Modernizing the Population Estimates Base: What We've Learned from Three Years of Updates

Christine Hartley, PhD, and Luke Rogers, PhD, U.S. Census Bureau

Introduction

The Population Estimates Program (PEP) at the U.S. Census Bureau is responsible for producing the official estimates of population and housing units in between the decennial census counts. The range of uses for these estimates is vast: they serve as denominators for key statistical rates and controls for major national surveys, guide planning at the state and local levels, and inform research across academia and the public and private sectors. Importantly, they also factor into the distribution of trillions of dollars in federal funding—although the Census Bureau has no involvement in the funding formulations themselves.[1]

PEP has been producing estimates for many decades, and that time has been characterized by countless methodological improvements. Over recent decades, the estimates are developed in an annual cycle which features a period of research to design and test updates to the methods. The outcome each year is a new time series, or vintage, of population estimates. The word *vintage* is used to distinguish a time series of estimates created with a consistent population starting point and methodology, as well as to indicate the latest year of estimates available. Each new vintage includes revised estimates starting at the date of the most recent decennial census and up through the vintage year.

One commonality across vintages has been to build the estimates off the results of the latest decennial census. This starting point for the time series is referred to as the "estimates base." However, the 2020 Census was subject to some notable and unexpected circumstances, including the COVID-19 pandemic, which had a significant impact on decennial operations. This impact combined with changes resulting from the modernized disclosure avoidance framework[2] meant that the 2020 Census detail PEP needed was not available in time to produce the Vintage 2021 estimates base for counties and higher levels of geography. For the first time, staff in PEP needed to create an alternative estimates base—one that was not exclusively drawn from the most recent census.

This paper describes the revolutionary "blended base" approach, which was implemented for the Vintage 2021 estimates, and explores the changes that have been applied to subsequent vintages as PEP transitioned from a reactive model for the estimates base to a proactive model.

---

[1] Villa Ross, Ceci. (2023). *Uses of Decennial Census Programs Data in Federal Funds Distribution: Fiscal Year 2021*. U.S. Census Bureau. https://www2.census.gov/library/working-papers/2023/decennial/census-data-federal-funds-fy-2021.pdf, accessed September 13, 2023.

[2] More information on the 2020 Census Disclosure Avoidance System is available at https://www.census.gov/programs-surveys/decennial-census/decade/2020/planning-management/process/disclosure-avoidance.html.

Methodology

A high-level understanding of PEP's general methodology is key to fully appreciating the methodological departure represented by the blended base.[3] Although PEP produces estimates for numerous levels of geography extending from the nation down to cities and towns, different methodological approaches are applied based on the specific geography. Counties and higher levels of geography are developed using a cohort-component method (Figure 1) whereby changes relative to the estimates base are captured through a variety of administrative records and survey data on births, deaths, and migration—known as the components of population change.

Figure 1. The Balancing Equation for the Cohort-Component Model

Related to this method, there are several types of updates that may be incorporated into each new vintage. First, updates may be applied to the estimates base. Regular updates of this type include the application of legal boundary changes or other geographic updates so that a given vintage reflects geography that is current as of January 1 of the vintage year (e.g., geographies across all dates in the Vintage 2022 times series are current as of January 1, 2022). The estimates base may also be updated due to the outcomes of official census programs such as Count Question Resolution (CQR) and the Post-Census Group Quarters Review (PCGQR), or changes to the method of how the base population is estimated.[4] Then, the components of change may be updated to draw from more recent or complete input data that have become available (e.g., the vital records that are used to estimate births and deaths). Additionally, improvements may be made to the methods of how the various components of change are estimated. Finally, details of the overall method for tying the components together and generating the estimates via the equation in Figure 1 may be modified.

The type(s) of change can have an impact on the magnitude of change. For example, data updates are more likely to affect the estimates for the most recent years of the time series, while method revisions—either specifically pertaining to the components of change or the overall methodology—generally affect the series cumulatively from the census date forward. However, some types of change have historically been more common than others: particularly, updates to the estimates base prior to Vintage 2021 were almost exclusively limited to geographic updates and the incorporation of CQR results.

[3] For more detail on PEP's methodologies, please see https://www.census.gov/programs-surveys/popest/technical-documentation/methodology.html.
[4] For more information on Count Question Resolution, please see https://www.census.gov/programs-surveys/decennial-census/decade/2020/planning-management/evaluate/cqr.html.

The Vintage 2021 Blended Base: A Reactive Model

When it became apparent that a major methodological adjustment would be necessary to create the base for the Vintage 2021 estimates, there was no clear precedent to follow. Furthermore, a looming deadline to publish the estimates by the end of 2021 meant that time was significantly limited for designing and testing an alternative.

One option to meet the deadline was to create estimates for July 1, 2021, by extending the Vintage 2020 time series by an additional year. Vintage 2020 was the final series built off the results of the 2010 Census. Instead of restarting at 2020, PEP could have incorporated updated data on births, deaths, and migration to cover the period of July 1, 2020, to July 1, 2021, and released a time series for April 1, 2010, to July 1, 2021. However, in advance of Vintage 2021 production, it was known that the Vintage 2020 estimates were lower than the results of the 2020 Census. Thus, taking this approach would have retained that difference, significantly limiting the utility of the resulting Vintage 2021 estimates.

Instead, PEP staff focused on using the Vintage 2020 estimates as a starting point and identifying how other data sources could be used to improve upon it. Within the limited window of time available for research, data needed to be readily available ("in-house") and previously vetted regarding the methodology for producing the data and their quality. Two datasets which met these requirements were the 2020 Demographic Analysis (DA) estimates and the 2020 Census Public Law 94-171 Redistricting File.[5,6]

The 2020 DA estimates for April 1, 2020, were produced by PEP staff using an administrative records-based approach. These national estimates rely on birth records instead of decennial census data to form the starting population, and so the resulting estimates are considered independent of the Census. For that reason, they are used as a measure of coverage. In the time available, PEP staff were successfully able to incorporate national age and sex detail from this source into the estimates base.

The 2020 Census Redistricting Data Summary Files became available while Vintage 2021 estimates research was underway. Although this dataset includes demographic detail, to use these data in the estimates base required sufficient time for PEP staff to (1) research whether the data were suitable for estimates production and (2) examine the impact of incorporating the detail into the estimates base. For example, although the Hispanic population as well as the population aged 18 and older were available around the time that estimates production got underway, early research on incorporating these data resulted in implausible demographic patterns within the base. Additionally, only very limited national-level coverage measures for this detail were available to address questions, particularly from stakeholders, pertaining to the quality of these data.

