NWX-US DEPT OF COMMERCE

March 10, 2022
10:00 am ET

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

Michael Cook: Good morning, I'm Michael Cook, chief of the Public Information Office at the U.S. Census Bureau. Thank you for joining us for today's news conference and technical presentation. Shortly we'll begin posting the first set of coverage estimates for the 2020 Post-Enumeration Survey, along with additional results for the 2020 Demographic Analysis estimates. These are key indicators for 2020 Census data quality. The results tell us how well we counted the nation in the 2020 Census and if we under- or overcounted certain groups.

In this morning's new conference, you will hear from Census Bureau's director, Robert Santos, who will provide high-level findings from our analysis. Then, Erika Becker-Medina, chief of the Decennial Communications

Coordination Office, who will provide important additional context about the challenges we faced and how we overcame them. We will then have a technical briefing and a deeper dive into the findings from Eric Jensen, senior technical expert for Demographic Analysis, Population Division, and Timothy Kennel, assistant division chief for Statistical Methods in the Decennial Statistical Studies Division. At the end of both presentations, we'll be joined by a few additional subject matter experts. They'll take question from the members of the media, stakeholders and Census partners.

When the Q&A session concludes, we'll then deliver a Spanish language presentation. One final note, before we begin, if you'd like to ask a question today, you'll need to dial into our phone line at 1-888-790-3166 and use the passcode you see on the screen. You can also find it in the Media Advisoryonline. Without further delay, I'll turn it over to Director Santos.

Robert Santos: Thank you, Michael. Good morning, everyone. Today we're releasing results from two efforts to further assess the quality of our nation's 2020 Census. Today's release illustrates our continuing commitment to transparency and scientific integrity. You deserve to know everything we know about the quality of the census. We need to know about the quality of the census to provide guidance to the public on uses of the data as well as for planning future data collection.

The results stem from two independent studies to measure the quality of the 2020 Census: The Post-Enumeration Survey, a study that explores the types of households and people who were counted correctly, and also when people shouldn't have been counted or were missed altogether; and the Demographic Analysis, a study that independently estimates the population by demographic groups as a benchmark relative to corresponding decennial counts. We are

releasing both analyses at the same time to give a fuller picture of the quality of the 2020 Census.

Now, because no census is perfect, we believe it's more productive to think about the usefulness of the data for intended purposes. This gets at the concept of fitness for use. Decennial censuses over time inherently have featured varying qualities of fitness for use. So today's presentation represents part of our effort to inform you on fitness for use by presenting findings on the strengths and limitations of the 2020 Census data.

As you know, the Census Bureau faced an unprecedented set of challenges over the last two years. Many of you, including myself, have voiced concerns. How could anyone not be concerned? Today's findings will put some of those concerns to rest and leave others for further exploration.

Since becoming director, I've had the honor to meet with many who helped carry out the census. They are extremely dedicated, and fully embraced their responsibility to conduct a rigorous count of the nation's population with the utmost scientific integrity. So, I want to express my support of and sincere gratitude to the Census Bureau staff for their tenacity through all the challenges they faced. I also want to thank our partners and stakeholders for their tremendous support.

Today, you will see statistical evidence that the quality of the 2020 Census total population count is robust and consistent with that of recent censuses. This is notable, given the unprecedented challenges of 2020. But we will also note some limitations. You will see evidence that the 2020 Census undercounted many of the same population groups we have historically undercounted, and it overcounted others. Specifically, undercounted groups include the Black population, the Hispanic or Latino population, the American

Indian and Alaska Native population living on reservations, and the population group that reported being of Some Other Race.
At the other end of the spectrum, the 2020 Census overcounted the non-Hispanic White alone population and the Asian population. The Native Hawaiian and Other Pacific Islander population experienced neither an over nor undercount. Like previous censuses, the 2020 Census undercounted children, especially young children ages 0 to 4.

Now, all censuses have limitations, yet can still provide valuable information to our society. This is certainly true in 2020. Taking today's findings as a whole, we believe the 2020 Census data are fit for many uses in decision-making as well as for painting a vivid portrait of our nation's people. Yes, there are areas of concern and we'll be exploring those further. That is part of our due diligence, our pursuit of excellence and our service to country.
We remain proud of the job we accomplished in the face of immense challenges.

And we are ready to work with stakeholders and the public to fully leverage this enormously valuable resource. In closing, please note that additional coverage estimates will be released this summer.

Next, we are going to hear from Erika Becker-Medina, chief of the Decennial Communications Coordination Office. Erika will provide some important additional context before we present our findings. Thank you.

Erika Becker-Medina: Thank you, Director Santos, and good morning, everyone. Before we go into today's findings, I'd like to step back and recall some of the events from the 2020 Census. As you know, 2020 required us to adapt our operations to meet the unprecedented challenges that arose from the COVID-19 pandemic. Even

under ideal circumstances, conducting a census is an enormous undertaking, involving hundreds of thousands of people and dozens of operations and systems, all with the goal of counting everyone once, only once, and in the right place. The historic pandemic led to stay-at-home orders, that forced us to temporarily cease all in-person field operations to ensure the safety of our employees and the public.

These delays compounded other challenges. Under normal census circumstances, we would have completed in-person field operations before hurricane season was in full swing. Instead, we hit our peak operations as the nation faced multiple devastating hurricanes. In addition to the hurricanes, we had devastating wildfires, with dangerous air quality issues from the smoke. And when we were in the field, concern about the virus transmission also caused people to be more hesitant about having discussions with strangers at their door, and it made it more difficult to collect data in person from households that did not self-respond to the census.

But we adapted to provide additional opportunities for everyone to respond, and extended data collection by two and a half months, to allow more time for households to respond.

We changed our procedures to minimize in-person contact with the public, by leaving census invitations in mailboxes, and trained our census takers to exercise social distancing. We expanded outreach through more than 400,000 national and local partners and expanded the paid advertising campaign, to engage more audiences and local media markets.

We deployed staff to places in low-responding areas to answer questions and help people respond to the census. We sent teams of skilled census takers

from other parts of the country that were closer to being finished to work in areas lagging after hurricane damage.

