



RISK FACTORS AFFECTING THE FAIRNESS AND ACCURACY OF THE 2020 CENSUS

**Statement of
Robert L. Santos
Vice President and Chief Methodologist, Urban Institute**

**before the
Committee on Oversight and Reform,
United States House of Representatives**

ENSURING THE 2020 CENSUS COUNT IS COMPLETE AND ACCURATE

December 3, 2020

* The views expressed are my own and should not be attributed to the Urban Institute, its trustees, or its funders.

Chairwoman Maloney, Ranking Member Comer, and members of the Committee, I am pleased to address the Committee on the accuracy and fairness of the 2020 Census. I am the chief methodologist for the Urban Institute. My remarks today are my own and should not be attributed to the Urban Institute, its trustees, or its funders.

I understand your interest in the data anomalies uncovered by Census Bureau career staff and reported by Director Dillingham in his November 19 press release. Before providing my commentary on that topic, I submit to you that the drivers of accuracy and completeness transcend the discovery of anomalies in the post-enumeration processing phase of the 2020 Census. The story of census accuracy is much deeper and complex than the current chapter, and its roots go back for years. So, I would like to review some of the risks that the Census Bureau has faced in the decennial survey operation because they provide clues as to the questions that need to be asked and the data that need to be requested to better understand the strengths and limitations of the counts that will inevitably be produced.

I'll begin with some research results that Diana Elliott, Steven Martin, and I conducted last year that considered possible outcomes of a 2020 Census. (This was obviously before the COVID-19 pandemic.) We used three "risk scenarios" based mostly on previous Census Bureau research to simulate what might occur in 2020. The lowest-risk scenario simply took the purported excellent performance of the 2010 Census, which was extremely accurate—within one-hundredth of 1 percent of an independent total population estimate. We superimposed those parameters onto a 2020 projection of the US population and observed a net undercount of 0.3 percent. The increase was because our wonderful nation has become much more diverse over the past 10 years. The percentages of the population that are African American, Latinx, or Asian have increased. And these are some of the folks that are historically harder to count. In fact, although the 2010 Census total population was spot-on accurate for the total US population, the accuracy came at the expense of *fairness*. White people were over-counted by 0.84 percent, and this conveniently made up for the net undercounts of people of color. Non-Hispanic Black people had a net undercount of 0.84 percent, Hispanics of 1.54 percent, and Native Americans of about 5 percent.) It is unfair to over-count one sector of society while undercounting another to achieve overall accuracy. Doing so only reinforces inequities in political representation, federal funding to local communities, and economic and public health opportunities. And these inaccuracies are baked in for the ten-year period between censuses.

Why does this matter now? Well, fast-forward to 2020 and the COVID-19 pandemic. We know that the economic downturn affected people of color and communities of color far more than others. The accompanying economic downturn—notwithstanding the stock market performance—eliminated millions of lower-wage jobs held by people of color, which in turn exacerbated their health and economic conditions. We see high racial-ethnic disparities in rates of job loss, hunger, risk of eviction and foreclosure, and of health and well-being in communities of color. Since March, millions of families in communities of color have and continue to suffer financially, physically, and mentally. It is safe to say that since the onset of the epidemic, completing a 2020 Census form was nowhere near the top of their priority list. Their daily lives revolved around meeting basic needs such as survival, shelter, and food.

That brings us to a basic quality indicator of the 2020 Census, the self-response rate. Self-response represents a household's completion and submission of their own census form. Self-responses are

considered the highest quality data in a census. And we know from past research that the lower the self-response rate is, the higher the risk of a net undercount for a given subpopulation group. Now, the Census Bureau reports a national self-response rate of 67 percent for 2020, which is even higher than the 66.5 percent achieved in the 2010 Census. But it should be no surprise that self-response rates varied across states and within cities. In fact, self-response rates were drastically lower—50–60 percent, and sometimes less—for inner-city, mostly poor neighborhoods with high concentrations of Latinx, Blacks, Native Americans, Asians, and other hard-to-count populations. In contrast, suburban areas showed ultra-high rates of self-response to the 2020 Census: 70–80 percent, and sometimes higher. Suburban families tended to retain their jobs, had great internet service, and worked from home. Guess who is likely to be counted accurately or overcounted, again? And guess who is most likely to be undercounted, again? I believe the die is cast and that miscounting will be significant and of a much larger magnitude than in previous censuses in our lifetime. All this, and differential self-response represents only one risk to accuracy facing the 2020 Census.

