

To: Robert Santos

Director

U.S. Census Bureau

From: Jay Breidt

Census Scientific Advisory Committee (CSAC) Chair

March 18, 2022

Subject: Recommendations and Comments to the Census Bureau from the Census Scientific Advisory Committee Spring 2022 Meeting

## **Introduction**

The Census Scientific Advisory Committee (CSAC) thanks the US Census Bureau for all their work in preparing the Spring 2022 CSAC virtual meeting. CSAC thanks the presenters for their careful preparations and clear presentations, for providing CSAC with specific questions on each presentation, and for keeping to their allotted time and leaving room for discussion. CSAC appreciated the background material that was provided and the high level of technical detail for the topics on the agenda, while noting that CSAC had insufficient time to absorb this information. CSAC would appreciate receiving materials well in advance of the meeting, ideally no later than the Monday two weeks prior to the CSAC meeting (discussants' slides have a deadline of Monday of CSAC week). For more technical content, such as much of the disclosure avoidance content for this spring meeting, one-way briefings to the full CSAC or the appropriate working group would be more effective. Further, publicly available background materials that require no internal Bureau approvals should be made available to CSAC as soon as possible after the agenda is finalized. The more technical the material, the more time CSAC members will need to process it in advance of the meeting. CSAC would like to ensure that our advice is ready in a timely fashion and is as useful as possible to the Bureau.

CSAC workload per committee member is unusually high at this meeting, not only due to the amount of content but because the committee currently has only thirteen members. According to its charter, CSAC can have up to 21 members. CSAC is therefore excited to welcome new committee members after this spring meeting. CSAC notes that the Bureau's orientation process for new members has been very valuable, and that an additional useful tool for engaging new members prior to a full CSAC meeting is through one-way briefings.

CSAC observes that many of our discussions focus on clear communications, especially across disciplinary boundaries and levels of expertise. Some small steps in the direction of clearer communication would be to minimize the use of acronyms, to define any essential acronym on first use, and to provide an updated acronyms list before each meeting. Further, figures and tables should be understandable from the slides alone, with fully identified content and sufficient detail in labels and text. Links to relevant documents might be included in the slides, to facilitate CSAC access to the information.

Finally, CSAC looks forward to learning more about several on-going developments and new initiatives brought up at this meeting. These include but are not limited to: (1) the FRAMES Program; (2) the implementation of the 2020 proposed criteria for urban classification; and (3) the Ask U.S. Panel and its potential utility for addressing census undercount; (4) Diversity, Equity and Inclusion initiatives.

## **Census Bureau Modernization and Transformation Activities**

CSAC was happy to welcome new director Robert Santos and to hear some of his strategic vision for the Bureau. Many of the Director's comments resonated with CSAC: the challenges of data collection, the rising expectations for tailored information, the greater awareness of privacy issues, the critical role of stakeholder engagement, the need to hear diverse voices, and the importance of equitable access to data and analysis. While the Bureau's continued commitment to data-driven innovation, transparency, and scientific integrity is not surprising, CSAC recognizes the challenges of building and maintaining the strength of the Bureau's brand in a landscape of many alternative data products of varying quality.

CSAC appreciated the Director's comments on the potential efficiencies of combining multiple sources of data. While the linkage problems across data sources are challenging, CSAC is intrigued by the possibility of freeing up resources from people who are well-represented in one or more sources and using those freed resources to better describe people who are poorly represented across sources. The FRAMES project is one such instance of combining multiple sources of data, and CSAC requests that we receive regular updates on the progress of the FRAMES program and the technical challenges that underlie this program: see recommendations from the Fall 2021 CSAC meeting.

Results from the Post-Enumeration Survey (PES) and demographic analysis (DA) released last week show that the 2020 Decennial Census continued to suffer from serious differential undercounts, despite the best efforts of the Bureau, for reasons anticipated and unanticipated. These differential undercounts have unequal and often inequitable impacts. Given these coverage issues, CSAC is interested in the potential improvements for the population estimates program, including changes to the population estimates base like those in the "blended base" approach for the vintage 2021 population estimates. CSAC is specifically interested in how the Bureau might partner with state and local governments to improve the population estimates. For

example, how might programs like FRAMES (which relies on data ingested from multiple sources and multiple partnerships) be used to improve annual population estimates at local levels? How might such programs be used to mitigate potential undercount in the 2030 decennial census?