The end result, dubbed the "blended base," consisted of April 1, 2020 estimates of total resident population for counties and higher levels of geography which matched the published results of the 2020 Census, and which featured demographic detail from a combination of the Vintage 2020 estimates series and the national 2020 DA estimates. This combination of data resolved the difference between the Vintage 2020 estimates of total population and the results of the 2020 Census, but this wasn't the

[5] For more information on the DA estimates, please see https://www.census.gov/programs-surveys/decennial-census/about/coverage-measurement/da.html.
[6] More information on the 2020 Census Redistricting Data Summary Files is available at https://www.census.gov/programs-surveys/decennial-census/about/rdo/summary-files.html.

only benefit. It also turned out that the incorporation of 2020 DA age and sex data at the national level produced a smoother, more demographically plausible age distribution than the 2020 Census results, which contained age heaping at the highest level since at least the 1980 Decennial Census.[7] Additionally, the blended base estimates exhibited a somewhat mitigating effect on the persistent undercount of young children in the census data[8] because the vital records which inform the 2020 DA age and sex data provide excellent coverage of this population. This translated into both higher population values for children ages 0 to 4, as well as ages 10 to 14 (children who were ages 0 to 4 in the 2010 Census, who had retained the undercount as they were "aged forward" to 2020 in the population estimates).

Most importantly, the blended base provided a path forward for the population estimates that did not rely on 2020 Census data availability.

Transitioning to Proactive Estimates Base Improvements: Vintage 2022 and Vintage 2023

Despite the limited time available to produce it, the blended base provided an effective and adaptable framework. The completion of the original model also opened the door to more deliberate and systematic improvements to the base population within the regular annual research cycle. Simply knowing in advance that base population research must occur for Vintage 2022 and having the Vintage 2021 framework to start from represented a huge advantage relative to the previous year.

The development of the Vintage 2022 blended base took several key factors into account:

- *Options for expanding the use of the current data sources*. The 2020 Census and DA data used in the Vintage 2021 blended base represented the greatest level of detail that could be confidently incorporated given the time and resources available. For Vintage 2022, it was possible to revisit these sources and evaluate what additional detail may improve the blended base. Specifically, whereas the Vintage 2021 blended base used age and sex detail from the 2020 DA estimates for ages 0 to 85+ (representing the group of individuals aged 85 and older), for Vintage 2022, it was possible to expand this to the full DA age range up to ages 100+.

  Additionally, Vintage 2021 used the 2020 Census PL 94-171 Redistricting File for the county resident population totals. The resident population is the sum of the populations living in households and group quarters facilities, which are typically processed separately and then combined to form the resident population. For Vintage 2022, the blended base method was improved so that it incorporated 2020 Census household and group quarters population by the 7 major facility types, derived from the estimates base used for subcounty estimates processing aggregated to the county level.[9] Not only did this improve the consistency across levels of

[7] Age heaping occurs when population counts for ages ending in preferred digits such as 0 and 5 (e.g., 20, 25, 30) are higher than would be expected based on known birth, death and migration patterns. For more details, see: https://www.census.gov/newsroom/blogs/random-samplings/2023/05/age-heaping-2020-census-dhc.html).
[8] More information on the coverage of young children is available at https://www.census.gov/library/stories/2022/03/despite-efforts-census-undercount-of-young-children-persists.html.
[9] For more information on the GQ population and descriptions of the facility types, please see the Group Quarters section in Appendix B: https://www2.census.gov/programs-surveys/decennial/2020/technical-

geography, it meant annual geographic updates for counties could be reflected in the blended base—something that wasn't possible for Vintage 2021.

- *Possible improvements flagged during the review of the Vintage 2021 blended base.* The process of producing the estimates includes rigorous review at several stages to ensure the validity and reliability of the output. During review, it is common for items to be flagged as areas where it may be possible to test and implement improvements for future vintages—which was the case during the review of the Vintage 2021 blended base. As such, PEP staff tested and implemented updates that would better approximate the size of the household and group quarters populations in the blended base by making changes to the order of processing and pulling in the additional 2020 Census data, as described above.
- *Required updates*. The development of Vintage 2022 followed the state of Connecticut's transition from 8 counties to 9 county-equivalent planning regions. Consequently, the blended base method was modified to process estimates for the 9 planning regions.

These improvements and their respective impacts are described in more detail in the next section.

The Impact of Coverage Concerns

While PEP staff were identifying, testing, and evaluating these changes, the Census Bureau was proceeding with its efforts to publish 2020 Census data products. Of note was the March 10, 2022, release of official measures of coverage from the 2020 Post-Enumeration Survey (PES) and DA. Among the findings from this release, it was confirmed that many of the same population groups that have been historically undercounted were again undercounted in the 2020 Census, although the total population was determined to be robust. Similarly, there was evidence that some populations had been overcounted. The implications of these results inspired thoughtful discussion among Census stakeholders, ranging from casual data users to government officials. Many questioned what could be done to improve the counts given the statutory limitations preventing post hoc changes to the Census results, especially as they are used for apportionment or redistricting. Soon, these questions were directed to PEP: could adjustments for coverage be made to the 2020 Census data used in the blended base? This was seen as a desirable outcome because it would ensure that the estimates built from the base carried these adjustments forward across the decade, resulting in more accurate and equitable data for their myriad of uses—including their use in funding level calculations.

This question was regarded seriously by Census leadership, who called for a team to be established to determine the feasibility of taking coverage measures from the 2020 PES, DA, and other sources into account in the development of the population estimates in order to inform decisions about how the blended base is improved. In response, the Base Evaluation and Research Team (BERT) was formed, consisting of Census Bureau subject-matter experts in the areas of population estimates, age and sex statistics, coverage measurement, race and ethnicity, demography, disclosure avoidance, and (added later) group quarters population.

documentation/complete-tech-docs/demographic-and-housing-characteristics-file-and-demographic-profile/2020census-demographic-and-housing-characteristics-file-and-demographic-profile-techdoc.pdf.

BERT's Research Agenda

The team established numerous research priorities, the first of which was to assess the history of post-decennial census adjustments to the estimates. This was an enlightening endeavor which confirmed that regular adjustments to the estimates base had been limited to updates to accommodate official changes from programs like CQR. In certain instances, survey controls created from the estimates were adjusted to account for the findings of coverage evaluations.[10] In others, a variant of the cohort-component approach known as the "inflation-deflation" method was applied to preserve the age distribution of undercounts in the base population rather than aging them across the decade.[11] However, BERT was unable to identify evidence that PEP had ever adjusted census totals in the estimates base specifically for coverage issues, although it had been considered on multiple occasions.

