

NWX-US DEPT OF COMMERCE

October 15, 2021

10:00 am ET

Coordinator: Welcome and thank you for standing by. At this time all participants are in listen-only mode until the question-and-answer session of today's conference. At that time you may press star 1 on your phone to ask a question. I would like to inform all parties that today's conference is being recorded. If you have any objections, you may disconnect at this time. I would now like to turn the conference over to Michael Cook. Thank you. You may begin.

Michael Cook: Good morning, everyone, and thank you for joining us today. I'm Michael Cook, Chief of the Public Information Office at the U.S. Census Bureau.

This is the second status briefing of the National Urban League plaintiffs in the case of the National Urban League vs. Gina Raimondo.

The Census Bureau has agreed in a joint stipulation in settling this lawsuit to brief plaintiffs every two months to allow an opportunity for questions and answers regarding the status of forthcoming data quality metrics and assessment of released data quality metrics. A final briefing will occur after the release of the Post-Enumeration Survey final results.

Today you will hear from Michael Bentley, Assistant Division Chief for the Census Statistical Support and the Census Bureau's Decennial Fiscal Studies Division, as well as Deborah Stempowski, Assistant Director for the Census Bureau's Decennial Programs.

Today's briefing is open to the plaintiffs, their attorneys and the public albeit only the plaintiffs and their attorneys will be able to ask questions after today's presentations.

Media members of the public, you can find more information in our online newsroom at [census.gov](https://www.census.gov). You can access today's presentation as well in this briefing's electronic press kit. Let's get right into the presentation. I'll turn it over to our first speaker.

Deborah Stempowski: Great. And thanks, Michael. It's great to be here with you all this morning. I'm Deb Stempowski and I'm going to get things rolling before I hand it over to Mike Bentley.

So I think we're ready for our next slide. I'm going to focus these first couple slides on our accomplishments and some updates since our last briefing.

So as you mostly - as you know, the Census Bureau provides the 50 states, the District of Columbia and Puerto Rico with population counts to use in their redrawing of congressional and state legislative districts. And that's a process called redistricting.

The Census Bureau released these data on our FTP site publicly on August 12, 2021. And it gave the country a first look at populations for small areas. To help tell that story of our nation, we also provided data profiles, data

visualizations and America count stories that highlighted how our nation has changed over the last 10 years.

The Census Bureau also released that same data in easier to use formats on September 16. I encourage you to visit the Web site to take advantage of the tools and stories there. It's a rich library of information.

Next slide. So we also released several rounds of operational quality metrics, which Mike will discuss in more detail in a few minutes. Since our last briefing, we had a pair of releases under the guise of our Release 3.

The August 18 release contained a selection of sub-state, county and tract level summary statistics based on specific metrics that we released on April 26 for the nation and at the state level.

On August 25, a wealth of information on item non-response rate by data collection operation for the nation's 50 states, District of Columbia and Puerto Rico were also released. So I'm going to stop there. Mike's going to dive into that topic in more detail.

But looking ahead to data products after redistricting, we've begun a concerted effort now to gather feedback from stakeholders on the proposed tables in the 2020 Census demographic profile, the demographic and housing characteristics file and the detailed demographic and housing characteristics file.

In September, we released an updated 2020 Census Data Product Planning Crosswalk. This crosswalk allows users to examine our current proposal for various levels of detail. The crosswalk shows the specific content and the

lowest level of geography we're considering for each table and how those proposals compare with 2010.

These plans, however, are not final. They're based on collecting feedback from data users as part of our process. You can see on the slide here the email address where we invite you to provide feedback through October 22. In particular we're seeking feedback on how you would use the data at the various levels of geography in the file.

We frequently refer to that as use cases. It's not that you need the data. We know that. We are trying to understand how you use it. So we appreciate that folks need that data. We just want a better understanding of the actual uses in your organizations.

All right. We can go to the next slide, (Kathleen). And so let me touch a little bit on the status of our Post-Enumeration Survey. Work continues for data collection for the PES with just a handful of matching and field operations left. The last operations run into early calendar year 2022 and the first results from the PES are planned for release in the first quarter of calendar '22, with additional results to be released in the summer of 2022.