Thankfully, through all the challenges, the public could still respond online, by phone and by mail, and respond they did. About two in every three households across the nation responded on their own, with a self-response rate of 65%, beating the 2010 rate of 61%. Of those households that responded on their own, 4 of 5 chose to do so online, and we did not experience a single minute of downtime, or any cyber intrusions for our online response option. But self-response is not the only way we count the nation. Even if self-response rates for an area were low, that does not necessarily mean that the population was undercounted, because we used additional methods, like sending out census takers to collect data from households that did not respond on their own.

As a result of all the extraordinary efforts, we completed the job we set out to do -- account for virtually all housing units assigned in the United States. And while we're proud of the work we did, we know it's only part of the story. That's why we've sharing the information all along to assess the quality of the data. The Census Bureau has a long established commitment to transparency. We maintained that commitment with the 2020 Census and have taken additional actions to communicate our understanding about the quality and fitness for use of the data we collected. Sharing what we know when we know it.

For the first time in our history, we released operational data quality metrics on the same day as the first census results. These metrics include information on how people responded to the census, as well as how the Census Bureau accounted for addresses that did not respond. Additional quality metrics followed, providing further insight into how housing units were enumerated.

We also released state-level findings for selected metrics, as well as metrics for item nonresponse, which occurs when a respondent provides some information but does not respond to all the questions. The release of these operational quality metrics provided data points related to how we managed census operations and the outcomes from those operations.

In addition to these metrics, we're completing a series of planned assessments and evaluations of the 2020 Census operations. Operational assessments provide data on workload volumes, production rates and cost related to operations, processing and systems. Evaluations determine the effectiveness of census components and opportunities for improvement and innovation.

We also continue our work with respected members of the scientific and statistical communities to provide independent external assessment of the 2020 Census. Today's first release of the first Post-Enumeration Survey results along with details from Demographic Analysis provides additional data points in understanding the quality of the 2020 Census, that when considered with other data points add to our understanding about the completeness, accuracy and quality of the 2020 Census and strengthen our belief that the 2020 Census data are fit for use.

The coverage patterns we're sharing today are at the national level. Although people may want to know if their city or neighborhood was undercounted, we cannot produce coverage estimates at such low levels of geography. While coverage is one of many data quality indicators we had for the 2020 Census, it's the only indicator we have of undercount and overcount. And we can only produce coverage estimates by the characteristics we collected in the 2020 Census, which means we won't have coverage error estimates for the foreign-born population or by poverty status.

Now, I'll turn it over to Eric Jensen, Senior Technical Expert for Demographic Analysis, to provide some of the latest findings. Eric.

Eric Jensen: Thanks, Erika. Good morning. Demographic analysis, or DA, is a long-standing program the Census Bureau uses to evaluate the quality of the census. DA was first used by the Census Bureau in 1960, and has been used every decade since. DA are national estimates of the population on census day by demographic detail.

The estimates are produced using current and historical vital records, data on international migration and Medicare records. It did not use any 2020 Census data from the DA population estimates, as it's completely independent of the census. Almost all of the data used for the 2020 DA estimates were produced before April 1, 2020, and are therefore, unaffected by the COVID-19 pandemic. We use the demographic analysis population estimates of the net coverage error at the national level by demographic detail. This metric gives us insight into the quality of the 2020 Census that I'll explain momentarily.

The methodology used to create the Demographic Analysis estimates is built on decades of extensive research and collaboration both internally and with expert demographers around the nation, and this makes it a valuable tool for evaluating the decennial census. To estimate the population ages 0 to 74, we used the demographic balancing equation, where population equals births minus deaths plus immigrations minus emigration. Birth and death records come from the National Center for Health Statistics.

We used several sources of data to estimate international migration, including our American Community Survey. We used a different method to estimate the oldest age groups because the birth records before 1945 were not as complete as they are today. We use Medicare enrollment records for the cohorts born

before 1945, which is the population aged 75 and older on April 1, 2020. We did make adjustments to Medicare records to account for people who are ineligible for Medicare, delay enrollment or never enroll in Medicare. Finally, to calculate the total population, we added the estimates for each birth cohort and then added those to the Medicare-based assessments for the older ages. Although we use all these different data sources, the birth records are the foundation of the DA estimates.

We used the population results from demographic analysis to estimate net coverage error in the decennial census. Net coverage error is calculated with the census count minus the DA estimate. Then we divide that by the DA estimate and multiply by 100 to get a rate. The DA estimates are used as the denominator because we use them as a benchmark for the census.

Net coverage error combines both undercounts and overcounts for the same group. This means that if a group had a large undercount and an equally large overcount, it would show that as a net coverage error of zero. However, groups that are consistently undercounted in census usually do not have large overcounts too. DA is useful for showing patterns about coverage across demographic groups and we can look at these within a census or across different censuses. Demographic analysis has historically been used to highlight coverage differentials by age, sex and race.

The 2020 DA population estimates were released by December 15, 2020. This is over four months ahead of the first results from the 2020 Census being released. We did this to show that the DA estimate were independent of the 2020 Census. On that day we released three sets of estimates and we produced a range of estimates, a low, a middle and a high, to reflect the uncertainty in the data and methods used to produce the estimates.

Each of these series was slightly different, had slightly different assumptions about the population. We also released the components of population change for all three sets and series. We've already released several of the DA net coverage error estimates. We released the DA net coverage error estimates for the total population on April 26, 2021, when the apportionment counts were released. In November 2021, we released a blog that used the redistricting data to calculate net coverage error for select age groups and also by Hispanic origin.

Today we're releasing the DA net coverage error estimates by age and sex for the 2020 Census. These estimates will be presented in three separate tables. The first table shows net coverage error estimates by single year of age and sex. The second table includes selected age groups, these are mostly five-year age groups.

And the third table shows age and sex in the same broad age categories used by the Post-Enumeration Survey. Because the single year of age and sex data from the 2020 Census has not been released yet, we are using a special 2020 Census file with differential privacy acquired to protect confidentiality. We're not releasing the DA estimates by race and Hispanic origin this time. In order to make comparisons between the DA estimates and the census counts, we used what is called the Modified Race File. This file was produced by the Population Estimates Program and will include the 2020 Census data in race categories that are consistent with the race categories used in the official population estimates.