Other research priorities focused on the analysis of 2020 Census data by subject area and the evaluation of the 2020 Census coverage measures. For the 2020 Census data, a systematic approach was adopted to determine whether using additional data from the 2020 Census would improve the blended base. Each subject area (age and sex, race, and Hispanic origin) sought to make comparisons to benchmark data, evaluate change since the 2010 Census, and assess the data's demographic plausibility. The underlying, longer-term strategy was to first confirm which 2020 Census results could be incorporated into the estimates base, and then to identify whether there are adjustments that should be made to these data to address coverage issues, either before or after they are used in the blended base.

Regarding the 2020 Census coverage measures, the PES and DA, the intent was to thoroughly evaluate the methodology used to generate the measures to determine whether, based on their design, the measures are fit for use in the estimates base. Additionally, the team would assess whether design changes would increase the utility of these measures for possible adjustments to the estimates base in future decades, should that be warranted. In this sense, suggestions could be shared to facilitate planning for coverage evaluation of the 2030 Census.

The intent was for BERT to pursue these research priorities so that recommendations could be made to PEP regarding what additional 2020 Census data were suitable for use in the blended base, and whether there were adjustments that could be applied to those data to address coverage issues.

Although stakeholders were eager for BERT research to improve the blended base, the timing of the team's formation and the scope of its research were not compatible with the production timeline for the Vintage 2022 estimates: by the point at which BERT began meeting in March of 2022, the Vintage 2022 improvements had already been defined and were being researched and tested in order to be implemented by the December deadline to release the data.

As such, the research completed by the team over its first eighteen months was oriented toward making recommendations regarding the Vintage 2023 estimates base. Over that time, the greatest strides were made in the evaluations of the 2020 Census results by age and sex and by Hispanic origin, as well as the 2020 PES (details below).

[10] For more details, see the section, "Adjustment of CPS Controls for Net Underenumeration in the 1990 Census" in https://www2.census.gov/programs-surveys/cps/methodology/tp63rv.pdf.
[11] For example, see https://www.census.gov/content/dam/Census/library/publications/1974/demo/p25-519.pdf.

It is important to note that all of BERT's progress was made in spite of fairly significant resource limitations and in addition to meeting regular estimates production milestones. However, it was recognized that to adequately tackle the scope of BERT's work—in addition to other estimates improvement research—would require additional staff. Thus, BERT and PEP jointly pursued options to expand resources, including seeking support from other areas of the Census Bureau with mutual interests and issuing a request for a dedicated budget initiative. Although the outcome of this initiative is still undetermined, it would enable a significant expansion to PEP that would establish an infrastructure which prioritizes estimates improvement (versus the current structure, which enables a limited amount of research to be accommodated around the annual estimates production schedule). In the meantime, some temporary support was secured from other parts of the Census Bureau, and so a small number of staff were added to PEP and BERT to contribute to the progress described below.

Age and Sex

The primary focus of this work was the age heaping in the 2020 Census, both at the national level and in many states. The research identified a small subset of cases that contributed to the majority of the heaped ages, investigated various ways to fix the heaped data, and recommended a particular approach using a Gaussian filter to smooth the age distribution. This method proved effective in reducing age heaping for nations, states, and counties. However, to apply this method to a 2020 Census file and have it available for use in the blended base requires adherence to the agency's disclosure avoidance guidelines; specifically, an appropriate differential privacy mechanism must be developed to protect the confidentiality of the data. This process extended beyond the window for Vintage 2023 production, and thus incorporation of these findings will be tested for a future vintage.

Hispanic Origin

The research on Hispanic origin sought to determine the reasonableness of the Hispanic origin data from the 2020 Census relative to the detail from Vintage 2020 present in the first two iterations of the blended base. The work included assessing the impact of updates to the Hispanic origin question as well as coding and editing procedures for 2020, leading to the conclusion that these changes did not affect the data's usability for the estimates base. The bulk of the analyses consisted of comparisons between the 2020 Census counts and benchmark data sources for the Hispanic resident, household, and GQ populations. The benchmarks included in these evaluations were the Vintage 2022 population estimates for April 1, 2020, and the 2021 American Community Survey (ACS) 1-Year and 5-Year estimates. Results at the national, state, and county levels indicated that the resident and household Hispanic populations were similar across the data sources and appeared demographically reasonable. However, counts of the Hispanic GQ population from the 2020 Census diverged notably from estimates of the GQ population from benchmark sources. These conclusions led to the recommendation that Hispanic totals for resident and household populations be incorporated into the Vintage 2023 blended base, but that further research was necessary prior to using 2020 Census data for the Hispanic GQ population.

2020 Post-Enumeration Survey

The 2020 PES research included an evaluation of the survey's history since 1990 and considered the impact of changes in sample size and methodology. It explored how and why the topic of census adjustment has been considered but ultimately never implemented in prior decades. Primarily, the work delved into the 2020 PES results for demographic groups and states, and particularly the issue of

synthetic estimation (the process that would be employed to translate national PES estimates to lower levels, such as counties or demographic groups within states). The limitations of this approach were explored in depth, namely that the coverage rate for a small area would be determined by the characteristics of the people in that area rather than a unique area-level effect, resulting in synthetic bias. For the 2020 PES, the synthetic bias of the state estimates was reduced by the inclusion of a state effect in the estimation model; this had not been the case in 1990, 2000, or 2010 where the PES estimates for states were synthetic because they did not include a state effect. This update for 2020 represented an improvement to the method, but came at the cost of adding variability (i.e., sampling errors) to the state estimates; adjusting the blended base using the PES results would transfer this variability/uncertainty into the estimates. Ultimately, the design of the PES was deemed inconducive to valid, reliable adjustments to the estimates base—which is not a shortcoming of the survey, since it was not designed for this purpose.[12] In fact, the PES still provides critical insight for BERT regarding the strengths and limitations of the 2020 Census counts and where targeted adjustments may be warranted, and thus may inform changes to upcoming vintages by helping to identify subpopulations of interest.