- 1. CSAC recommends that the Bureau research possibilities for making adjustments (for differential undercount/overcount) to the population estimates base and brief the committee on the Bureau's plans for improved population estimates (including any use of the blended base) as soon as feasible, via a one-way briefing or at the next CSAC meeting.**

CSAC noted with interest the Director's reference in his formal remarks to the Census Bureau's Post-Secondary Employment Outcomes (PSEO) program. As the Bureau describes it, "PSEO data provide earnings and employment outcomes for college and university graduates by degree level, degree major, and post-secondary institution. These statistics are generated by matching university transcript data with a national database of jobs, using state-of-the-art confidentiality protection mechanisms to protect the underlying data. The PSEO are made possible through data sharing partnerships between universities, university systems, State Departments of Education, State Labor Market Information offices, and the U.S. Census Bureau." As such, PSEO promises to provide to its end users key earnings data in comparative perspective, allowing academic subject-matter distinctions to be utilized in assessing future earnings prospects, earnings growth over time, and in cross-institution comparisons.

- 2. CSAC recommends that the Census Bureau hold a one-way briefing to present its plan for the further expansion of the Post-Secondary Employment Outcomes program beyond its current coverage in states (n=17) and their institutions of higher education. As part of this briefing, CSAC recommends that the Bureau explain the relationship between the PSEO program and the US Department of Education's College Scorecard (<https://collegescorecard.ed.gov/>).**

## **Differential Privacy Working Group Deliverable - Task 4**

The CSAC Working Group on Implementing Differential Privacy for the 2020 Census Data Products was tasked to deliver comments on four specific tasks laid out in our convening charter. The Working Group presented recommendations on the first three tasks at prior meetings. At the Spring 2022 CSAC meeting, the Working Group presented recommendations associated with "Task 4: Developing strategies for communicating the use of differential privacy for the 2020 Census data products."

The working group's task was to provide recommendations for messaging and communication strategies for informing data users, of all skill levels, about what differential privacy is, how the specific disclosure avoidance algorithms implemented in the 2020 Disclosure Avoidance System

(DAS) work, the data protection the 2020 DAS provides, the privacy/accuracy tradeoff, and the strengths and limitations of the 2020 Census data products.

CSAC recognizes that public conversations about implementing Differential Privacy (DP) for the 2020 Census have raised serious concerns about both data accuracy and privacy. Such ongoing conversations in the press, among data users, and among professionals, risk eroding public trust in the Census Bureau. For example, some users are confused as to why more data cannot be held invariant when it is difficult to imagine any risk of disclosure from a particular table/value being released, such as county population totals. Users may not find it intuitive, and may (correctly or incorrectly) call the application of differential privacy into question. This broader context sets the stage within which Bureau communications around differential privacy must take place now and for the foreseeable future.

3. **CSAC recommends that the Census Bureau consult with stakeholders to identify common concerns about DP and develop concise and clear public responses. In doing so, the Census Bureau should consider and respond to broader public discourse around differential privacy. This might be done by expanding the “Latest Frequently Asked Questions” (<https://www.census.gov/programs-surveys/decennial-census/decade/2020/planning-management/process/disclosure-avoidance/2020-das-updates/2020-das-faqs.html>) to catalog direct responses to frequently expressed public concerns.**
4. **CSAC recommends the Census Bureau develop communications that clearly and concisely explain why aggregated population totals for relatively large geographies, including counties, could not be held invariant and why differential privacy protection is necessary for such totals.**
5. **CSAC recommends that the Census Bureau collaborate with a partner organization to create a public forum (data user group) where users can ask/answer questions about using differentially private Census 2020 data and engage with one another with regard to how they use census data. Such a forum might look similar to the American Community Survey (ACS) Data Users Group. This would be a good opportunity for outreach and to engage directly with data users in between Census years on the data user community’s terms.**

CSAC commends the Bureau on its efforts to increase transparency and to communicate clearly about differential privacy with various users.

