

**Reengineering the Census Bureau’s Annual Economic Surveys: Annual Integrated Economic Survey (AIES)
Census Bureau written responses to Committee comments and questions**

Committee questions:

There is concern that business data collection is company based and establishment level data will be lost.

Although the sample is company-based, AIES will request data at the company, industry, and establishment levels. Data will be collected at the establishment level to produce select geographic data. Key data items for all four trades will be collected at the establishment level. Manufacturing will continue to collect all data items at the establishment level.

What is the impact on total factor productivity (TFP) measures from the Annual Survey of Manufactures (ASM)? For historical consistency, will the new system be flexible enough to continue TFP analysis?

AIES will produce the same manufacturing data as the ASM; manufacturing data will be collected at the establishment level. AIES outreach includes both internal and external stakeholders. The Census Bureau’s Center for Economic Studies has representation on appropriate AIES teams and will continue to be consulted to reduce or eliminate impacts to programs.

How will AIES and the Economic Census be coordinated?

Census is standardizing and aligning with the economic census where possible for all aspects of the survey lifecycle. Over the last year, Census launched a large-scale effort to align content between AIES and the economic census. Other projects include but are not limited to aligning variable names, edits, the use of enterprise solutions, analysis plans and tools, and dissemination. Once AIES designs publication tables and understands time series impacts we will define a benchmarking strategy to benchmark to the economic census.

How will Census maximize the use of administrative data and 3rd party data?

Census has been evaluating administrative data and alternative data sources for AIES, the economic census, and other programs. We will continue to evaluate and identify 3rd party sources to use as a source of imputation for nonresponse or an alternative source for response data, as well as an input for model-based statistics (e.g., small area estimates). AIES is planning to survey all sampled units; we are not planning to impute with administrative data for unsampled companies such as is done in the economic census.

How do you plan to resolve or minimize multiple reporters in the sample unit?

We expect that larger companies will have multiple reporters for the same form as this is common practice now. We did consider the 2018 National Academy of Sciences report recommendation for a two-stage sample design: a PPS sample of companies within sector, followed by a SRS-WOR sample of establishments within companies. However, the required sample maintenance activities for such a design appeared to offset any potential reduction in response burden, as sampled establishments could go out of business or be idle. Furthermore, this approach makes it difficult to objectively compute response rates (how many sampled establishments must respond for the company to be considered a respondent) and has proven quite difficult to implement in other economic census programs such as the Commodity Flow Survey.

Each existing survey has a sample size – will there be mismatches to AIES with company as sample unit?

We are not attempting to achieve concordance between the AIES sample size and the consolidated surveys' sample sizes and cannot guarantee that the AIES sample size will equal the union of the six surveys' sample sizes that it is replacing. In fact, we expect that the AIES sample size will be larger so that this unified survey can achieve both the national industry and geographic industry reliability constraints. In addition, we are planning for some expected nonresponse in our allocations, a practice that has not been implemented historically in our economic surveys. Finally, it is important to note that the ASM is the only current survey with establishment as the survey unit. Three of the remaining surveys include entire companies with survey and subsample tax-units (EINs) and two surveys sample entire companies.

It is already difficult to determine who is the respondent for the survey. Does this become more complicated with consolidation?

Census is looking across programs for the best contact, and in some cases, we may need multiple contacts. Census plans to utilize what we learned from selecting contacts during the AIES Coordinated Collection Pilot. Delegation functionality will be available in the collection instrument. Split mailing for companies is also an option to reach multiple contacts.

How does the Census Bureau's organizational structure impact the structure of the survey organization?

The Economic Directorate was reorganized in 2014, the affected trades for this survey all reside in the same division now.

What are the implications for coordination with monthly surveys and data collection methods?

Indicator survey areas have representation on AIES teams and are consulted on a regular basis. Census is working to incorporate indicator area requirements into the sample design. Indicator areas will be heavily involved in reviewing time series implications. Census is working to determine if monthly surveys will be subsamples of the AIES sample (as they are currently subsamples of the annual surveys) or if they will be independent samples.