Stakeholder Engagement

Given the level of interest and investment in BERT's progress expressed by stakeholders, the team consistently prioritized stakeholder engagement. Frequent briefings were provided to a wide range of internal and external groups, including the Federal-State Cooperative for Population Estimates, the Census Scientific Advisory Committee, the Census Quality Reinforcement Taskforce, the Centers for Disease Control, the Office of Management and Budget, the Government Accountability Office, the State Data Centers and Census Information Centers, and congressional staff. Presentations were made via public webinars as well as at professional meetings, including the Population Association of America and the Southern Demographic Association. Blog entries were also published on www.census.gov describing research efforts and findings.[13]

For the first phase of research, these efforts culminated in an ad hoc expert panel on Vintage 2023 base population updates, which took place via virtual meeting on November 3, 2023. The purpose of this meeting was to engage in discussion with external experts to help assess the sensitivity of findings, develop external communication regarding BERT research progress, and inform further research/changes for Vintage 2024 and beyond. The panel consisted of nine external experts with diverse backgrounds in terms of their specific experience and professional affiliations. BERT members from across the Census Bureau presented background on the team and the blended base; an in-depth look at expected Vintage 2023 improvements to the base population, which was as-yet-unreleased; and findings from the evaluation of the 2020 PES. The resulting discussion with the panel participants was robust and engaging, covered a wide range of feedback, and a indicated a strong degree of support for the team's efforts thus far and intentions for future research. BERT intends to leverage this meeting format again to share further findings as they become available.

[12] More details on this decision are available at https://www.census.gov/newsroom/blogs/random-samplings/2023/12/recommendations-2020-pes-coverage-results-in-vintage-2023-pop-estimates.html.
[13] For example, see https://www.census.gov/newsroom/blogs/random-samplings/2023/05/age-heaping-2020-census-dhc.html; https://www.census.gov/library/stories/2023/06/blended-base-methodology.html; and https://www.census.gov/newsroom/blogs/random-samplings/2023/12/recommendations-2020-pes-coverage-results-in-vintage-2023-pop-estimates.html.

The Evolution of the Blended Base

Table 1 provides a high-level summary of the way the base population has evolved, including Vintage 2020 for reference. This section explores the changes for Vintages 2021 through 2023 and illustrates the impact of the various improvements.

Table 1. Vintage Estimates Base Population Sources: 2020 to 2023

				Most Recent Census (Nation, States, and Counties)						2020 Demographic Analysis (National)						Vintage 2020 (Nation, States, and Counties)					
Vintage	Geographic Updates	CQR	PCGQR	Pop. by Universe (HH/GQ)	Total Pop.	Age	Sex	Race	Hispanic Origin	Pop. by Universe (HH/GQ)	Total Pop.	Age	Sex	Race	Hispanic Origin	Pop. by Universe (HH/GQ)	Total Pop.	Age	Sex	Race	Hispanic Origin
2020	X	X		X	X	X	X	X													
2021	Subcounty & HU only				X							0 - 85+	X			X		X	X	X	X
2022	X	X		X	X							0 - 100+	X			X		X	X	X	X
2023	X	X	X	X	X				X			X	X			X		X	X	X	X

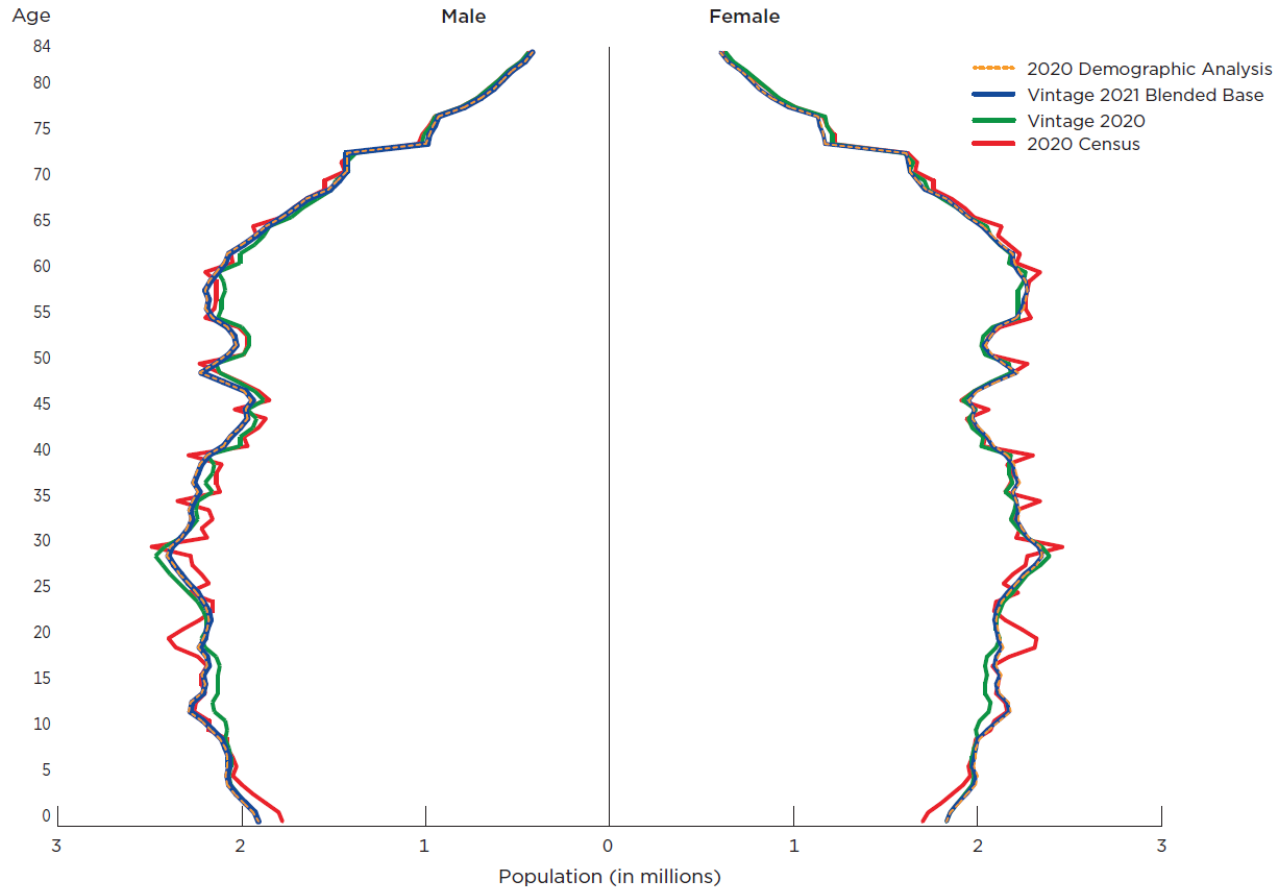
Note: CQR = estimates base updates resulting from the Count Question Resolution Program. PCGQR = estimates base updates resulting from successful Post-Census Group Quarters Review cases. HH = household population. GQ = group quarters population. HU = housing units.

Many distinctions are evident between Vintage 2020 and Vintage 2021. The most significant, of course, is the introduction of the additional sources of data. It is possible to examine the impact of combining these data sources by plotting them on a population pyramid, as in Figure 2, which features the age and sex distribution for the 2020 Census, 2020 DA estimates, Vintage 2020 estimates, and the Vintage 2021 blended base.