6. **CSAC recommends that the Census Bureau prepare and publish a detailed report on the public engagement and iterative process that the Bureau engaged in while developing the 2020 DAS. Such a report would put the full story of DP development in one place and show the progress the Bureau made over time and how the Bureau adapted to public response, which would further increase transparency. It should include an executive summary that summarizes the process for a broad, general audience.**

- 7. CSAC recommends that the Census Bureau provide a regularly updated roadmap and timeline for coming data releases, including demonstration data as well as census data products, so that users can plan accordingly.**

Press coverage of overcount and undercount in the 2020 Decennial shows that the public is quite concerned about implications for underserved populations. An example provided by the Partnership for America's Children notes that the phrase "net undercount" suggests that the count worked well, when in fact a number of important populations experienced significant under-reporting. There is likely to be equal interest in the accuracy, bias, and outliers associated with race, ethnicity, and age (especially young children) in the Demographic and Housing Characteristics file (DHC) as well as potential confusion about sources of error in the census counts.

- 8. CSAC recommends that the Census Bureau develop communication materials that address multiple sources of error (undercount and DP in relation to one another) as they relate to race, ethnicity, and age.**

The Census Bureau specifically requested that the Working Group provide feedback on the handbook "Disclosure Avoidance for the 2020 Census: An Introduction," which was published by the Census Bureau in collaboration with the Population Reference Bureau (PRB) in November 2021. CSAC commends the Census Bureau and PRB for the work to produce this important handbook. It is a useful guide in explaining what DP is and how it works using clear and accessible language. It offers guidance and valuable practicalities, and is probably most useful for mid-skill users.

- 9. CSAC recommends that the Census Bureau and PRB make some revisions to the current version of the handbook and publish the revised version. The specific revisions CSAC recommends are summarized below.**

The handbook effectively describes the tradeoffs between privacy and data accuracy, describing how privacy risks have evolved over time. This is important, but it's also possible that readers could misunderstand exactly what information is at risk of disclosure from the Census Bureau. For example, a reader might assume that highly sensitive information, such as social security numbers, addresses, income, etc. were at risk. It is important to eliminate the potential for such misunderstanding.

- a. CSAC recommends that the handbook (and future versions thereof) recognize what information is collected in Census 2020 (e.g. age, race, sex, household relations), and so is at risk of disclosure.**

CSAC noted that the limitations of DP as a disclosure avoidance tool applied to the census are not described in the handbook.

- b. CSAC recommends that the limitations and a concise statement on the real world impacts (on release of data products, data accuracy, etc.) of applying DP be included, early in the handbook.**
- c. CSAC recommends that the handbook should review various sources of inaccuracy in the 2020 and prior censuses, to put differential privacy in 2020 in perspective.**

CSAC Working Group members carefully reviewed the handbook, but we are not the most appropriate group to offer detailed editing nor do we have the perspective of the variety of data users whose feedback should be considered. A full editorial review is outside the scope of CSAC.

- d. CSAC recommends that the Census Bureau commission a critical review to address audience usability from the perspective of diverse data users (skill levels) and offer feedback on detailed editing. Some specific editorial changes CSAC identified that should be revised include:**
  - i. p. 7 photo metaphor blurring is incorrect**
  - ii. The appendix should be reconsidered. Some of that material may be addressed through links to outside sources. Material that is included should be described with text interpretations so that the meaning and take-home message is clear.**
  - iii. Adding links to additional resources that are more appropriate for different users/audiences is one way to target different information to different audiences.**

Beyond the current handbook, CSAC would like to make recommendations about future communications and supplementary documents.

**10. CSAC recommends that additional supplementary documents and future handbooks that address Demographic and Housing Characteristics (DHC), additional use cases, and new data releases be created and published on a common website used to communicate about using Census 2020 data.**

**11. CSAC recommends the Census Bureau make a more general and concise version of the handbook for a novice audience/data user.**

CSAC found the sections of the handbook that discuss fitness for use and related examples to be particularly helpful.