Finally, we'll release the results from some experimental DA series. These estimates use new data and methods about population estimates for groups we haven't been able to look at in the past. They'll provide more accurate race

and Hispanic origin detail than what we currently produce. And also, we're working on state and county DA estimates for young children.

Now I'll talk about the results. This table shows the national results. Which again, these were first released on April 26 of last year as part of the apportionment release. For the low DA series, we see an overcount of .22% and for the middle and high series, we see an undercount of -.35%, -1.21%. We show an overcount in one series and undercounts in the other series. To make sense of this we need to understand how we developed the range of estimates.

We did not produce standard errors for DA, like we would for survey estimates. Instead, we did a sensitivity analysis to create a range of estimates to account for uncertainty. To produce the range, we developed the middle series first and then varied the levels of different components to create the low and high series. These components include the historical births, international migration, and Medicare-based estimates. All three series are plausible estimates of the population living in the United States on April 1st, 2020. To choose between the three, it's important to understand how they are different. This graphic breaks down the differences between the low and middle series and the middle and high series by population component.

For example, 56% of the difference between the low and middle series, comes from international migration. The difference between the middle and high series is also driven by international migration, as well as adjustments to the Medicare enrollment records. So, for the low series, where we see an overcount, the DA populational estimates for this series have less international migration, fewer historical births, and fewer people in the oldest ages. For the middle and high series, where we see undercounts, there are higher levels of

international migration, more historical births, and more people in the oldest ages.

Next, I'll present the DA net coverage error estimates by single year of age. This graph shows the net coverage error estimates by single year of age, for our three DA series. If the value is below zero, that indicates an undercount. Values above zero are overcount. The largest undercounts that you see are for the youngest ages, which is consistent with past censuses. You see large overcounts for college ages, and also for retirement ages in the low and middle series. These are the blue and red lines.

As mentioned previously, the different series were created based on slightly different assumptions about populations. The differences in the net coverage error estimates across the series reflects this assumption. Additionally, estimates for the age of 75 and older are unique because they use Medicare enrollment records and not birth records to produce the population. The larger range across these series for the older age groups reflects this methodological difference, as well as the specific adjustments that were made to each series to reflect the different assumptions about Medicare data.

Overall, the patterns we see in the single year of age net coverage year estimates are kind of hard to interpret. We see large overcounts for certain ages and large undercounts for the surrounding ages. This pattern is caused by "age heaping" in the 2020 Census result. Age heaping refersto distortion of age distribution of a population where the number of ages reported that end in zero and five, so, for example, 20, 25, 30, 35, is higher than what would be expected to naturally occur.

Age heaping often happens when people reporting for someone else, either another member of their household or a neighbor giving a proxy response. We

can think about it this way, if you don't know exactly how old someone is, you're more likely to guess a rounded number, like 45 or 50, than to guess 46 or 49. Age heaping happens in every census, but what we're seeing in the 2020 Census, is a little more pronounced. This graph shows the 2020 Census results by single year of age in gray, with the DA population estimates, on the top which are the dash lines. These arrows highlight the spikes in the age distribution that are caused by age heaping.

You'll notice the population, the DA population estimates, the dash lines, did not show any age heaping because they're produced using administrative records. To help minimize the effects of age heaping, we present our results in age groups. This graph shows the DA net coverage year estimates by selected age groups, and again, these are mostly five-year age groups. A key finding is that the 2020 Census undercounted young children aged 0 to 4. This is an age group that is persistently undercounted in decennial censuses.

The Census Bureau did a lot of work this past decade to try to improve the count for young children in 2020 Census. We conducted research on this issue which led to operational changes in 2020 Census. Additionally, we made counting young children an important part of the 2020 Census integrated marketing campaign. Throughout all of this, we worked closely with stakeholder groups. But the coverage of young children is a complex problem.

Today we released an American Counts Story that discusses our strategy to improve data on young children, and I'd encourage you to read it. We find a large overcount for the population age 18 to 24, the DA estimates are produced for the entire population living in the United States on April 1, 2020. They include people living in both housing units and group quarters, such as college dormitories. We do see an undercount for working-age adults, and I'll talk more about this group later. Finally, we see a large overcount in the low

and middle DA series for the retirement ages. However, for the high series we see undercounts for the oldest ages. For ages 75 and older the range of estimates comes exclusively from change in the levels of under enrollment in the Medicare records.

Finally, I'll present the DA net coverage error estimates by age and sex. As mentioned earlier, a strength of demographic analysis is that it highlights differential patterns in coverage in a decennial census. In this graph we see the differences in the DA net coverage error estimates by age and sex. The blue line is for males and the red line is for females.

The child population less than 18, we see very small differences in coverage rates by sex. However, if we look at the adult working ages, we see big differences in coverage for males and females. Previously we showed an overall undercount for these ages, but when we look at it by sex, we see that males were undercounted, while females had net coverage error estimates close to zero. Again, this is a good example of how DA helps us see differences in coverage patterns between demographic groups.

So, to summarize, the 2020 DA net coverage error estimates by single year of age show a lot of age heaping in the 2020 Census results. Young children aged 0 to 4 were undercounted, which is a persistent problem in decennial censuses. The DA results showed that the college-aged population was overcounted. Working-age adults were undercounted, and this is mostly because of the larger undercounts for males in these ages. We also see a large overcount for the retirement ages and older cohorts.

I'll now pass over to my colleague, Tim Kennell to talk about Post-Enumeration Survey.

Tim Kennell: Thank you. It's my privilege to share the 2020 Census coverage results from the Post-Enumeration Survey, which I will often refer to as the PES. Before I begin my remarks, I would like to acknowledge and thank my co-workers who worked here for the past years to ensure that we have a high-quality and useful coverage estimates for the 2020 Census. And I also want to acknowledge the thousands of respondents who entrusted us with their data and responded to the PES five field operations over the past two years. I will give a brief introduction to the Post-Enumeration Survey before I present two types of coverage results.

Net coverage and gross components of coverage. Net coverage tells us whether the overall census counts were too high or too low. Gross components of coverage provide estimates of how many people were correctly counted, erroneously counted, or missed in the census. I'll end with some remarks about how the PES overcame challenges.