As noted previously, by relying on 2020 DA estimates by age and sex at the national level, the blended base (represented by the blue line) does not feature the telltale spikes resulting from age heaping that are apparent in the red 2020 Census line. This is also the reason that the dotted yellow line representing the 2020 DA estimates appears to fall directly on top of the blue blended base line.

Additionally, the impacts on children ages 0 to 4 and 10 to 14 can be observed by comparing the blue blended base line against the red 2020 Census line and green Vintage 2020 line, respectively. The way the 2020 Census line falls below the blended base line for both males and females indicates that there are more children ages 0 to 4 in the blended base than in the 2020 Census. Similarly, the blended base line extends beyond the Vintage 2020 line for males and females ages 10 to 14, indicating there are more children in this age group in the blended base. These children would have been ages 0 to 4 in the 2010 Census, which formed the base population for the Vintage 2020 estimates. Then, they were aged by ten years to April 1, 2020, in the cohort-component method, but retained the undercount from the 2010 Census.

Figure 2. Vintage 2021 Blended Base, Base Inputs, and 2020 Census Data by Sex for Ages 0-84: April 1, 2020



Note: These 2020 Census data by age and sex represent a special tabulation of the 2020 Census with confidentiality protections applied using the 2020 Census Disclosure Avoidance System. The U.S. Census Bureau reviewed this data product for unauthorized disclosure of confidential information and has approved the disclosure avoidance practices applied to this release. (DRB clearance number CB-FY22-DSEP-001.)
Source: U.S. Census Bureau, 2020 Decennial Census; 2020 Demographic Analysis; Vintage 2020 and 2021 Population Estimates.

The most significant update to the blended base for Vintage 2022 was the incorporation of subcounty 2020 Census counts for the GQ population by the 7 major facility types and for the household population. The original Vintage 2021 blended base matched 2020 Census totals for the nation, states, and counties, but specifically for the resident population—which represents the sum of the household and group quarters populations. The GQ data in the Vintage 2021 blended base were held constant from the Vintage 2020 estimates base. However, at the subcounty level (which is estimated with a separate methodology), it was possible to maintain a decennial census-based starting point by infusing a small amount of differentially private noise into the 2020 Census household population and the GQ population by the 7 major facility types. This subcounty estimates base update was completed after the publication of the original blended base, but it meant that for Vintage 2022, these totals could be summed to the county level to serve as updated April 1, 2020 totals in the blended base. This change improved consistency in the estimates across levels of geography.

Table 2 illustrates the results of this update by population universe. Overall, the change to the resident population is very minimal (239 or 0.0%) and represents the net impact of introducing differentially private noise to protect confidentiality in the 2020 Census results. The most notable change is to the GQ population, which increased by 478,046 or 6.2%. This was the outcome of incorporating the updated GQ data from the 2020 Census versus the previous iteration whereby the GQ

population was derived from Vintage 2020 (which had the 2010 Census as a starting point and then featured annual updates to capture change in this population from 2010 to 2020). Vintage 2022 adjusted for an overestimate of the household population and a corresponding underestimate in the GQ population seen in Vintage 2021.

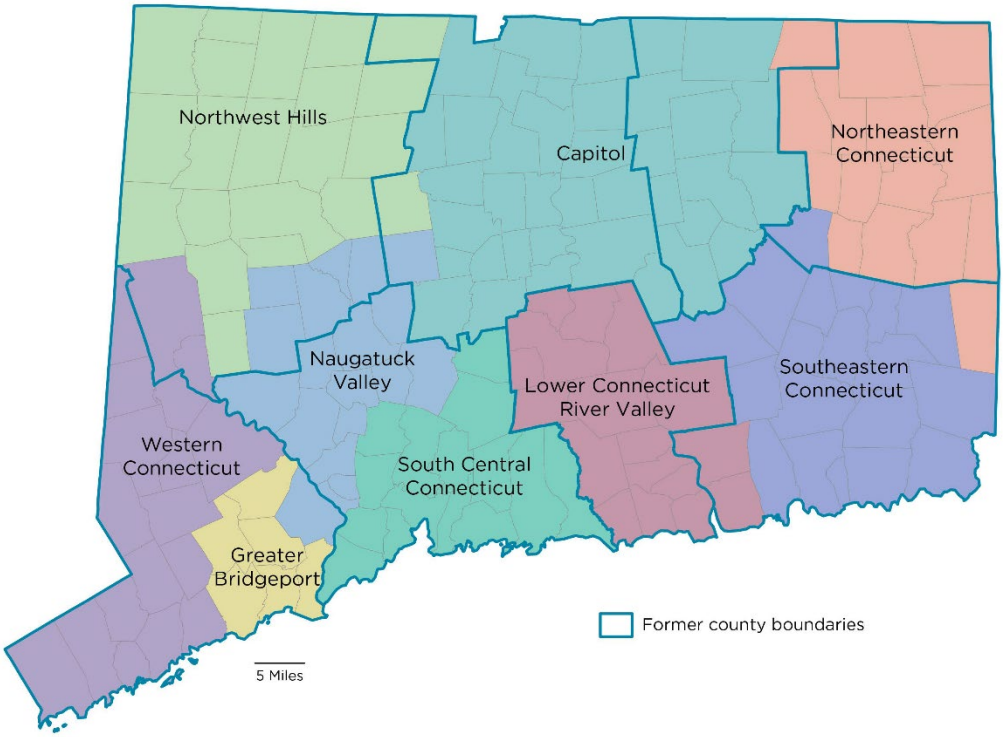
Table 2. Population Estimates and Differences by Vintage and Universe: April 1, 2020

	Population Universe		
	Resident	**Household**	**Group Quarters**
Vintage 2021	331,449,281	323,688,422	7,760,859
Vintage 2022	331,449,520	323,210,615	8,238,905
Numeric Difference	239	-447,807	478,046
Percent Difference	0.0%	-0.1%	6.2%

Source: U.S. Census Bureau, Vintage 2022 and 2021 Population Estimates.