**12. CSAC recommends that for future publications these examples be expanded to include more use cases and to include a discussion on how to assess fitness for use in custom geographies (e.g. aggregating blocks or block groups). One way that the Bureau could offer such guidance would be to develop an online user tool**

**based on a model (fed by 2010 demo data in comparison to 2010 SF1) that estimates uncertainty associated with any data point. Such a tool might allow users to enter the geography (or group of geographies) of their choice, a variable of interest, and the cell value to inform the tool which would then estimate the level of uncertainty associated with that estimate.**

**13. CSAC recommends that in future communications the Census Bureau address systematically a variety of audiences with diverse skill/knowledge levels and how the Bureau can best engage with these audiences in different outlets. There are several ways this could be done.** Some suggestions include:

- This could be an opportunity for community engagement, perhaps in collaboration with the Partnership program, where the community relationship goes beyond convincing people to respond to the census, but now engages people in how the Bureau can best meet community needs for their own specific uses by providing data relevant to the community and communicating its fitness for use in ways that match community needs and understanding.
- The Bureau should consider the audience when presenting technical details. Some users need this information, others do not. Adding links to additional resources that are more appropriate for different users/audiences is one way to target different information to different audiences.
- The Bureau should engage more extensively with the press and be responsive to journalist requests. Engaging with the press is an important avenue through which to reach a very general audience.
- The Bureau should create one-page documents with specific guidelines, each of which addresses a different common use case and offers clear guidance on what data users should consider in terms of data accuracy (fitness for use), the need to aggregate geographies (or not), population thresholds, or what alternative tables may be available that could provide more accuracy (tables with fewer breakdowns but that address the same bigger issue). These one-page documents could supplement the handbook and might look similar to the current handbook p. 19, but for different kinds of users.

**14. CSAC recommends using call-out boxes (exhibits) to highlight examples or key issues, particularly in longer documents.**

**15. CSAC recommends that a summary of the limitations of DP and a concise statement of its real world impacts (on release of data products, data accuracy, etc.) be included in all future communications.**

**16. CSAC recommends that future communications products include several specific examples of use cases and talk through the implications of DP for those uses. Graphical representations should be used throughout communications whenever possible. The DHC file will require many more (and different) examples, in comparison to the redistricting file.**

## **Demographic Profile and Demographic and Housing Characteristics File (DHC)**

CSAC commends the Census Bureau for significant progress in the release of demonstration products related to the Demographic Profile and Demographic and Housing Characteristics (DHC) file. Demonstration products related to the DHC “person” tables were released on March 16, with a webinar scheduled for March 22. The DHC “housing” tables will be released in April. CSAC is impressed with the extent of the data to be released as part of the DHC demonstration products (114 Person File tables, 121 Housing Unit File tables, 66 geographic levels). There will be a lot to digest. In the March 16 notice, the Bureau proposed a comment period of 30 days after the April release.

- 17. Because of the importance of the DHC tables, the amount of detail released, and the fact that although “housing” tables were included in the first demonstration products, they were not included in the second set of demonstration products, CSAC recommends a comment period of 45 days after the April release.**

CSAC congratulates the Bureau on the substantial reduction in net mean error at the county level compared to earlier demonstration products.

- 18. CSAC recommends that the Bureau provide commentary on why the reductions are so much greater for “persons” than “units.”**
- 19. CSAC also recommends that the Bureau share summary tables that report other measures of accuracy, bias, and outliers for counties.**

CSAC understands that the newly released demonstration products are and will be responsive to comments and use cases submitted last fall, at least in part. More than 400 were submitted! CSAC appreciates the careful review these are receiving. The March 16 notice calls out specific areas that the Bureau is working to address, including additional accuracy for ages 18-25 for the group quarters population, for centenarians, and for relationship to householder for small populations (e.g., foster children or same-sex unmarried partners). CSAC understands that after a careful assessment of cost for privacy and accuracy, the Bureau may not release the detail needed for all desired uses of the data. There is continuing concern about whether the levels of geography needed by those who produce estimates and projections for planning purposes will be available. In this connection, CSAC is pleased to see that the Bureau is considering options to share information from comments it received. CSAC believes that information about use cases and their disposition—i.e., whether supported or not--will be very helpful to users.

- 20. In the interest of transparency, CSAC recommends that the Bureau share information about use cases, both those supported by special tabulations and those which are not.**



CSAC enthusiastically welcomes the news that planning is underway to develop a Public Use Microdata Sample (PUMS) based on the 2020 Decennial. PUMS data are a vital national resource, used extensively in demographic, economic, and social research, and important to policymakers and policymaking.