Other risks to accuracy faced by the 2020 Census include the following:

- *The citizenship question fracas and anti-immigrant policy climate:* Litigation on the inclusion of the citizenship question and possible exclusion of undocumented immigrants generated fear in the immigrant population; the Census Bureau’s own research suggested that these would have a chilling effect on census participation by immigrant families.
- *2020 March mobility:* In reaction to the widespread COVID-19 lockdown in March, millions of people moved just before Census Day, April 1. This migration included college students returning home, along with many others such as grandparents and adult children choosing to ride it out with family.
- *Scheduling disruptions to the 2020 Census field operations:* Pauses, starts, and stops in the nonresponse follow-up field effort wreaked havoc on hiring, training, and management of field staff and field protocols to resolve nonparticipating households.
- *Natural disasters:* Serious wildfires, tropical storms, and even a Midwest derecho caused local delays in census field operations.
- *Condensed processing period:* In an attempt to adhere to the statutory deadline of December 31 for apportionment counts, the Census Bureau has reorganized and shortened its processing by half to ten weeks, putting enormous pressure on career staff working 7 days a week and up to 15 hours in a day.

Despite this plethora of risks, the Census Bureau touted a 99.98 percent accounting of all households and addresses in the 2020 Census. But this figure should be taken with a grain of salt. “Accounting” simply means that a nonparticipating household or address was brought to final disposition according to field protocol. This could happen through self-responses, enumeration of the household by field staff, whole-person imputations via administrative records, or proxy interviews where only partial enumeration information is obtained by someone not living in the target household. While self-responding and interviewed households provide the highest quality, all other “accountings” have lower data quality and pose a risk to accuracy. This is especially true for proxy interviews, where enumerators take whatever information they can get from a neighbor or landlord. Sometimes the only information recorded was a guess of the number of persons living in the recalcitrant household. As such,

it is important to understand the components of the 99.98 percent accounting to get a sense of the quality of the counts. That brings us to transparency in the 2020 Census.

It is imperative that the Census Bureau be transparent by providing to researchers and the public the data needed to assess the quality of the counts. In fact, the American Statistical Association and the Census Bureau's own Scientific Advisory Committee have independently written reports recommending the release of specific quality indicators and the need for transparency. For instance, a simple statistic that could be reported immediately is the number of proxy interviews conducted by field enumerators, broken down by the proportion of full household enumerations by proxy and the proportion of proxies for which only a count of persons was obtained.

As for the anomalies disclosed but not described by Director Dillingham, there should be no surprise that such problems were encountered given all the risks facing the 2020 Census. I understand from media reporting that the anomalies are related to group quarters counts. This would be consistent with difficulties in sorting out the residency of college students, many of whom attended schools out of state but returned home before April 1 and remained there, possibly to this day. Disentangling which students should be counted at home and which should be counted at their school location could be quite challenging and potentially alter the state counts that make up the apportionment counts due on December 31. But this is an educated guess, as more information is needed.

And this current challenge does not take into consideration the likely over- and under-counting of racial ethnic and other subpopulations that will almost surely emerge as a result of combining self-reports and other enumerations of lesser quality, as I mentioned earlier.

Unfortunately, we know much less about the performance and quality of the census than we could because this census is the first with a totally digitized field operation. The public deserves to know the accuracy and quality of the census counts. Researchers outside the Census Bureau stand ready to help. We only lack the data. I recommend that the Census Bureau adopt the recommendations in the reports of the American Statistical Association and the Bureau's Scientific Advisory Committee.

In closing, I commend the career staff of the Census Bureau for their dedication, their commitment to scientific integrity, and their oath to uphold the Constitution. The scientific community holds these scientists and those of the other federal statistical agencies in high esteem. They deserve to be allowed to do their jobs without compromising their ability to execute their due diligence, so they can achieve the highest degree of quality in the 2020 counts and can divulge the strengths and inevitable limitations of those counts.

I thank the committee for inviting me to testify, and I look forward to your questions.