The Post-Enumeration Survey is one way to assess the quality of the census. We use the PES to estimate the number of people in the country. Then, we compare the PES estimates to the census to determine if the census counts were too high or too low. The Post-Enumeration Survey also gives us information about how many people were correctly counted in the census, missed, or erroneously enumerated. Consistent with our prior practice, we will not be adjusting the census counts for apportionment or redistricting based on results from the PES.

The Post-Enumeration Survey is a probability survey where we interviewed people in about ten thousand blocks across the country independently of the census. We then looked for these people in the census to determine who was missed or counted in error. We're releasing these results today so you can see the strengths, limitations, and errors in the 2020 Census. We've conducted a

Post-Enumeration Survey to measure the quality of the census since 1950. Today's census coverage estimates help us understand the 2020 Census quality and will inform our plans for the next census.

Although both demographic analyses in the PES provide independent estimates of the population, they differ in who is in scope. One of the primary differences is that the Post-Enumeration Survey excludes people living in remote Alaska areas and group quarters such as college dorms, nursing homes and prisons. While demographic analysis includes both of those groups. In this slide, we see the 2020 Census counts for the demographic analysis and the PES universals. The bottom row of the table shows the estimated population slides from DA and the PES. The Post-Enumeration Survey and demographic analysis define net coverage error the same way.

In the past, the PES only reported the undercount. So, the sign of the coverage estimates from previous Post-Enumeration Surveys have the reverse sign as I will report in a moment. A few things to note is that when I show a negative number, it means the census count was too low: an undercount. A positive number means the census was too high. And a number close to zero means the census count was about right. This chart shows the 2020 Census count excluding people living in group quarters and remote Alaska areas and that count was 323.2 million people. The PES estimated that the population really had 323.9 million people, as seen in the bar chart on the right.

A few bars look like they're about the same height and in fact statistically we can't say with confidence whether there was an overall undercount or overcount at the national level. This graph shows the estimated net coverage error for the past four censuses. The vertical lines or towers at the end of each bar show the 90% margin of error for the estimates. If the vertical lines touch the zero line, we can't constantly claim that there was an undercount or

overcount. The 1990 Census undercounted the population. The 2000 Census overcounted the population.

The 2010 and 2020 Censuses did not have a statistically significant overall undercount or overcount.
Here we see the net coverage error rates by race and Hispanic origin. Except for the non-Hispanic White Alone, all rows show race alone or in combination. In this and the other tables I showed, the 2020 estimates will be to the right and the earlier estimates will be to the left. People who select more than one race are included in this table for each race they selected. This ensures accurate coverage of all who selected any specific race.

The 2020 Census undercounted many of the same population groups we have historically undercounted, while overcounting others. For the 2020 Census we estimated undercounts for the groups Black or African Americans, American Indians or Alaska Native, Some Other Race and Hispanic or Latino.

Overcounts were estimated for White, Non-Hispanic White Alone and Asian. For some groups, undercounts and overcounts by race and Hispanic origin were more pronounced in 2020 than in 2010. Rows highlighted in green show the groups for which there was a statistically significant change from 2010. We saw statistically significant changes for the following groups: non-Hispanic White, Asian, some other race, and Hispanic or Latino. The 2020 Census undercounted young children ages 0 to 4, despite major efforts by the Census Bureau and stakeholders to improve the count this decade.

More information about the undercount of young children and how we're addressing it are available in the America Counts blog released today.
The 2020 Census undercounted some age-sex groups and overcounted other groups. The 2020 Census undercounted male and female young adults and

males aged 30 to 49. People over the age of 50 were overcounted in the 2020 Census. For the most part, demographic analysis and the PES agree on which age sex groups were undercounted and which were overcounted. Even though, specific estimates are not always the same. In the past, demographic analysis and the PES have not always agreed.

But in 2020 Census there is considerable agreement in terms of which groups were undercounted and which were overcounted. A noteworthy difference between the PES and DA results is for males and females aged 18 to 29. In that category, the PES says an undercount of 2.3% for males and an undercount of 1% for females, while the DA estimates show an overcount for these populations, except for the high series for males.

I'll note that in 2010 the PES and DA estimates also disagreed about whether there was an undercount or overcount for young adults. The PES and demographic analysis have different universes, methods, and errors, so there are likely many reasons that contributed to these differences, and we are continuing to research disagreements between the PES and DA.

Consistent with prior censuses, the 2020 Census overcounted homeowners and undercounted renters. In general, we know that the Pandemic affected people's job situations and housing. Many people moved temporarily or even permanently as a result of the Pandemic. However, even with the Pandemic-related challenges, the 2000, 2010 and 2020 Censuses all showed overcounts for homeowners and undercounts for renters.

At the national level, net coverage error of the 2020 Census count was not statistically significant from zero, so we can't constantly say there was an overall undercount or overcount. The total population count appears robust and consistent with recent censuses. This is an important finding reflecting a

notable accomplishment amid the unprecedented challenges of 2020. Of course, as with previous censuses, there are limitations. There are statistically significant undercounts or overcounts for specific groups.

The 2020 Census continued to undercount some race groups and ethnic groups, while overcounting others. The 2020 Census also undercounted children, especially young children. The Census continues to undercount renters. Even with the limitations, the 2020 Census are fit for many uses and decision-making, as well as for painting a vivid portrait of our nation's people. In addition to estimates of net coverage error, the Post-Enumeration Survey also estimated components of coverage.

This slide shows the components of coverage for the census on the left and the PES components of coverage on the right. The census counts on the left is divided into three components: direct enumerations, erroneous enumerations and whole-person census imputations. The PES estimate on the right is divided into people who are correctly counted in the census and omissions.

This slide and the next will help describe the four components of coverage. Let's look at the big blue bars at the bottom. The PES estimated that the census correctly counted 305.1 million people. Correct enumerations refer to people counted in the census who were living in the U.S. on April 1, 2020.

According to the PES, the people should have been and were counted in the census. Because the proportion of the other components of coverage is relatively small, the next slide will show them separately. However, this slide shows that the correct enumerations are the largest component of census coverage. There were 7.17 million erroneous enumerations. Erroneous enumerations include duplicates, as well as people who are counted but should not have been.