**21. CSAC recommends that the Bureau share its plans and timetable for a PUMS file as soon as possible.**

CSAC also appreciates the presentation by Michael Hawes, “Reconstruction and Re-identification of the Demographic and Housing Characteristics File (DHC)”. Unfortunately, CSAC was not able to fully assimilate the information provided given the very short time frame. This presentation contained new information about the Reconstruction Attack and its efficacy that does not seem to have been previously made public. There are questions about the Census Edited File and what it proxies, the consequences of restricting reidentification efforts to records in the CEF with a Protected Identification Key, and that providing more disclosure protection (cutting rho in half) buys little reduction in risk. There has been a lot of interest in understanding attacks and attack efficacy, but it is difficult for the broader research community to evaluate, as the original dataset and more importantly, the details of the 2010 DAS are not public. This makes the Bureau uniquely placed in terms of conducting these studies. Sharing the results of these attacks is therefore all the more important.

**22. CSAC recommends that the Bureau present results of the reidentification and reconstruction exercise in a one-way briefing so that CSAC could better understand the methods and results.**

**23. CSAC recommends that the Bureau publish a detailed research paper describing the results of their reconstruction experiment.**

### **Detailed Demographic and Housing Characteristics File (Detailed DHC)**

CSAC was pleased to receive a thorough and detailed update of the Detailed Demographic and Housing Characteristics File (Detailed DHC). CSAC recognizes the tremendous amount of work involved by multiple groups over the past year. The Detailed DHC presentation provided information relating to the development of detailed race, ethnicity, and American Indian, Alaska Native tribes and villages counts summarized in individual-level Tables T1 (total population counts) and T2 (sex by age for selected age categories), as well as household-level tables T3 (household type) and T4 (household tenure). The presentation outlined the prioritization of Tables T1 and T2 and provided a timeline and series of next steps for development, differential privacy adjustments, and public engagement through communication and demonstration product releases. Next, representatives from Tumult Labs provided details on the differential

privacy process for Tables T1 and T2 (supplemented by a technical report with additional details on specific steps). Finally, the Bureau summarized the development of model-based adjustments to provide “joins” linking individual and household summaries. The presentation outlined the challenges in applying differential privacy approaches to the very wide variety of uses (some unknown and unexpected) to which the joined data will be applied.

CSAC appreciates the detailed updates on the development of the Detailed DHC and agrees with the prioritization of the completion of Tables T1 and T2 over Tables T3 and T4. CSAC notes that the brief summary provided comparing the numbers of tables and geographies considered in the 2010 Census to those from the 2020 Census needs additional details (such as those provided in the 2020 Crosswalk Summary) in order to fully communicate the proposed changes to future data users.

**24. CSAC recommends public summaries of 2010 to 2020 data products contain links to the 2020 Crosswalk Summary information and additional detailed descriptions in order to allow interested users to identify clearly which 2010 data products have 2020 analogues and which do not.**

In addition to the general information and technical details provided, the Bureau posed three specific questions to CSAC relating to the Detailed DHC, namely:

- What parameter settings (e.g., MOEs, levels of geography) are acceptable to meet data user needs?
- Does CSAC have any feedback on the Tumult Labs methodological work for developing tables T1 and T2?
- Does CSAC have any feedback on the methodology to model Join tables?

Regarding Question 1, CSAC notes that Tables T1 and T2 are proposed for the following levels of geography: Nation, state, county, American Indian, Alaska Native, and Native Hawaiian (AIANNH) territories. CSAC notes that there is wide heterogeneity in population sizes (and associated margins of error (MOEs)) within each level of geography and, as a result, it is difficult to define uniform guidelines for MOEs within each level of geography.

In addition, CSAC notes that smaller levels of geography (in particular, census tracts and minor civil divisions) are widely used in administrative applications and for allocation of local funds, and there will continue to be great demand for census tract and minor civil division data. CSAC agrees with summary statements by the presenters that there is a need to provide data at smaller geographies. In addition, CSAC notes that some counties have smaller population sizes than most census tracts and that these low population size (high MOE) counties are not uniformly spread across the nation, so that impacts on performance of Detailed DHC data products are likely to vary geographically. As a result, CSAC notes that Question 1, defining MOEs and levels of geography for best performance, is best addressed in the context of families of use cases and in the context of geographic variations in this performance. Such discussions relating to Question 1 would best be addressed in a series of focused, ongoing follow-up briefings and meetings with CSAC.