We are planning to continue the annual benchmarking procedures and are investigating the relationships between annual samples and indicators samples.

Are there potential issues with using payroll as a measure of size?

There are definite drawbacks to using annual payroll as our building block for measure of size such as weaker correlation with receipts in selected industries as well as for smaller businesses. However, there are advantages to using payroll. First, recent administrative data are available for all establishments on the frame; other variables, such as total receipts, are only available at the company level. Second, there are extensive validated edit procedures in place for annual payroll, including comparisons to COS data.

As part of the sample design, we create national industry rosters and geographic industry rosters for each company, with each unit on the roster representing the company's total payroll in the designated NAICS6 or NAICS4 by geography category. The company inclusion probabilities account for the company's contribution to

either set of domain estimates. Consequently, the sampling design resembles a probability proportional to estimated size (PPES) rather than being a strictly probability proportional to size (PPS) design.

The adjusted AIES sampling weights will account for unit size with respect to annual payroll, as AIES will produce two separate sets of final weights for each sampled unit: (1) ratio adjusted or calibrated weights that produce unbiased estimates of total annual payroll at the national tabulation level (TBD, but likely NAICS5 or NAICS6) and (2) ratio adjusted or calibrated weights that produce unbiased estimates of total annual payroll at the geographic industry tabulation level (TBD, but likely NAICS 3 or NAICS4 by geographic group). At this time, we do not believe that similar calibration to total receipts (accounting for output) is possible for the reasons discussed above.

How does AIES mesh with BLS job surveys? What discussions have been held regarding integration of AIES with the BLS job surveys?

Census consults with BLS on a continuous basis to identify critical requirements for content needed for collection and other aspects of the survey life cycle; BLS has provided input to the content for AIES. Census provides regular updates to BLS staff on the progress of AIES.

Industry classification codes don't always correspond with how businesses define their activity. Can suggested changes be submitted to help refine NAICS codes and their associated definitions?

We agree some companies have difficulty with classification; we work with these companies on a case by cases basis to ensure proper classification and reporting. AIES plans to use the 2017 NAICS definitions for the initial release and will update using the 2022 NAICS in future releases. Questions regarding current (and former) industries and definitions are typically handled either by email (naics@census.gov) or phone (888) 756-2427. This contact information is listed at the bottom of the main page of the NAICS website (census.gov/naics).

Comments geared towards creation of new industries or revision of definitions (for 2027, the next revision cycle) are not formally recognized until the revision process begins, which won't be for another few years. But in the meantime, those types of comments can be sent to econ.naics2022@census.gov.

What about changes in NAICS coding and data continuity?

NAICS coding is an existing challenge. Census continues to refine instructional text in collection instruments to assist respondents. Ongoing cognitive testing and research continues to gather information regarding opportunities to clarify NAICS definitions for respondents. The sample design team is aware of the disconnect between versions of NAICS. Possible mitigation strategies include more frequent resamples or including rotation groups in the sample design. The most recent plan is for a cross-sectional sample. However, if the sample size requirements are prohibitive, or the planned NAICS revisions are extensive, rotation groups become more likely. Census will continue to communicate NAICS changes to data users on its website through NAICS concordances and via the planned AIES methodology report.

What about changes in the sample of establishments, or their change in ownership, and data continuity?

Economic programs are heavily skewed, so the consistently largest companies are generally included in the sample and have continuity between sample revisions; the actual establishment composition will change due to mergers, divestitures, acquisitions, etc. Each new sample will be independently selected, however; we are not attempting to introduce any longitudinal requirement to ensure that the most recent sample is reflective of the economy at the time of sample selection.

What are the plans for metadata collection? The identity of respondent should be included.

Contact information is collected and retained. Census internal files cannot be shared, including the contact information for the responding company.

There seems to be optimism that integrating surveys will improve response rates. Is there evidence that integrating these 6 surveys will improve response rates?