For example, they may have died before April 1, 2020, or were just visiting the country. The erroneous enumerations are shown in red. There are 10.85 million whole person census imputations. For some records in the census, we didn't receive a response with enough characteristics. So, we used this statistical technique called whole person imputations to fill in the blanks. Some of you may be wondering how the PES dealt with administrative record enumerations and proxy responses.

Generally, they are included in estimates of correct and erroneous enumerations, just like the other responses. In the summer, we plan to release some tables that will show components of coverage specifically for administrative record enumerations and proxy responses. Omissions - People who were counted in the population, but not correctly counted in the census are shown in green. The PES estimated 18.8 million omissions.

But many of them were accounted for in the census as whole person imputations. This chart shows that even though the census counts are about the same as the PES estimates, this is only because the 18.8 million omissions on the right are balanced by 7 million erroneous enumerations and 10.8 million whole person imputations. This chart shows the components of census coverage.

One noteworthy change from 2010 to 2020 is the decrease in the erroneous enumeration rate. The 2010 Census had a duplication rate of 2.8%. The 2020 Census duplication rate was 1.6%. While reviewing the 2020 Census data the Census Bureau determined that there was a need for an additional round of unduplication because of Pandemic-related migration. This effort certainly contributed to the decrease in the number of duplicates and erroneous enumerations in the 2020 Census.

Before I share results and census components, I want to note a few things about whole person census imputations. The Census had 10.85 million whole person imputations. All characteristics were statistically filled in for these census person's records. We break down the imputations into two groups. Households where the number of people in the household was already known, and households where we didn't have the household size.

Whole person imputations where the population count was already known included situations where a proxy responded, or a household resident knew the number of people in the household but had very limited information about the occupants. The majority of the whole person census imputations were in households where we already knew the number of people in the household. We had 1.86 million whole person census imputations where we imputed the household count and all of the people in the household.

Here we see components of coverage by race and Hispanic origin. As we already saw, correct enumerations are the largest component of census coverage. This graph just shows the components of error. But I want to remind people that the correct enumerations are the largest portion of the census count. Erroneous enumerations are split into two groups: duplicates, shown in red, and other erroneous enumerations in orange.

One goal of the Post-Enumeration Survey is to measure these components of coverage. For the most part, the PES cannot answer why some groups had different amounts of correct or erroneous enumerations. Here we see the components of coverage for the age groups. As a results of the Pandemic, many college students and others moved at the beginning of the census reference day of April 1, 2020. This caused some challenges with counting young adults.

Young adults had an erroneous enumeration rate of 3.6% and whole person imputation rates of 3.9%. In 2010, we saw similar patterns of duplication and erroneous enumerations for males and females in the 18 to 29 age group. So, it's unclear how much of the challenges with correctly counting young adults was related to the Pandemic, or other factors.

The 2020 Census did a better job of correctly counting homeowners than renters. Renters are more mobile than homeowners and may have experienced the Pandemic and census differently than homeowners. There remain opportunities to improve counting renters in the census.

This slide shows components in census coverage by relationship to the householder. There are differences between the various groups. More work could be done to increase the correct enumeration rate for people who are not the householder, spouse, or unmarried partner. As previously mentioned, we divide the population into people who are correctly enumerated in the census and omissions.

Overall, the omission rate was 5.8%. Here we see the omission rate by race and Hispanic origin from the 2010 in baby blue. The 2020 omission rates are in dark blue to the right. As a reminder, omissions are people who were in the population but weren't correctly enumerated in the census. Some may have been accounted for in the census as whole person census imputations. Nevertheless, we clearly see from this graph that omission rates vary by race and ethnicity. The omission rates for renters are higher than owners, as was the case in 2010. The coverage results I mentioned today are included in a report on the Census Coverage Measurement webpage and can be downloaded on data.census.gov. So far, we've focused on releasing census coverage estimates by demographic characteristics that I presented today. We are still

working on additional coverage reports and tables which will be released when they're available this summer.

This summer we'll be releasing three reports with more coverage estimates. The first report will include coverage estimates by state. By that I mean we'll have one table with a row for each state and D.C., and the percent net coverage error rate for each state. The PES sample size is not large enough to produce unbiased estimates of any characteristic at the state level. The report will also include national coverage estimates for people by various census operations. For example, there will be components of coverage for internet responses, non-response follow-up enumerations, administrative record enumerations, proxy responses and householder responses.

The second summer report will include coverage rates of housing units. This report will include tables showing how many housing units in the census were correctly counted, erroneously counted, or missed. These will be broken down by various housing unit characteristics such as vacancy and occupancy.
The third report will contain coverage estimates for people in housing units in Puerto Rico, by many of the breakouts in the other reports.

I'd now like to turn our attention to the quality of the Post-Enumeration Survey. The 2020 Census had many challenges and so did the PES. No survey is without challenges. Because of the COVID-19 Pandemic, some changes were made to the PES. Probably the most visible modification was delaying four of the five field operations. Many surveys doing in-person interviewing in the summer and fall of 2020, suffered from large amounts of non-response. Delaying field work and extending deadlines probably contributed to the PES having higher response rates than many other surveys in the field during the summer and fall of 2020.

Nevertheless, delaying the PES schedule also increased the time between the census and the PES interviews. To lessen the impact of recall bias we equipped our staff with calendars to help people more accurately remember back to April 1, 2020. Many colleges and universities either closed or pivoted to virtual learning in 2020. This contributed to a major migration of young adults, often back home and into the household population.

This migration made it challenging to determine who should be included in the PES and who was out of scope because they should have been counted in college dorms or other group quarters. Another challenge was related to increasing levels of people not answering specific questions or item non-response. It's hard to conduct a PES matching and follow-up work when characteristics are missing. On the positive side, we overcame many challenges over the past two years.

We adapted our operations given the changing Pandemic environment and continued to implement quality safeguards into all of the field and clerical matching activities. The PES results today provide valuable insights into how census coverage differs by a variety of demographic characteristics at the national level.

In addition to releasing the report on census coverage today, we're also releasing a technical document called a Source and Accuracy Statement that describes many of the possible errors in the PES data. Here is one table based on data in the Source and Accuracy Statement. One major difference between the 2010 and 2020 Post-Enumeration Survey is the rate of insufficient interviews.

