAF (includes updating 501's to 502's). At the time we were uncertain who will create the GQAC ADDUP (Address Update File) thus concluded that an ADDUP would not be created to update the MAF. However, Dave Tarr will make the address update data available to DITD-AUSB (Thammarak). Thammarak will create the ADDUP. Thus all eligible GQAC updates will be updated to the MAF." Decision: Approved to move forward. 4. 2020 Post-Enumeration Survey Recommendation not to Disseminate Sub-State Coverage Estimates. It is planned that the estimates will be produced but not publicly disseminated. Jim noted that due to normal post-Census drops in funding that a new risk is introduced that resources may not be available to produce estimates in FY 22. DSSD will work to ensure the estimates are produced. Decision: Recommendation Approved. Other attendees: Melissa Therrien DCMD, others on the phone.
Invitees	
Attendees	Adam E Bacon - DCCO; Alan Lang --OCIA; Alexa Kennedy Jones-Puthoff - DCMD; Burton Reist-ADDC; Cynthia Davis Hollingsworth--ADDC; David J Peters - ADSD; Deborah M Stempowski - DCMD; Deidre C Hicks - ADDC; Deirdre Dalpiaz Bishop - GEO; Donna M Daily - ACSO; Everett G Whiteley - BUD; Gerell L. Smith - DCCO; James Treat-ADDC; James Christy --FLD; Jill Renee O Brien - DCMD; Karen Battle - POP; Luis J Cano -- DCEO; Maryann M Chapin - DCMD; Michael T Thieme --CAD; Patrick J Cantwell - DSSD; Phani-Kumar Atri Kalluri - DITD; Robert W Colosi - DSSD; Robin A Pennington- ADDC; Sheila M Proudfoot --DCMD; Sherri J Norris - DCCO; Tasha R Boone - DCMD; Victoria Velkoff - ACSO; Carlos LaCosta --CAO; Tim Trainor - GEO; Robin Bachman PCO FED, Jennifer Reichert ACSOD, Diane Barrett DSSD, Denise Smith DCMD, Mary Frances Zelenak DSSD, Dora Durante, DCMD, Deborah Fenstermaker DSSD, Scott Fifield DCMD, Julianne Zamora DSSD, Timothy Kennel DSSD, Quyen Nguyen CTO
Date	12/13/2017
Chair	Jim Treat
Location	4H485
eVoting	No
Attachments	2018 E2E CT placemat_v3.7_CR 623_Draftsn.pptx 2018 E2E CT placemat_v3.7_CR 623_Draftsnpdf.pdf 2018 E2ECT GQAC Study Plan Final Draft _v1.3_121117.docx 2018 E2ECT GQAC Study Plan Final Draft _v1.3_121117pdf.pdf 2020 Census Architecture Document Version 3.6 12_06_2017.docx 2020 Census Architecture Document Version 3.6 12_06_2017snpdf.pdf 2020 Census Architecture Transition Plan 3.6- 12_07_2017sn.docx 2020 Census Architecture Transition Plan 3.6- 12_07_2017snpdf.pdf PMGB Recommendation not to disseminate sub-state coverage estimates.pdf PMGB Recommendation not to disseminate sub-state coverage estimates.pptx Recommendation to Not Disseminate Sub-State Estimates of Coverage from the 2020 PES d3.docx Recommendation to Not Disseminate Sub-State Estimates of Coverage from the 2020 PES d3.pdf

Created at 12/11/2017 9:01 AM by Gerell L Smith (CENSUS/ADDC FED)
Last modified at 3/30/2018 9:00 AM by Sherri J Norris (CENSUS/ADDC FED)

[Close]

Action Items

☐ 📎	Type	Title	☐ Assigned To	Status	Due Date	% Complete	Predecessors

There are no items to show in this view of the "PMGB Action Items" list. To add a new item, click "New".