We anticipate reducing response burden will help with survey response rates; sample design and informing respondents about changes should help with response rates. Furthermore, since the design is coordinated across all sectors, we can impose time-in/time-out constraints on smaller units that were not available with the other disaggregated designs.

When Census conducted nonrespondent debriefing interviews, one of the common issues was “data dispersion”, the concept that a single person at a company may not have access to all the data being requested. This is a barrier to completion. By combining the surveys and grouping the questions by topic, and then providing the opportunity to delegate those topics to the appropriate person, we are hoping to mitigate the responses loss to dispersion (that is, making it easier to get the right questions to the right person at the company). These same debriefing interviews identified "mixed topics" in surveys as an issue; again, respondents may only have access to some of the data, so by combining surveys and then slicing by topic, we may cut down on nonresponse due to multiple topics in a survey. In other interviews, respondents mentioned they were unaware they had not completed the totality of our request; that is, they answered some forms and not others, either because they don't notice that there was another form to complete, or they found additional forms duplicative and so don't answer them. Again, by combining the surveys, we are cutting down on the number of "entry points" to our requests, everything is in one place where the respondent can find the entire request together.

AIES cognitive testing that is currently underway is resulting in positive feedback. Respondents like the idea of having all the requests for data come at the same time so their response process, gathering the requested data, can be done at one time. This was also identified during previous focus groups that respondents wanted to provide data for the entire company at one time instead of piecemeal or Census not requesting data for some portions of the company.

Census plans to learn about response through the SY21 AIES Pilot survey and debriefings. Through information obtained from respondents we can identify obstacles and develop mitigation strategies to support response rates during production.

Census has learned quite a bit regarding response through the AIES Coordinated Collection pilot. The AIES Coordinated Collection pilot took place during calendar years 2020 and 2021 for survey years 2019 and 2020, for three current annual surveys: Annual Retail Trade Survey (ARTS), Annual Wholesale Trade Survey (AWTS), and

the Services Annual Survey (SAS). The purpose of this research was to investigate the impact of implementing a contact strategy that encompassed multiple survey requests. Respondents were assigned to various experimental arms within this initial research, and researchers conducted follow-up debriefing interviews to learn more about the impact of the modified contact strategy. This effort helped the team to better understand the processing, setup requirements, data collection operations, and response processes necessary to implement coordinated collection.

The first pilot, conducted in 2020, was mailed in January and February 2020 (2 separate groups to test both a combined and a staggered due date) with various due dates in February, March, and April 2020. The subsequent coordinated collection operation in 2021, again only with selected respondents assigned to two or more surveys within ARTS, AWTS, and SAS was mailed on February 16, 2021, and was due March 23, 2021.

Debriefing interviews with 2020 pilot respondents made it clear that the COVID-19 global pandemic had an immeasurable impact on both the experimental contact in the pilot and data collection writ large. This impact hindered our ability to fully test and understand the effectiveness of various contact methods; however, we did learn some lessons and documented improvements that were implemented for the next cycle, such as initiating email communication as soon as possible, contacting respondents via phone earlier than usual, and moving forward with one mailing and one due date.

Debriefings with non-respondents and partial respondents after the 2021 operations revealed continued struggle with COVID-19, signaling that we need to build additional time into the mail timeline to accommodate USPS changes and continued telework. Other respondent feedback is leading us to consider a pre-survey contact to alert respondents to upcoming changes, capture any structural changes for the company, and validate contact information as well as discuss building in additional delegation features and continuing to test a topic-driven approach. These actions would provide a more user-friendly instrument that allows the respondent to easily report data.

The Annual Integrated Economic Survey (AIES) Collection Team, in collaboration with the respondent research and other teams, will continue to consider respondent input while developing processes and implementing improvements for production AIES, set to mail out in early 2024. Along with the coordinated collection operations planned during 2022, there is an upcoming pilot survey and related research, and cognitive testing, all focused on ensuring that respondents are able to answer the survey, as envisioned, identifying potential barriers to completion, and evaluating response burden.