These were interviews where we reached the respondent, sometimes a proxy, who answered some questions, but did not provide enough detailed

information about anyone in the household for the interview to be used. Nevertheless, despite challenges, the PES got completed interviews from 83.2% of the occupied housing units we visited. Today's Source and Accuracy Statement and future methodology reports will provide a lot of information about the quality of the PES.

Overall, the Post-Enumeration Survey achieved its goal of highlighting some of the strengths and weaknesses of the 2020 Census. The coverage estimates I discussed today will be helpful in planning for the 2030 Census and I hope those fund new innovations and research to improve future census operations.

In conclusion, although no census is perfect, the total census count appears robust and consistent with recent censuses. This is an important finding reflecting a notable accomplishment and the unprecedented challenges of 2020. Of course, there are limitations. The 2020 Census undercounted some populations groups and overcounted others. Even with the limitations, the 2020 Census data are fit for many uses and decision-making, as well as for painting a vivid portrait of our nation's people. Thank you and I'll now turn it back to Mr. Cook.

Michael Cook: Thanks, Tim. We'd like to start taking questions now from the media first and then open it up to partners and stakeholders and for this question-and-answer session, in addition to Erika, Tim and Eric, we'll be joined now by Dale Kelly, chief of the Field Division, Jennifer Reichert, chief of the Decennial Management Division, and Karen Battle, chief of the Population Division, and Robert Santos, director of the U.S. Census Bureau.

Reminder, that to ask a question, you must call the phone number displayed on the slide on your screen. You can only ask a question by dialing into our phone line. Before asking a question, please state your name and either your

news outlet or your organization. We want to fit in as many questions as possible, so please only one question with one follow up allowed per organization. Operator, I now hand it over to you for additional instructions.

Operator: Thank you. If you would like to ask a question, please press *1. If you need to withdraw your question, press *2.

Michael Cook: And while we wait for our first question to be queued up, a reminder to check out our press kit online, that's where you'll find a number of resources, including links to where you can find the PES and DA data. Our news release link to today's slide deck and in-depth report on today's data release.

You can also find links to the webinar we conducted last week that explains the purpose and the methodology of the PES and DA, along with an explanatory blog on the components of coverage for the PES and DA. We also have additional blogs, as mentioned earlier and other content that show how PES and DA fit into the overall 2020 Census picture. Operator, do we have our first question?

Operator: Yes. Our first question comes from Hansi Lo Wang from NPR. Your line is open.

Hansi Lo Wang: Thank you, (unintelligible). Hi, this is Hansi from NPR. For each of the groups by race and Hispanic origin, can you tell us these overcount and undercount rates, can they be measured in the hundreds of thousands of people left out? Millions? I understand that because of disclosure avoidance issues, the Census is not releasing the total populations of the people who were used for this Post Enumeration Survey, so can you give us a ballpark sense? Are we talking about hundreds of thousands, millions of people who were overcounted or undercounted by race and Hispanic origin groups here?

Michael Cook: To answer that question on overcounting and undercounting is specifically on a raise, so with that, it is a topic dealing with the Post Enumeration Survey, I'm going to turn that over to Tim.

Tim Kennel: For this release we're focusing on the undercount and overcount rates and not on total levels. The total levels have not been released from the Census yet and so I can't make any statements about that, and they have not gone through disclosure avoidance processing either.

Hansi Lo Wang: Okay, then I guess the follow up question, is the Census Bureau planning to use any of these over or under counting rates by race and Hispanic origin to adjust the baseline used to calculate annual population estimates or any other non-2020 Census data products? And if not, why?

Michael Cook: Thanks for that follow up and as you know and stated in our reporting, the Post-Enumeration Survey and DA does not affect the apportionment in the official counts of the 2020 Census, but for things that it will affect and adjust, I'll turn that back over to our subject matter expert, Eric and Tim, but also, better yet, let me toss this to Karen Battle, the Chief of Population Division who can speak to that directly. Karen.

Karen Battle: Hi. Yes, thank you, Michael. So, as Tim mentioned earlier, we of course do not have plans to adjust the official 2020 Census population count. However, we are taking a closer look at the net coverage error information that's been presented today and we know that there's additional net coverage error information that is yet to be produced. So, we want to take a look at all of that and think through the research that we would need to do to help us figure out whether or not this information can help us adjust the base of our population estimates program as we move forward into the decade.

And we have already taken some steps in this direction because we are using a blended population estimate base for our Vintage 2021 estimates where we have blended together 2020 Census data, 2020 demographic analysis data and 2020 population estimates data and that is helping to mitigate somewhat the effect of the undercount of young children. So, we're taking steps in that direction, but we have to do research so that we can understand whether or not we can do that.

Michael Cook: Thank you, Karen. Operator, do we have our next caller?

Operator: Our next question comes from Mike Schneider from Associated Press.

Mike Schneider: Hi, good morning. Good morning. I had a question for Director Santos. I wanted to ask you if you had any theories about the increase in the Hispanic undercount and do you think the role of the policies of the Trump Administration was a factor?

Director Santos: Thank you for that question, Michael. I think that the condition of the population during the Pandemic was quite profound. We had families of all races and ethnicities, but especially among Latinos who were really suffering during this period. They were out of work. There were issues of housing stability, there were hunger issues, and so forth, and I think that played a role in the ability to secure participation. Despite I think the wonderful job that was done by our partners and stakeholders to help bolster the participation on the ground level. We need more of that for next time and so I think that really it was more of the Pandemic that we had to deal with.

Michael Cook: Thanks for that Director Santos. Operator, do we have our next caller?

Operator: Our next caller comes from Michael Macagnone from CQ Roll Call.

Michael Cook: Hi, Michael.

Michael Macagnone: Hi, Michael, can you hear me?

Michael Cook: Loud and clear.

Michael Macagnone: All right, perfect. I actually wanted to follow up a little bit on Mike Schneider's question. Director Santos, prior to taking on your role at the Agency you spoke at Urban Altered Report raising concerns about specific decisions made by the Trump Administration that raised the possibility of an undercount among the Hispanic and Latino population. What do you think the results released today say about the concerns that you had back then, specifically about the citizenship question and then later decisions made such as cutting in-person counting short?