**25. CSAC recommends scheduling one-way briefings relating to the choices of MOEs and levels of geography in the context of specific categories of use cases. Such briefings should include a clear discussion of the roles of the MOEs in defining fitness-of-use for the differential privacy in order to allow CSAC to frame detailed recommendations.**

Regarding Question 2, CSAC appreciates the summary presentation and accompanying technical information regarding the development of Detailed DHC Tables T1 and T2, however, allowing CSAC more time to review materials would allow for fuller discussion in future CSAC meetings. CSAC agrees with the proposed prioritization of Tables T1 and T2 (individual level tables) over Tables T3 and T4 (household tables). The information provided by the technical documentation is quite helpful in explaining the details. The discussion of the “adaptivity” criteria raised the possibility of different levels of aggregation for different population subgroups. The idea is appealing but CSAC would benefit from a detailed one-way briefing on such adaptivity might be built into data releases.

**26. CSAC recommends a one-way briefing on plans for incorporating adaptivity into the disclosure avoidance system and its implications for anticipated data product releases and timelines.**

CSAC found the model-based approach to calculating the join tables builds on past Bureau excellence in the development of small-area estimation techniques. The presentation and documents illustrate how model-based approaches can serve as a hybrid estimation scheme building off of elements developed under strict privacy budgets might be extended to provide model-based predictions of totals and proportions associated with other geographies such as census tracts and other, custom geographies. The presentation noted (correctly) that the full range of anticipated uses cannot be fully enumerated (making formal privacy budgets difficult to define) and that there is great demand for tract-level data. CSAC encourages further development of the approach with particular attention to formulating and communicating what data products (especially at the census tract level) the model approach would enable. The modeling approaches also offer an opportunity to provide model-based margins of error with reported values (similar to the MOEs provided with ACS population counts) and CSAC looks forward to the upcoming focused briefings providing further details on the approach and its anticipated uses) in the coming months.

**27. CSAC recommends the Bureau provide updates on technical details, anticipated data releases, and timelines for the model-based results in upcoming briefings to CSAC relating to the model-based approaches. Invitations for upcoming briefings on model-based approaches should be extended to all members of CSAC.**

## **Update on the 2020 Post-Enumeration Survey (PES)**

CSAC commends the Census Bureau for completing a very credible 2020 PES despite numerous challenges stemming from substantial delays in field operations. While there is ample evidence that these delays affected the quality of data collected in the PES interviews in several respects, the PES estimate of the national net undercount is close to that of the Demographic Analysis (DA) middle series, and the estimates of correct enumerations, erroneous enumerations, and omissions are similar to the estimates for the 2010 Census.

Despite targeted efforts to improve the count of young children (0 to 4), the census continues to undercount this population. The PES estimate of the net undercount, at 2.79 percent, is just over half the estimate of 5.4 percent obtained from DA. That DA is most accurate for young children implies that the PES estimate of the net undercount of this population is low, albeit improved over 2010, when the coverage measurement program estimated a net undercount of 0.72 percent compared to 4.6 percent for DA. A 2019 internal planning document on the PES design, recently released to the public as Memorandum 2022.06, indicated that “correlation bias adjustment factors are currently being researched for young children and Hispanic origin.”

**28. CSAC recommends that the Census Bureau publish the results of this research and, if an adjustment for correlation bias among young children was implemented, document the adjustment and, if not, why not.**

If an adjustment for correlation bias among young children was included in the 2020 PES, it was clearly insufficient. While improving the quality of the PES estimate of young children ranks much lower in priority than improving the enumeration of this population, obtaining more accurate estimates of the undercount of young children is nevertheless important.

**29. CSAC recommends that the Census Bureau seek to develop a better adjustment for correlation bias among young children for the 2030 Census.**

The PES estimated a net undercount of persons of Hispanic origin of 4.99 percent, which is more than three times the estimated undercount of this population in 2010 (1.54 percent). The magnitude of the undercount reflects unique challenges in counting this population in 2020; nevertheless, it raises serious concerns about the quality of these data and carries potential implications for federal funding allocation affecting members of the Hispanic population. Understanding how the net undercount is distributed will be important to addressing potential implications, but as with the 2010 Census, state-level PES estimates to be released this summer will not include breakdowns by demographic groups. In 2010, no state coverage estimate was significantly different from the national estimate. However, comparisons of 2020 census counts with state population estimates suggest that this may not be true in 2020. If state differences are observed, Hispanic undercounts could be a contributing factor.