And you can see here that the first release will have national results coverage estimates for people, including by race and Hispanic origin. And the second release will include state level coverage estimates for housing units, data estimates for people and coverage estimates for Puerto Rico.

So I think now I'm going to turn it over to Mike. We can go to the next slide. And he's going to give you a deep dive into our operational quality metrics. Mike.

Michael Bentley: Thank you, Deb. Good morning, everyone. So as Deb said, today I'm going to give a brief but in-depth overview of our 2020 census operational quality metrics.

As Deb said, we've released three batches of metrics. First, the metrics coinciding with the apportionment results in April focused on final address resolution of census addresses, including housing units and group quarters, and also dove into housing unit status results from non-response follow-up.

Next on May 28 we released indicators focusing on average household size by census operation as well as more information on housing unit status by operation.

Lastly we had a pair of releases under the guise of Release 3. First a selection of sub-state county and tract level summary statistics of selected Release 1 metrics. In the interest of time, I won't be going into further details on those numbers today. And then finally, a wealth of information on item non-response rates by data collection operation.

Next slide, please. For each of the quality metrics, we release data for the nation but also for each individual states plus D.C. and Puerto Rico. This is a first for many of these results. For some of these statistics, it's the first time we've ever produced them before in these specific ways and certainly not this soon after the census.

For the most part in the past, such as in the 2010 or 2000 Censuses, it's usually been as part of our operational assessments program and very few of them below the national level.

Because the pandemic and other factors created concerns about the 2020 Census' quality, this is an important part of our continuing transparency effort following along with the goal of providing what we know and when we know it.

And for comparison, with the exception of the sub-state summary metrics, we've also released corresponding apples to apples data from the 2010 Census.

Next slide, please. We know that no census is perfect, but we use well-established and proven methods to get as close as we can. Remember, we have to get everyone in the country to cooperate and that's no small feat.

We also check our work and have numerous quality checks built into collecting the data. And we also conducted a very comprehensive review during data processing.

All that said, there just isn't one number that we can use to describe the quality of the census. The census is too complex and has data too layered for that.

So I want to talk a little bit about things to consider when we've been making comparisons between the 2020 and 2010 Censuses or from one state to another.

The first point is the differences across geographies from one state to another where by sub-population are expected. Further each state is unique and things can change over time. Moreover differences may be the result of operational changes like the differences in the way we collect the data during the census.

For instance we modernized the 2020 Census in various ways, including adding an Internet response option and using high quality administrative records to resolve some non-responding housing units.

Differences may also be the result of environmental changes, such as changes in the population or population growth or movement or trends and respondent behavior or various acts of nature, like, of course, the COVID-19 pandemic or wildfires or hurricanes. Different doesn't necessarily mean better or worse. It can just mean different.

So next I'm going to summarize some details on what is included within the quality metrics and then show a few highlights. So Release 1 primarily focused on the final resolution of census addresses by data collection operation.

First the operational metrics showed the percent of occupied, vacant and deleted addresses by how that status was determined. Occupied and vacant I think are obvious.

But I believe sort of census jargon, which includes cases where housing that doesn't exist or is a duplicate of another or other cases where ultimately it's not included in the final census tabulations. By operation, this could be from self-response, from non-response follow-up or group quarters or other smaller census operations or was left unresolved.

Then we look at the self-response resolution by mode. In the 2020 Census, we offered three ways for households respond on their own, online, by returning their paper form in the mail or by using their telephone through our Census Questionnaire Assistance Service.

We also looked closely at addresses resolved in non-response follow-up and by respondent type. By respondent type I mean this can be from a household member that lives there at the house, a proxy respondent such as a neighbor or building manager or landlord or was determined for the use of administrative reference.

And lastly we provided a look at the percentage of what we call “pop count only” cases among the occupied households from non-response follow-up.

Next slide, please. This is the bottom line upfront, a few highlights from the 2020 metrics are shown here. Overall of all addresses in the census, over 65% were resolved from a self-response.