Director Santos: What I can say immediately is that the Census Bureau's own research in terms of focus group research raised concerns very early on about the inclusion of the citizenship question and so therefore all of the publicity surrounding the efforts to place it on may well have had an impact. And so, I am personally not surprised to see the results that we see today. Having said that, I think it's more appropriate to look at what's going on today, what we have found in terms of the strengths and limitations of the 2020 Census counts for various populations and start talking about how do we use and leverage that to understand and make the most use of our valuable 2020 Census data and how can we move forward in order to create methods and other types of partnerships with communities to increase participation. Thank you.

Michael Macagnone: Also, if I can follow up on that, Director Santos, what do you think is the Census Bureau's role in addressing problems that communities with high Hispanic and Latino populations – I know concerns have been raised about the Rio Grande Valley, in addressing problems that they might have as a result of the undercount in terms of getting needed program funds, in terms of getting grants, things like that.

Director Santos: Thank you for that question also. The 2020 Census is a starting point for population estimates. The population estimates can incorporate known errors in the counting of certain jurisdictions.

So, for example, if a city did not properly have its boundaries included and so had an undercount according to that, then evidence can be brought through our Count Question Resolution Program so that we can take that into account and adjust our population estimates accordingly and further as we go on throughout the decade there are also mechanisms like the Census Challenge Program for jurisdictions to also make sure that we're doing the best we can in terms of providing the most accurate population estimates because it's the population estimates program that's typically used for the allocation of federal funds as well as for calibrating our really important surveys, like the American Community Survey, the Current Population Survey and so forth. Karen may have more to add to that, but otherwise that is my response.

Michael Cook: Thanks for that, Director Santos. Operator, do we have our next caller?

Operator: Yes. Our next question comes from William O'Hare from Count All Kids.

Michael Cook: Hi, Bill.

William O'Hare: Hi, Mike, can you hear me alright?

Michael Cook: Loud and clear, sir.

William O'Hare: I want to follow up on the video released about the undercount of young children and I want to quickly thank the bureau for focusing on this. As you know it's an issue that we have shared for a long time, but I know the bureau did a lot to lift this issue up in the 2020 Census and it's a little bit disappointing and unfortunate that the net undercount in young children increased by a percentage point or so over the last decade, but kind of understandable because of all the factors.

The question I have is, can anyone at the bureau talk about how the newly constructed cross-directorate team in the undercount of young children, or working group, whatever it's called, is going to use the data that was released today and subsequent from PES and DA and their plans to improve the count in 2030?

Michael Cook: Thanks for that question, Bill, specifically about the undercount in young children and you mentioned the group that's currently working in the Census Bureau to try and head this off at the pass and I'll speak or send this over directly to Eric Jensen, specifically about, talk about DA and then open it up for others that might need to chime in. Eric.

Eric Jensen: Yeah, thanks, Bill for your question. So, we released an America Count Story today, which talks specifically about the coverage of young children and the decennial census and as was mentioned in the presentation earlier, this is a persistent issue. We've seen undercounts in 2010, in past censuses.

So, the Bureau has recently formed a cross-directorate team, so it's made up of subject matter experts in demography, in statistics, in survey design, in census operations, in partnerships, in communications, and these experts are coming together, and the focus is -- we have three main things we're working on. One is research into this. Why were young children missed? What is it about young children that makes them hard to count in the census?

Two, looking at future data collection. So, you mentioned specifically 2030 Census, but we're also looking at ways to improve how we make the household roster in the American Community Survey and that can impact the coverage of young children. And then finally, through data products, we're looking to improve data on young children. As mentioned earlier, the population estimates program, to run this blended based approach is using some controls from demographic analysis by age and that should mitigate some of the issues we see for the undercount in children in the 2020 Census.

And as you know those population estimates are really important not only for allocating federal funds, but they're also used as controls on demographic surveys, which is the ACS and the CPS, and they're used as denominators for vital rates, so for information produced by like the National Center for Health Statistics, those vital rates use the population estimates. So, this new focus on not only data collection like we did last decade but also on data products I think is what sets this work apart from the Census Bureau's past efforts to improve the data on young children. Thank you.

Michael Cook: Thank you, Eric.

William O'Hare: Thank you.

Michael Cook: Operator, do we have our next caller?

Operator: Our next question comes from Dr. Yvette Roubideaux from National Congress of American Indians.

Michael Cook: Hello, Dr. Roubideaux.

Dr. Roubideaux: Hello, I hope you can hear me.

Michael Cook: Loud and clear.

Dr. Roubideaux: Thank you. Well, it is very concerning to see that American Indians and Alaska Natives living on reservations have again the highest undercount in the 2020 Census and even though it's not statistically different compared to 2010, it's probably because there wasn't enough in the sample, it's still concerning for us. Are you planning on immediately consulting with tribal nations on this persistent undercount and the actions you will take to address these persistent undercounts? It's very concerning because of the lost potential resources and funding, and I know our communities are likely very concerned.

Michael Cook: Thanks for that line of questioning, Dr. Roubideaux. Yes, we're very grateful for the tribes and to the tribes and many tribal citizens who worked as census takers, and our partners who helped us get a count of people living on reservations. I'll turn your line of questioning over to the panel to see if there's any remarks that those would like to make specifically about your line of questioning, about the undercount of American Indians and Alaska Native population.

Director Santos: I can actually start with that at a high level, in that part of the reason I accepted the Directorship when I was asked to serve the country was because I wanted to focus on outreach and so I see it as the high priority. To conduct

outreach to various stakeholder groups, including people on tribal lands so that we can better understand and form partnerships and reduce issues like mistrust or misconceptions about how census data are used and why they are collected. So that is high on my list, and I expect to be doing that in coming months. And I leave it to the rest of the panel to add whatever they would like.

Dr. Roubideaux: Just on a follow-up. I know that there is a tribal consultation scheduled later this month. I hope that the results of the PES can be presented to the tribal leaders in general public or lay language and a robust conversation can happen about the implications of these results as you work on DHC and and DDHC data.

Director Santos: Indeed. Thank you.