**30. CSAC recommends that if state estimates show significant differences in net coverage error, and one or more states with large Hispanic populations show larger than average undercoverage, the Census Bureau should estimate and**

**publish the contribution of Hispanic undercoverage to states' total undercoverage.**

Because the onset of the pandemic in 2020 resulted in colleges sending their students home, the task of enumerating college students once and only once and in the right place was made very difficult. With improved unduplication implemented in the 2020 census, duplicate counting of college students should have been less of an issue than it would have been if a pandemic had occurred during an earlier census. Nevertheless, DA estimates indicate an overcount of 3 to 4 percent among persons 18-24. The PES shows an undercount of 1.62 percent for the somewhat broader population 18-29, excluding persons in group quarters, who account for a large portion of the college population. That the PES cannot support substate estimates, despite its similarity to the 2010 design (which included such estimates), eliminates the possibility of using estimates of the net coverage error in counties with large college populations to assess how well the 2020 Census performed in counting college students in the right place.

**31. CSAC recommends that the Census Bureau use all data sources available internally to assess the accuracy of the 2020 Census counts in counties and, if possible, smaller places with large college populations.**

The Census Bureau has asked CSAC for recommendations regarding alternative PES designs or changes to lower the risk of field or clerical problems and has specifically asked CSAC how the PES might use administrative records to help measure census coverage. PES operations are extensive, complex, and highly integrated. They have been developed over a period of decades with contributions from some of the Census Bureau's leading researchers. While CSAC may be able to offer more substantive suggestions given sufficient time, at the present we can note only that we are struck by the apparent extent of the Bureau's continued reliance on clerical matching, which is time consuming and may introduce new sources of error while tackling matches that could not be reliably handled by computerized techniques.

**32. CSAC recommends that the Census Bureau explore the development of more sophisticated computer matching procedures to reduce the PES's dependence on clerical matching.**

The 2020 Census Experiments and Evaluations include an evaluation of dual-system estimation using administrative records in place of the PES. The Census Bureau has previously explored (and may be continuing to do so) triple-system estimation, in which administrative records provide a third source to supplement the census and PES. Much more is known about the strengths and limitations of administrative records for census uses than was the case when triple-system estimation was first proposed. In particular, despite their extensive coverage, administrative records do not cover the population uniformly and do not include the information needed to exclude persons in group quarters, both of which complicate their use in the PES. At the same time, correlation bias per se is less of an issue with administrative records than with a replication of census methods, although other biases in administrative records cannot be ignored.

- 33. CSAC recommends that the Census Bureau provide CSAC with an update on its experiment(s) to incorporate administrative records into its future census coverage evaluation.**

## **Seasonal Adjustment of Time Series during the Pandemic**

CSAC thanks the Census Bureau for providing a detailed and well-motivated presentation of the methodology used for outlier removal and seasonal adjustment during the COVID-19 pandemic. Overall, CSAC finds the adopted approach to outlier removal for seasonal adjustment both well-reasoned and practical.

- 34. CSAC recommends that the Census Bureau detail suitable criteria for evaluating the quality of seasonal adjustment procedures deployed in anomalous periods, selecting amongst competing procedures, and declaring a procedure suitable for deployment.**
- 35. CSAC recommends that the Census Bureau publish an evaluation of the selected outlier removal procedure and any competing procedures considered using the aforementioned criteria.**

Several possibilities for evaluation include:

- Judging the accuracy of a procedure on prior anomalous periods (e.g., the period following Hurricane Katrina for state food stamp uptake) with agreed-upon ground-truth seasonality,
- Judging the accuracy of a procedure using standard simulation settings with realistic anomaly patterns and known seasonality, and
- Judging the accuracy of a procedure using a semi-synthetic evaluation. For example,
  - Begin with real time series data with no expected outliers, estimate seasonality in the usual way, and treat the estimated seasonality as ground truth
  - Introduce a synthetic outlier sequence (e.g., modeled after COVID-19 pandemic behavior), and
  - Evaluate how well each method reconstructs the ground truth seasonality in the presence of these outliers.