The majority of households, almost 80% that self-responded did so online in our Internet instrument. Also among occupied households that were resolved in non-response, about 55% were enumerated by household member, about 26% by a proxy such as a neighbor or landlord and about 18% using our administrative records.

And at the end of all data collection, only about 0.23% of addresses were left unresolved. Another 0.71% were left unresolved as a result of unduplication processes. And combined the unresolved rate was 0.93%. And these are the cases that went to our count imputation.

And next, this is the first of a few slides with a lot of numbers on them. I do want to note that all of these can be found within the quality metrics available on our Census Web page.

Here we are showing the percentage of census addresses resolved by self-response in each state in 2020 compared to the 2010 Census. Self-response

metrics are important because self-responses generally provide the highest quality response data, and they also comprise the largest portion of the census count.

Every state had a higher self-response compared to 2010. The national rate was about 65.3% after post-collection processing whereas in 2010, it was 61.1%. Minnesota led all states with 73.6% followed by Wisconsin and Washington State.

Nevada had the biggest improvement from 2010 at almost 10 percentage points with New Hampshire following next and Washington close behind.

The states with the larger gains tended to be those where we were able to move a chunk of blocks from other enumeration areas, such as update leave or update enumerate into the self-response type of enumeration area where we mailed out our well-tested contact strategy of letters and follow-up reminders to encourage responses.

On the next slide, as we've mentioned frequently in our updates, almost 80% of self-response came in via the Internet in 2020. In every state this is over 60%. The remainder was about 18% paper and about 2% by calling our phone lines.

Some of those differences from state to state were in part by design. Whether there are more rural areas that are not in our self-response mail out area or areas where we sent a paper questionnaire in the first mailing to places that we expected would be less likely to go online and we wanted to make the census easier for them. That's what you might have heard us refer to as our Internet choice strategy.

Moving on to the next slide, I want to talk a little bit more about unresolved addresses in the quality metrics. These are the cases that went to the last resort process of count imputation in order to determine status and fill in the number of people living at an address using our established statistical methods.

While we stand fully behind those procedures to fill in those numbers, count imputation is ultimately the last thing we want to be doing after our self-responses or non-response follow up. Overall for the 2020 census, as I mentioned, the count imputation rate was less than 1%. And here you can see how that varies slightly from state to state.

Next I'm going to talk a little bit about our quality metrics from the second release in May. Release 2 included sort of a hodgepodge of different metrics.

First we reported on the average occupied household size by data collection operation. We also looked closer to household sizes by showing the percent of households with just one person in them or with two people living there. And we provided details on housing unit size status by operation, including the distribution of how those occupied and vacant units were enumerated.

Some of those highlights are shown here on the next slide. We reported that the average household size is about 2.4 people per household for both households that self-responded and those resolved in NRFU.

Among households that self-responded, about 26% of occupied households had just one person living there and about 35% had two people. In comparison for households in non-response follow-up, about 33% had one person and 27% had two people living there on Census Day.

Additionally, about 8% of all occupied or vacant housing units were counted by self-response without a pre-assigned Census ID that was sent in the mail. This is sometimes referred to as our non-ID response option. And about 77% of all occupied housing units were enumerated by self-response.

Virtually, as you might expect, most vacant units fell into our non-response follow-up operation. In the interest of time, I won't be going into greater detail on Release 2. But generally we did not find any results that were very surprising or raised any flags about potential data quality issues.

So finally I want to move on from our last release of quality metrics from August on our item non-response rates. As this release was timed to coincide closely with the first release of the redistricting data, we reported item non-response for those data items, including population count, age or date of birth and Hispanic origin and race.

The rates were split out by different operations, including self-response and mode or non-response follow-up and NRFU respondent type and for other operations as well as group quarters.

And we also provided statistics on the rate of how often we received only population count and no other information for each of these operations.

I do want to note that as part of our evaluation and assessment program, we are set to release the Non-Response and Imputation Assessment Report next summer. This report will include comprehensive results on item non-response with metrics for all data items, including those being discussed today plus household tenure, sex and relationships as well as detailed imputation and substitution rates.