Dale Kelly: And if I may add to that, thank you for that question, because we are definitely grateful for the collaboration and the efforts and partnership that we had with the American Indian and Alaska Native community during the decennial and what we're putting in place going forward is actually identifying specific resources, creating tribal specialist partnership positions who will work with our tribal nations and our communities throughout the non-decennial years and keep informed about any upcoming data releases. What are new efforts that we need to make going forward that will help us collect and make sure we get a more accurate count in 2030. So, we will continue this effort that we've not necessarily done in the past through the non-decennial years. Thank you.

Michael Cook: Thanks for that Dale and thank you Director Santos. Operator, if you could check for the last question and while we wait for that last question to be asked, I'd like to remind everyone that we have a lot of events coming up that I think will be of interest. Later this afternoon at 2:00 pm EST, we'll host a webinar

giving a preview of what to expect for the release of the American Community Survey five-year estimates. Again, that webinar is just a few hours away, at 2:00 pm EST.

The embargo for ACS five-year data will begin March 15 at 10:00 EST, with public release at 12:01 am EST, March 17, that Thursday. On March 14, we'll hold a webinar in advance for the release of the 1950 Census records on April 1. Our Country and Metro Micro Area Population Estimates are scheduled for embargo March 22 at 10:00 EST for embargo subscribers. The data will then be publicly released on March 24 at 12:01 EST. Be sure to check out Census.gov for more details on all of these events and, Operator, do you have that one last question?

Operator: Yes. Our last question comes from Tara Bahrampour from the Washington Post.

Michael Cook: Hi, Tara.

Tara Bahrampour: Hi, thanks for taking my call. Thanks for the session. My question is you've mentioned a few times during the session that the census is fit for use for many purposes. Can you tell us what purposes it is not fit for?

Michael Cook: Thanks for that line of questions about fit for use and I'm going to turn that over to our panelist and I've got Dale Kelly here. I'm sorry, Karen Battle, rather, sorry, who can talk a little about the quality of the count and any others that would like to chime in. I know that we have a full panel of experts today. Karen.

Karen Battle: Yes. I would just say that we know that the census data has many, many uses, but the census data that we have already published of course are fit for use for their mandatory purposes, for apportionment, for the redistricting file and also for just helping us have a better portrait of the population in the United States. So, I don't know that I could tell you specifically what uses the data are not fit for because I see the data are fit for many, many uses.

Tara Bahrampour: It just seemed like there was a qualification. You know, instead of saying, all uses. It seems like there's some possibility that it might not be fit for other uses. Is there anyone else who can land on why that wording was used?

Director Santos: Yes, I'm happy to weigh in on that. There are technical issues regarding the level of geography because we will be producing a data that are subject to things like disclosure avoidance.

Some of the smallest units, like at the block level, we may have issues if you take for example all of the smallest blocks in the country and put them together in the data set to do small population block analysis because of the way we do data disclosure, that's not advisable and there are other things like that we can very well provide some guidance and we intend to provide guidance on the appropriate uses of census data or not. When it comes to today's topic, like the estimates of over and undercount, it's more of a matter of making sure we understand the strengths and limitations. So, for example, we can use the 2020 Census results to identify cities that have certain percentages of the Latino population or Asian population, etcetera, keeping in mind that there were undercounts for some and overcounts for others.

And it's simply a matter of tempering and not taking something, some data results and acting as if they're absolutely, 100% true, so to speak. All censuses

are imperfect in some point and so in order to use them and leverage the most insight out of them, we need to understand the limitations and therefore take that into account. That's what I would have to say. Thank you.

Tara Bahrampour: Thank you.

Michael Cook: Thank you for that, Director Santos, and Karen. I've been told that this event has a lot of interest today. I know I called for one last call, but I've been told that we do have others that have jumped in and if the panelist can indulge me, I'm going to hold you hostage just a little longer. Operator, can we have our next caller please?

Operator: Yes, our next question comes from Dr. Arturo Vargas, from NALEO Education Fund.

Michael Cook: Hello, Dr. Vargas.

Arturo. Vargas: Thank you, and it's not doctor, it's Mister Vargas, and thank you for taking my call. As you can imagine, we are terribly… I can't even find the right word… upset about the extent of the Latino undercount that has been now confirmed by the Post-Enumeration Survey.

In my over 35 years of having worked on promoting a fair and accurate survey and census, I don't believe I've seen such an extensive undercount rate of Latinos as we're now seeing in 2020. This represents a major step backwards on this and while the data overall for the national population may be fit for use, given the significant undercount of Latinos, to what extent are the data on Latinos fit for use for all the purposes that we see going forward over the next eight years until the 2030 Census?

Michael Cook: Thank you for that line of questioning, Arturo. I'll toss that over to Eric.

Karen Battle: Hi there, this is Karen. I wanted a comment about that. So, one of the things that I wanted to point out is that the 2020 Census counts of the Hispanic or Latino population at the national level and at the state level are in line with and look demographically reasonable when compared with our April 1, 2020, population estimates.

And I'll just note that when we released the redistricting data last year we did publish an American Count Story that looked at this a little bit closer and what we saw was that for about 40 or so states, the 2020 Census count for Hispanic or Latinos was actually higher than our population estimate, again which is a typical population benchmark that we would use and then for the ten or so states where the 2020 Census counts where Hispanic or Latino was less than our population estimates, we looked at the percent of the state population that was Hispanic or Latino and saw that the difference between the census and the population estimates was less than two percentage points for each of those states, so I just wanted to mention that as we of course rely on DA and we rely on PES as unofficial coverage measures, we do also look at other types of measures and indicators of quality. And one of those ways is taking a look and comparing the data with our population benchmarks. Just wanted to mention that.

Michael Cook: Thank you, Karen. And, Operator, do we have any more callers?

Operator: I'm showing no further questions at this time.

Michael Cook: Well, I'd like to thank everyone for their questions and for joining us for today's news conference and we hope that today's presentations gave you a

deeper understanding of the data and a better context of what PES and DA tell us about 2020 Census data quality.

A reminder, go to census.gov to access our press kits for more information and details. If you have questions, additional questions, please call the public information office at 301-763-3030 or tollfree at 1-877-861-2010 or email us at pio@census.gov. If you're joining us for the Spanish language presentation, please stay tuned. That will begin shortly. Otherwise, thank you for attending today's news conference and have a great rest of your day. Hopefully we'll see you later this afternoon for the ACS five-year estimates pre-released webinar.

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