Additionally, by using the aggregated subcounty estimates base, it also became possible to capture the geographic updates which are applied annually to the subcounty population. This was not the case for Vintage 2021, which featured county-level boundaries from Vintage 2020 (i.e., as of January 1, 2020). The timing of this improvement was fortuitous because it ended up coinciding with a rather major geographic update for counties (which is otherwise not nearly as common as it is for subcounty geographies). Specifically, the state of Connecticut transitioned from eight counties to nine county-equivalent "planning regions," as is seen in Figure 3.[14]

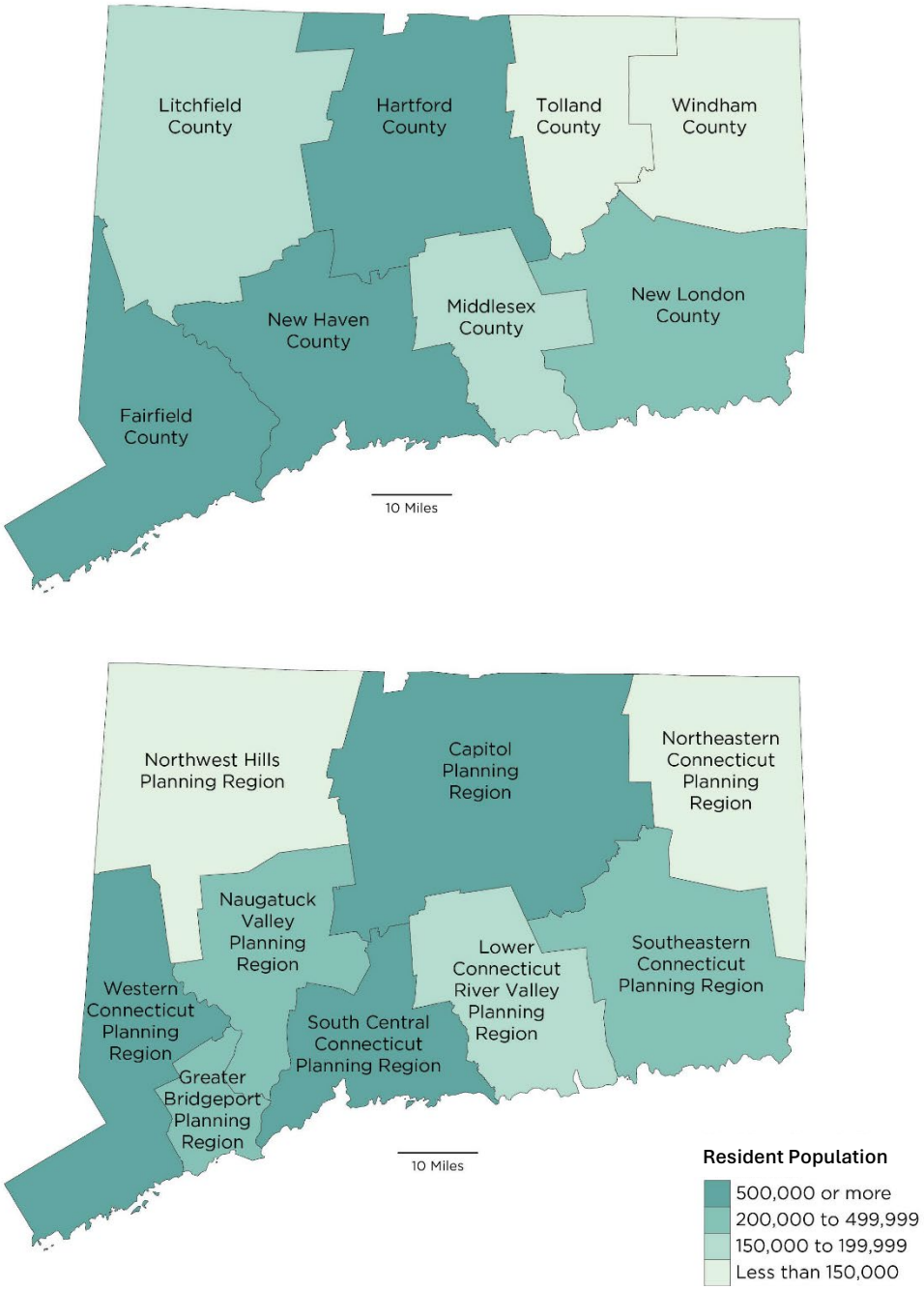
Figure 3. Connecticut Planning Regions with Former County Boundaries Overlay



[14] https://www.federalregister.gov/documents/2022/06/06/2022-12063/change-to-county-equivalents-in-the-state-of-connecticut

The ability to capture geographic updates in the Vintage 2022 blended base enabled PEP to estimate the population for the planning regions. The impact that the new planning regions had on the population levels is evident in Figure 4. Of the population thresholds featured on these maps, change is evident in the middle two: the number of geographies falling into the category of 200,000 to 499,999 resident population increased from one county to three planning regions. Meanwhile, one planning region, as opposed to two counties, had an April 1, 2020 population of 150,000 to 199,999.

Figure 4. Vintage 2022 Population Estimates for Connecticut by County and Planning Region: April 1, 2020



Source: U.S. Census Bureau, Vintage 2022 Population Estimates.

By Vintage 2023, the changes were a combination of processing improvements and updates resulting from BERT recommendations. Notably, two of the processing updates aimed to improve the distribution of the GQ population by the 7 major facility types. Consequently, in Vintage 2023, some counties saw shifts in their April 1, 2020 population—although these shifts were largely concentrated in the group quarters population. This is illustrated in Table 3.

Table 3. Degree of Population Change for Counties between Vintage 2022 and Vintage 2023 by Universe: April 1, 2020

Absolute Percent Difference	Change in Resident Population			Change in Household Population			Change in Group Quarters Population		
	Gain	Loss	None	Gain	Loss	None	Gain	Loss	None
5.00% or more	1	0	0	0	0	0	54	43	0
2.50 to 4.99%	1	0	0	0	0	0	82	84	0
1.00 to 2.49%	2	0	0	1	0	0	220	196	0
0.00 to 0.99%	1,408	1,437	295	1,320	1,394	429	927	960	511

Source: U.S. Census Bureau, Vintage 2023 and 2022 Population Estimates.

Although the majority of counties saw changes to their estimates base between Vintage 2022 and Vintage 2023, Table 3 distinguishes these changes by the type, degree of difference, and population universe. As a result, it is clear that most counties experienced very little change to their overall resident or household populations—less than 1%. Meanwhile, although more than half the counties also saw little or no change in their GQ population, a notable amount (416) saw gain or loss ranging from 1 to 2.49%, with 166 counties experiencing 2.5 to 4.99% change in their GQ population, and 97 seeing an increase or decrease of 5% or more.