Response rates for the two years of coordinated collection operations carried out thus far are below:

2020 (SY19) final response rates:

Coordinated Collection overall response rate: **74%**

Average of the non-coordinated ARTS, AWTS, and SAS response rates: **69%**

(ARTS: 61%, AWTS: 77%, SAS: 70%)

2021 (SY20) final response rates:

Coordinated Collection overall response rate: **81%**

Average of the non-coordinated ARTS, AWTS, and SAS response rates: **68%**

(ARTS: 67%, AWTS: 79%, SAS: 67%)

Is Census planning to use the same estimation process for all sectors?

Yes, AIES will use the same estimation, imputation, and variance estimation process for all sectors.

In the long run, is there an opportunity to innovate on data collection?

Yes, efforts are underway to identify and implement innovative data collection processes as well as other aspects of the AIES survey lifecycle, including processing, analysis, and dissemination, including the identification of alternative data and other 3rd party sources.

How do you determine the contact with sampling and sending surveys at the company level (to whom do you send the survey?)

Census is looking across programs for the best contact; in some cases, we may need multiple contacts. Census plans to utilize what we have learned from selecting contacts during the AIES Coordinated Collection Pilot. Also, delegation functionality will be available in the collection instrument. Split mailing for companies is also an option to reach multiple contacts.

Outreach for additional feedback to others beyond BEA, BLS, etc. would be helpful.

Census has presented on AIES at FCSM, CSAC, and FCSM in addition to the BEA, the BLS, and the FRB. A stakeholder engagement plan has been developed to identify and document outreach to key federal stakeholders, internal stakeholders, and other external stakeholders. There are plans to publish a federal register notice, create AIES webpages, pod casts, and various other outreach efforts. Outreach suggestions are welcome.

Public questions:

To what extent and how are Census, BEA, and BLS contributing to their respective departments' 2022-2026 Learning Agenda, as required by the Evidence Act? The Commerce agenda draft emphasizes improving data collections.

The Department of Commerce Evaluation Officer coordinates with all bureau-level evaluation officers and the Census Chief Data Officer is a member of the Commerce Data Governance Board. Within the Census Bureau, two groups have been formed to review and provide feedback on the Learning Agenda, Capacity Assessment and Evaluation Plan. The first group is the Evidence Building Working Group who works to highlight and coordinate evidence building efforts throughout the Census Bureau. This is documented on the following webpage: <https://www.census.gov/about/what/evidence-act.html>. The second group is the Census Data Governance Board that's described in their charter: <https://www2.census.gov/about/evidence-act/data-governance-board-charter.pdf>. Both groups meet monthly to discuss the implementation of the Evidence Act within Census and coordinate with the Department of Commerce, as needed.

Thanks to the Census staff for a very comprehensive presentation. My question today deals with sampling issues with company surveys, an area I'm quite familiar with from my prior work with Census surveys. After sample selection, there will be mergers, spinoffs, acquisitions, births, and deaths. Question: How will you preserve time series consistency of the industry/product data in a company-based survey, in contrast to an establishment-based survey? For example, will spin offs of certainty companies be kept as certainty

companies? What happens to acquisitions of non- sample companies? I know Census is aware of these issues, and I hope that the proposed solutions will be made available for comment well in advance.

These are challenging sample maintenance issues now for five of the six annual surveys covered by the AIES design and they are not handled consistently by all current programs. Certainly, having a single sampled company for an economy wide survey will “resolve” some of the inconsistencies, but timeliness in the updates (and historic corrections) will need to be addressed. Treatment of acquisitions, mergers, spinoffs, and deaths in an economy-wide sample has not been discussed yet; we will need a standard practice to reduce problematic reliability concerns due to stratum jumpers. As mentioned earlier, we are discussing sampling frequency. That decision needs to be made before AIES can determine whether an annual sample of births is needed. There will be an established process to incorporate organizational changes on the frame and provide them on a routine basis to AIES.