So I want to talk about some highlights from the item non-response rates. First I do want to be clear that the numbers we are discussing today from the quality metrics for the final item non-response rates after all data cleanup and processing to reconcile things such as multiple responses and other post-processing.

Some of you may have seen numbers that were put out earlier this year as part of a FOIA request. The bottom line is that those numbers were very preliminary and somewhat premature overall and not intended to represent the final picture on item non-response in the 2020 Census.

Overall the item non-response rates for most questions were higher in the 2020 Census compared to 2010. But they were still in the single digits overall, including less than 1% for a population count.

For most questions item non-response is lowest for households with the self-respondents. This is especially true for our Internet and phone respondents likely in part because the online questionnaire, which we also use to capture phone responses and remind people to provide a response if they tried to skip a question.

Among households enumerated in NRFU, rates were highest when a proxy responded and provided the data. And this is not really that surprising. A neighbor, for instance, can usually say with confidence if a home or address is occupied or vacant and maybe perhaps exactly how many people live there. But answers for the demographics such as age, sex, race, Hispanic origin or whether owned or rented will tend to be of lower quality.

And finally another highlight that I want to note is that group quarters tended to have relatively high item non-response across the board.

So next, this graphic just compares that at a national level that item non-response rates by ...between the 2010 and 2020 Censuses for self-responding households.

If you focus on the blue bars here, these represent 2020. You can see that the item non-response was actually lower in 2020 for three of the four items, only age being a little higher for the 2020 self-respondents compared to 2010.

In 2010, of course, we did not have an online response option. So almost all of our self-responses were from returning a paper form in the mail. With paper, we're happy to have those responses, of course, but there are no real-time checks we can do to remind respondents when they skip something to please provide a response. And our analysis shows that it worked well to reduce that level of missing data.

So next we have another state by state table. Using again the race question as an example showing the 2010 versus 2020 numbers. A couple of things I want to point out.

First there's clearly variation from state to state. Some show an increased in item non-responses, some a decrease. Some were lower than others, with North Dakota, Minnesota and Wisconsin leading the way with the lowest item non-response for race among those self-response households.

In general we found from past research leading up to 2020 that states with higher Hispanic or Latino populations tended to have higher item non-response to the race question as those respondents are a little bit less likely to identify within the existing race categories.

So there also tended to be the states that saw the largest decrease between 2010 and 2020. We presume in part because of enhancements such as the online functionality.

Next, here's a graph showing item non-response rates by race between the two censuses for non-response follow-up households. Again, the blue bar showed 2020 and the orange showed the 2010 Census. Here, unlike with self-response other than population count, which was required for the NRFU responses, we saw an increase in item non-response for age, Hispanic origin and for race.

Moving to this next slide with state-by-state rates for the race question, this increase was true in every single state for those households enumerated in non-response follow-up. Arkansas, Maine and West Virginia were the three states with the smallest increase compared to 2010.

Also this summarizes the item non-response for all non-response follow-up households. But as I mentioned before, the complete set of data available on our Web site also breaks down by cases completed with a household member, with a neighbor or other proxy and with those resolved with administrative records.

So moving on. Next I want to briefly go over to item non-response rate for people enumerated in group quarters. Again, the blue bar showed the 2020 numbers and orange 2010.

You know this really gets into our transparency mantra where we are trying to be as proactive as possible with sharing important information about the census. Part of that means not just showing the good things that went super well and the item non-response rates for group quarters don't exactly paint a perfect picture.

But as you can see here, not everything was great in the 2010 Census either. Twenty-five percent item non-response to the Hispanic origin question for instance and now in 2020 over 40%. This isn't very surprising either though when you think about it.

Unlike with most of our household data collection where someone tends to know a lot about the people they are living with, for instance, that's not exactly (as true in group) quarters. And the level of information that the group quarter facilities have on their populations varies very widely.

Some of this demographic data we were asking about, they may not even have. And it also varies widely by locality and by state. But we are studying all of this very closely and also looking at ways to enhance our design as we are with all of our operations as we begin the 2030 Census planning.