Overall, CSAC also finds the Census Bureau's approach to testing for the end of COVID-19 pandemic effects both well-reasoned and practical.

- 36. CSAC recommends that the Census Bureau detail suitable criteria for evaluating the quality of procedures for detecting the end of anomalous periods, selecting amongst competing procedures, and declaring a procedure suitable for deployment.**
- 37. CSAC recommends that the Census Bureau publish an evaluation of the selected procedure for detecting the end of anomalous periods and any competing procedures considered using the aforementioned criteria.**

Several possibilities for evaluation include:

- Judging the accuracy of a procedure on prior anomalous periods (e.g., the period following Hurricane Katrina for state food stamp uptake) with agreed-upon end dates
- Judging the accuracy of a procedure using standard simulation settings with realistic anomaly patterns and known end dates
- Judging the accuracy of a procedure using a semi-synthetic evaluation. For example:
  - Begin with real time series data with no expected outliers
  - Introduce a synthetic outlier sequence (e.g., modeled after COVID-19 pandemic behavior)
  - Evaluate how well each method detects the end date of the anomalous period

CSAC agrees that it may be difficult to disentangle seasonal changes from anomalous COVID-19 pandemic shifts without more post-pandemic data. However, some speculative possibilities for determining if seasonal patterns have changed include:

- Sharing strength across multiple time series to estimate seasonal patterns and detect changes
- Identifying external indicators of a target variable's seasonality from other time series or data sources
- Leveraging sectors that are known to be back to normal (that is, no longer significantly impacted by COVID-19 pandemic effects) and the historical seasonality relationships across sectors
- Leveraging data from countries and regions that were impacted by the COVID-19 pandemic at different times

## **Impact of Alternative Inflation Adjustments on Current Population Survey (CPS) Annual Social and Economic Supplement (ASEC) Income Statistics**

CSAC thanks the Bureau for bringing this timely topic for discussion.

CSAC agrees with the Interagency Technical Working Group (ITWG) recommendation that a chain-type price index has less bias than a traditional Consumer Price Index (CPI), and therefore more accurately captures the true impact of inflation on consumers, and is therefore a better statistical choice of index than a traditional CPI.

Given the widespread use of this data, and the limited public feedback to date, CSAC is strongly supportive of the Bureau robustly engaging stakeholders to both confirm the consensus around adoption of a chain-type price index, and to further develop a consensus about the specific chain-type price indices to use over time.

Moving to a chain-type price index will have the effect of making median household income growth over time look better than with the current index. Consumers may find it perplexing that substitutions for lower price goods are positively associated with increases in income. Given that the United States is currently experiencing 40-year high inflation levels, there may be significant public scrutiny of methodology changes, and concern about the timing of changes, especially if those changes have the effect of reducing the stated impact of inflation on federal statistical measures of well-being.

While CSAC's concerns are primarily with the scientific considerations, public confidence in the motivations behind changes to reporting methodologies is also crucial to the Bureau's mission. Given that reality, CSAC would emphasize:

- The importance of building as much of a scientific and statistical community consensus as reasonably possible before making changes to the index used in primary reporting; and including endorsements of these changes in the public communication.
- The importance of providing unadjusted numbers, in addition to inflation-adjusted numbers, so researchers and other consumers of the data can make their own choices about inflation adjustments.
- The careful consideration, in light of current inflation trends, of what timeline is optimal for deploying an index change, and to what extent this is deployed as a change in methodology, versus reporting on additional indices.

**38. CSAC recommends the Census Bureau explicitly consider the pros and cons of implementation next year versus this year.**

**39. CSAC also recommends the Census Bureau clearly articulate both “why” and “why now” in public communications if/when the Bureau implements an index change.**

With regard to the choice of the Personal Consumption Expenditures Price Index (PCEPI) versus the current index for the period prior to 2000, CSAC notes that the PCEPI tracks the Chained-CPI-Urban Consumers (C-CPI-U) more closely than does the current method. Use of the PCEPI for the earlier years appears more appropriate than the current method, which would have the effect of linking two disparate series to generate long-term adjustments.

**40. CSAC recommends that the Census Bureau use the PCEPI as the inflation adjustment for the years prior to 2000.**