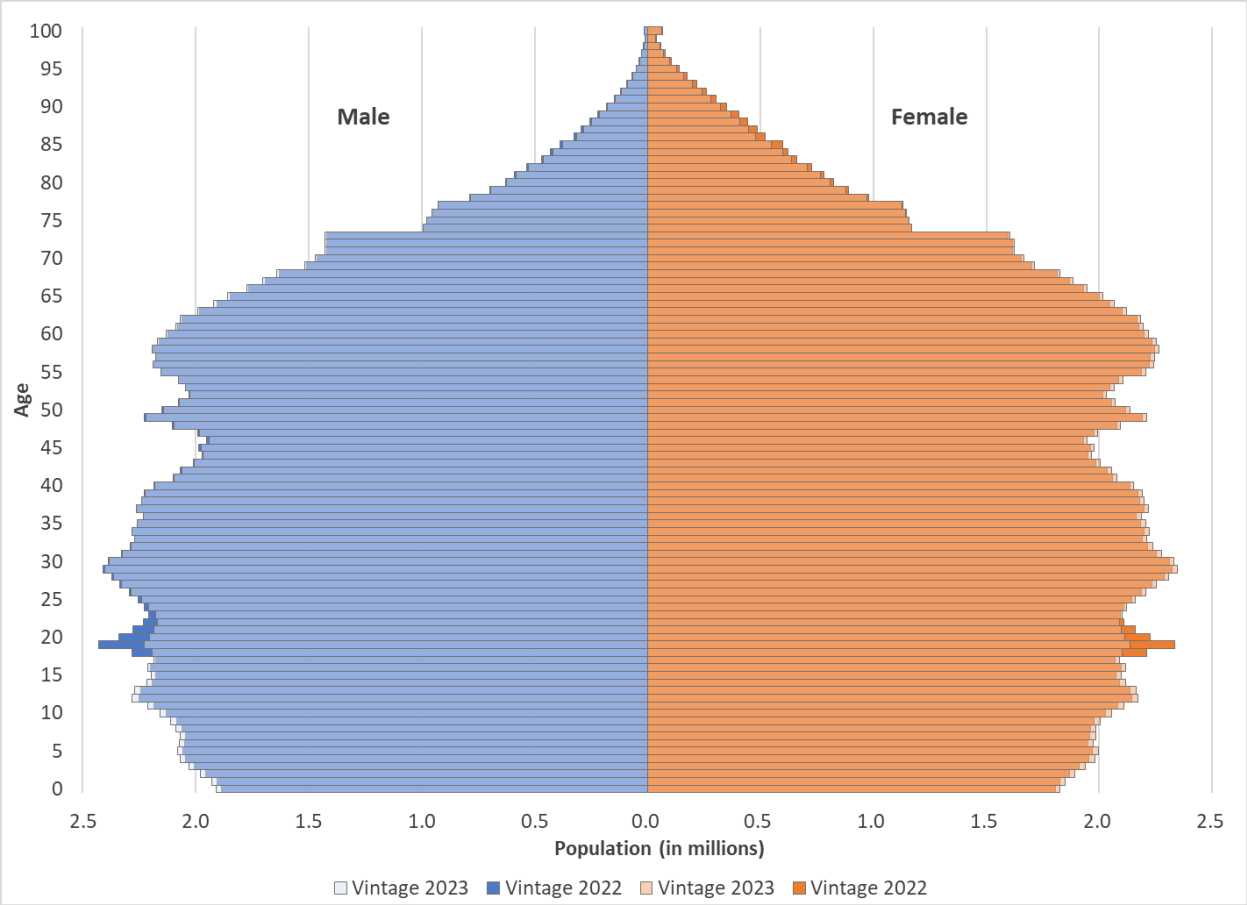
The processing improvements would have directly contributed to these changes, but Table 3 technically reflects the outcome of all combined updates to the Vintage 2023 estimates base. Thus, the impact of geographic updates or changes resulting from the CQR or PCGQR programs would also be captured in this table—as well as the effect of the major update informed by the BERT research, to incorporate Hispanic or Latino resident population totals from the 2020 Census. The Vintage 2023 estimates by age, sex, race, and Hispanic origin have not yet been released to the public (currently scheduled for June 2024), and so the specifics of this change cannot be discussed.

However, it is possible to observe the impact of incorporating Hispanic or Latino resident population totals from the 2020 Census on the national age structure of the blended base, depicted in Figure 5. This figure features the population distribution by single year of age and sex in the April 1, 2020 estimates base for Vintages 2022 (darker bins) and 2023 (lighter bins). Over the younger ages, the lighter Vintage 2023 bins extend beyond the darker Vintage 2022 bins, while at the older ages, the darker Vintage 2022 bins extend beyond the lighter Vintage 2023 bins. Both instances illustrate that incorporating the 2020 Census data by Hispanic origin results in a younger population for these age groups. This is due at least in part to the fact that the 2020 Census had a slightly greater proportion of Hispanics than the Vintage 2020 estimates did (18.7% vs. 18.6%), and this population is typically younger overall than the non-Hispanic population. For point of reference, in the Vintage 2022 blended base, the median age for the Hispanic population was 30.0 while non-Hispanics had a median age of 40.7. Introducing the 2020

Census Hispanic origin data and increasing the proportion of Hispanics in the Vintage 2023 base population impacted the population nationally, resulting in differences like the slight increase in 0 to 15-year-olds relative to Vintage 2022 evident in the population pyramid (Figure 5).

In addition to the impact of the 2020 Census Hispanic origin population, Figure 5 also features evidence of the GQ population processing improvements referenced earlier. Particularly for ages 18 to 23, there is a noticeable difference between the two versions, with Vintage 2023's population size being lower overall. This reflects the processing update which made the method of distributing demographic characteristics more sophisticated, thereby improving accuracy, in conjunction with a modification to address cases where Vintage 2020 data were too sparse to support the calculation of distributions. While both of these updates impacted the resident and household populations, they are primarily visible in the GQ population, especially among demographic groups with high concentrations of population in GQs (e.g., the college-aged).

Figure 5. Estimates of Resident Population in Vintage 2022 and Vintage 2023 by Age and Sex: April 1, 2020



Source: U.S. Census Bureau, Vintage 2023 and 2022 Population Estimates.

Lessons Learned

Over the course of designing, testing, implementing, and evaluating the changes described in the previous section, many lessons were learned which shaped, and continue to shape, decisions regarding BERT research and the development of the blended base. Among these were the following:

1. **A "gold standard" can evolve.**
   For much of PEP's history, the decennial census data—as the most comprehensive account of the nation's population—were regarded as a veritable gold standard across all analyses and applications. For example, each decade, PEP's vintage estimates were evaluated against the results of the latest census to determine the degree of "error" in the estimates. This information would inform research and methodological improvements over the subsequent decade to try to minimize that difference, known as the error of closure, for the next census. Similarly, these data were incorporated into the base population mostly "as is." The program did not otherwise adjust the census data, and by some accounts, actively sought to avoid any appearance of doing so.