So moving on to the next slide, at the state level for group quarters while most states had a higher item non-response for race it was a little bit of a mixed bag. Interestingly some states actually showed a decrease in the race item non-response compared to the 2010 Census.

Idaho, for instance, saw a drop of over 30 percentage points. Overall the states with the lowest item non-response for race in group quarters were in Idaho, Montana and Mississippi.

I do also want to mention that the final assessment report that I mentioned that's due next year will also include information on item non-response by group quarters type. This includes correctional facilities for adults, juvenile facilities, nursing facilities and college and military housing.

So before we finish up our briefing today, for those who haven't seen the operational quality metrics, I just want to briefly describe how and where you can find them.

You can move the next slide, please. If you go to [census.gov](https://www.census.gov) and then select 2020 Census under surveys and programs and then select the 2020 Census data quality option, once there you should bookmark it like I have.

I have that page open on almost every single work day. There's a lot of really useful information on the page. But if you select operational quality metrics at the top, you'll drop down to the relevant data quality releases that we've been talking about today.

And then on the next two slides, we see two different ways that you can view the numbers. First are downloadable Excel tables. There are four in total. And then second for the first two releases in April and May, we have a really neat, interactive dashboard that makes it easy to compare your selected metrics of interest.

And with that I am going to turn it back to Deb, who will be summarizing and talking about other ways that we are studying the quality of the 2020 Census.

Deborah Stempowski: Great. Thanks, Mike. We can go the next slide. So just to recap, what have we learned from the operational quality metrics?

We know there's not one single number that can quantify or grade the quality of the census. It's a massive undertaking with many nuances and complex pieces to consider. And it was unique and had a lot of challenges.

That being said, all the metrics that we've studied so far suggest that the quality is on par with the 2010 Census. And as Mike said before, and I think it's worth reiterating again, we will continue to share what we know when we know it and keep this level of transparency in place.

You know, historically, the United States has been a very unique melting pot and we know there's a lot of variation in the statistics and metrics from one place to another. And then although the work on the quality metrics started just about a year ago, this wasn't part of the plan. But we knew we had to change because things changed.

It does position us very well looking ahead to the 2030 Census and how we can make this data available perhaps faster and build it in as part of our processes.

Go to the next slide. So I do briefly want to remind everyone about other ways that we're measuring and analyzing the quality of the census.

We've released reports comparing the census count to the estimates from pop estimates and from our demographic analysis program and that provides a measure of coverage using birth, death and migration data alongside the release of the redistricting data.

We prepared a number of in-depth reports on various characteristics, as I referenced before, and other improvements in terms of how we measure those characteristics for the 2020 Census.

And of course, I discussed earlier we're eagerly awaiting the results of the Post-Enumeration Survey, which will start in quarter one of calendar year '22 and that will provide more information.

We're working with various groups in the scientific and statistical community on independent assessments of the Census. And I'll touch on that in a little more detail on the next slide.

And as we do each decade, although I think this decade our regular assessments and evaluations will garner more attention than ever before. We have that robust portfolio of work going on and we'll share that. That's generally what we do in the two years following the census so work on a number of fronts continues.

You want to go then, (Kathleen), to the next slide. And so here, let me just recap in a little more detail. We've talked at great length what the Census Bureau has been doing in terms of quality. So let's just end with the partnerships that we have with our external groups and what they've been doing to help us assess the quality.

First, way back in February, which seems a long time ago, the JASON Group released their report. That was quick turnaround. We needed something quickly. They provided that and that was on our processes, procedures and metrics. And we shared that with the public.

Just recently the ASA Task Force released their assessment report on national and state data. We thanked them for that work. That was a robust effort that we very much appreciate. But this was a different type of report compared to the JASON report.

The ASA data team had national and state level data available for their review, as we had planned, to support an assessment on apportionment. There

was to be a second report on sub-state analysis but that level of analysis has actually been handed off to our third group, the last bullet on the slide.

We established in May a panel with the Committee on National Statistics that will conduct an evaluation of their own on the quality of the 2020 Census. We've held several rounds of public meetings thus far. I hope you've been able to join us for some of those, and these have covered various topics that are determined by the panel.

The next public meeting in that work will be held in mid-November, and final information about that meeting will be available on their Web site when it's made - when it is available. In total, this panel's work will continue and takes about 24 months. So, we thank our external experts.

It's critical that we invite folks in to help us do the assessment and continue that tradition. And so, that sums up the external group. And then I believe we're at our last slide. So, I think now we're ready to open it up for questions.

Michael Cook: Thank you, Deb, and thank you, Mike. Before we begin taking questions, I'd like to remind everyone that only the plaintiffs and their attorneys will be able to ask questions. For those of you asking questions, please announce your name and who you are representing.

And as a reminder, everyone can visit our Web site, [census.gov](https://www.census.gov), where you can access today's presentation within the briefings electronic press kit. Operator, we will now begin taking questions.

Coordinator: Thank you. We will now begin the question-and-answer session. If you would like to ask a question, please press Star 1 and record your name. If you need to withdraw your question, press Star 2. Again, to ask a question, please press

Star 1. Our first question comes from Thomas Wolf from the Brennan Center for Justice. Your line is open.

Michael Cook: Hi, Tom.

Thomas Wolf: Thanks. Hi. Good to hear from you again. Thanks for everyone for putting together this presentation and talking with us today. I was hoping that we could circle back on the item non-response issue, I think, was flagged. We saw substantial increases in the item non-responses for date of birth, Hispanic origin, and race, in both NRFU and group quarters in Pages 24 and 26 of the slideshow.

Could you talk a little bit more about what you attribute the differences between 2010 and 2020 to? These are very large, as you've noted yourself.

Michael Cook: Thanks for that question. As you can see, we've navigated to that Slide 24. And I'm going to turn it over to, I believe, Mike, to address that line of questioning about our item non-response rates. Mike?

Mike: Sure. Thank you for the question. So, in terms of non-response follow-up, we know obviously that operation had to be delayed a little bit during data collection last summer a bit. And, you know, we were in deep in the middle of the COVID pandemic. We believe that part of the reason is because there was hesitancy with some parts of the community and the public actually to be interacting face to face with the door-to-door census takers.

And so, there might have been some hesitancy to fully cooperate when the census taker was at the door. So, I think that's part of the reason for non-response follow. For group quarters operations,

I don't know if Deb, you wanted to say more, but definitely there's a lot of - there was impacts for our group quarters operation as well, and it might have also played a hand with the comprehensiveness of the data that was collected.

Deb Stempowski: Right. I think I would just add on to that, Mike. We did see - and I'll use the colleges as an example, of a more robust effort on the side of the administrators to use the privacy protections that they provide for their students and to just give us then what we call directory information for those living on college dorms, on the campuses.

And so, that inevitably then drove item non-response and things that weren't what they classified as, you know, essentially just a handful of data items. Of course, we're continuing, Mike referenced that we're going to have one of the assessments out that's on this specific topic next summer.

And so, we continue to dig into both the household side and the group quarter side on this topic through that assessment process as well. Thanks for the question.

Michael Cook: Thanks for that. Operator, do we have our next question?

Coordinator: I'm showing no further questions at this time, but as a reminder, please press Star 1 if you'd like to ask a question.

Michael Cook: And while we wait for anyone to ready themselves or to present their questions, would just like to remind the people that are on the line or listening in, that in the electronic press kit, not only will you find a copy of today's presentation, but you'll also see a copy of the deck with the last presentation that we did, the July 6th briefing. Operator, do we have any questions?

Coordinator: I'm showing no further questions.

Michael Cook: Very well. Just a reminder for anyone listening, that you can sign up for our subscriptions in our online newsroom to receive alerts about Census real news to stay up to date about future briefings and the vast information that is released about our nation's people, places, and economy.

Again, for those media - members of the media, rather, that I invite you to contact us in the Public Information Office at PIO@Census.gov with any questions you might have about today's briefing. And on behalf of Deb Stempowski, and Mike, I'd like to thank everyone. I am Michael Cook. I thank everyone, and this concludes today's briefing. Thank you.

Coordinator: That concludes today's conference. Thank you for participating. You may disconnect at this time.

END