   However, this paradigm was very much disrupted by the exceptional circumstances impacting 2020 Census enumeration. PEP staff were faced with the reality that the available decennial census results did not include all the detail necessary for estimates production, and furthermore, may not adhere to the previous quality standards required for the base population. The process of designing an alternative estimates base underscored that combining data sources can ultimately compensate for limitations in the individual sources. Regardless of its status as the most comprehensive account of the nation's population, the decennial census was never truly perfect—the results of coverage measures such as PES and DA over time highlighted this. Thus, by systematically building an estimates base, it becomes possible to target specific populations or patterns for improvement based on known demographic trends—to objectively redefine the gold standard and use each research cycle to strive for it.

2. **The blended base approach is complex.**
   The premise of the blended base is straightforward: separate data sources are combined to produce estimates of the population on April 1, 2020. However, the reality of integrating these sources is far more complex. To begin, it requires identifying the elements of each that will contribute to the blend. In this particular circumstance, the Vintage 2020 estimates served as a starting point, and detail from the 2020 Census Redistricting File and DA estimates were introduced. Assessing the strengths of each data source to isolate the detail for the blend is challenging, especially since comparisons against benchmark data sources are typically used for this purpose. Yet what can stand as a benchmark for the 2020 Census—particularly for the purpose of improving the population estimates? Without sufficient benchmarks to evaluate the files, it can be difficult to distinguish between detail that is common to more than one source—to determine which data are the "best," in the sense that they offer the most accurate depiction of the population. In fact, determining how to appropriately and responsibly use benchmarks to guide decisions has been one of the foremost challenges for BERT.

   Additionally, the order and way the alternative sources are introduced affects the outcome: are totals from the other sources replacing Vintage 2020 totals? Are distributions from the other

sources being applied to the Vintage 2020 totals? At what level of geographic and demographic detail? Once this has been determined, the actual process of integrating the sources must be written out in statistical software (PEP currently uses SAS). In some instances, desirable combinations of data may not be programmable, or may introduce anomalies into the base population. As such, not all recommended improvements can necessarily be incorporated.

Similarly, not all perceived improvements result in true improvements. Changing one part of the base—perhaps targeting an adjustment for one variable, one group, or one geography—can create issues elsewhere, such as implausible demographic patterns. Thus, each update must be introduced independently and carefully evaluated.

Once the output has been produced, it must be reviewed. Here, again, the lack of a benchmark data source introduces challenges: what can the blended base be compared against to assess its validity? If no single benchmark exists, what other sources could be informative?

Each time the blended base is improved, this process must be repeated. Naturally, such an iterative process is intellectually demanding and can quickly become grueling for staff. As such, it is even more important to be deliberate about what, how often, and the manner in which simulations are run and reviewed.

3. **Once the base population becomes adaptive, the resulting possibilities are numerous.** Although the estimates base was not previously eligible for significant changes or adjustments, there is something to be said for the fact that the prospect was considered so regularly across PEP's history. Namely, there was a persistent notion that elements of the decennial census-based population could be improved to serve as an even stronger starting point for the annual estimates. However, to enact such a change required a justification worthy of shifting the paradigm upon which the estimates methodology was built. Until the 2020 Census enumeration was faced with such unprecedented challenges, no justification had risen to this level. After all, using a single source of data is far more straightforward. Once additional sources are introduced, how is the scope of eligible changes defined? How is the full scope of possible changes managed? Serious considerations must be made, including but not limited to:
    - What evidence is sufficient to determine an adjustment to the base population is warranted?
    - How broad do adjustments need to be? Is there a minimum population size for a specific demographic subgroup or geography to qualify for an adjustment?
    - What types of datasets may be used to inform adjustments? Can datasets specific to a single group or geography be used?
    - Does it raise issues of equity if adjustments are researched and applied for some groups and not others? Does the same apply if datasets only exist to support adjustments for some groups and not others?

Ultimately, the intent of BERT's research is to support development of the most accurate estimates base possible. By transitioning to the blended base approach, questions such as these underscore that even though the method is adaptable, adjustments must be deliberate and justified.

Future Research

Taking these lessons into account amidst many others, BERT research and resulting improvements to the blended base are ongoing. Future research also factors in the findings that have been amassed so far and stakeholder feedback.

Generally, the BERT research will be delving further into patterns for specific demographic groups. For example, the team will closely examine age patterns in the 2020 Census results by race, ethnicity, and sex. The impact of data imputation will be considered across demographic characteristics. Where differences are evident between the 2020 Census results and benchmark data sources, the team will look for evidence of geographic patterns.

At the same time, work will continue to investigate potential causes for the increase in heaping in the 2020 Census relative to previous censuses with an eye toward improving the 2030 Census data collection. BERT will also take a closer look at the results of the 2020 Census by sex to determine if an adjustment is needed in order to incorporate those data into the blended base.

The pace of research on incorporating 2020 Census results by race into the blended base is expected to increase. Progress in this area has been limited thus far by the lack of a 2020 Census file featuring race data in categories that are consistent with those used in PEP. Whereas the census race data includes the category "Some Other Race," PEP produces estimates in the categories determined by the Office of Management and Budget (OMB), which do not.[15] To reconcile these differences, a bridging system known as "modified race" distributes the "Some Other Race" group across the OMB race categories. Research to develop a file featuring the modified race variable is currently underway, and upon completion, it will enable more informative benchmark analyses. Metrics on diversity will also be examined between decennial censuses at multiple levels of geography as another means to assess the changes observed over time.

The availability of the modified race file will also enable the calculation of additional DA estimates of net coverage error, specifically by age and sex for the Black/non-Black populations (all ages) and the Hispanic/non-Hispanic populations (ages 0 to 29). The results of this work will provide additional insight into instances of under- or overcounts in the 2020 Census, which could potentially be targeted for an adjustment in the blended base. Exploring alternative data sources to develop such adjustments will be a priority for BERT.

Finally, further GQ research will take place as a dedicated subject area across characteristics. This was inspired by previous research findings indicating distinct differences in the GQ population compared to the household population. The work will, for example, investigate any potential issues with sex reporting for GQs, and continue benchmark comparisons of the GQ population by Hispanic origin and by race with a focus on GQ type.

Ultimately, BERT is striving to improve the blended base by building off its existing research and targeting recognized gaps in knowledge. By leveraging the team's wide range of expertise and strong connections with stakeholders, consistent improvements to the base population have now become a regular, formalized part of the annual estimates production process.

[15] Information on the OMB standards for the classification of federal data on race and ethnicity can be found at https://obamawhitehouse.archives.gov/omb/fedreg_1997standards.