

SUPPORTING STATEMENT
U.S. Department of Commerce
U.S. Census Bureau
Annual Integrated Economic Survey (AIES)
OMB Control No. 0607-1024

PART B: Collections of Information Employing Statistical Methods

This revision request documents the modifications that transpired between the U.S. Census Bureau obtaining OMB clearance and the present moment. The revised portions of this document are highlighted in gray.

1. Universe and Respondent Selection

The target population of the AIES covers all domestic, private sector, non-farm employer businesses in the United States (50 states and the District of Columbia) as defined by the 2017 North American Industry Classification System (NAICS). Exclusions are foreign operations of U.S. businesses headquartered in the U.S. territories and most government operations (including the U.S. Postal Service), agricultural production companies and private households.

The Annual Integrated Economic Survey (AIES) is selected from a frame of approximately 5.4 million companies constructed from the Business Register, which is the Census Bureau's master business list. In the AIES, the company is the ultimate sampling unit (USU). Approximately 384,940 companies will be included in the sample. Of the sampled companies, approximately 36,500 are self-representing (included with probability = 1), based on the diversity of their operations. The remaining companies on the frame are stratified within sector by geographic category within 3-digit industry (NAICS) classification. This is an unequal probability sample, with company inclusion probabilities accounting for contribution(s) to both national and subnational estimates of annual payroll.

The measure-of-size (MOS) variable used for the AIES sample design is based on the administrative annual payroll, which is available for the majority of establishments on the AIES sampling frame and is generally positively correlated with the AIES key data items. The company-level MOS is determined by summing the establishment values within the company. Each company on the frame is assigned to a single sampling stratum and has a single sampling MOS. Each non-self-representing company on the sampling frame has a sampling industry, defined as its most detailed unique NAICS code.

To minimize erroneous industry classifications in the sample design, in-scope companies are classified as certainty using the complexity criteria described in the following table:

Table 1: AIES Designation of Initial Certainty Stratum by Complexity

Company operates in	Number of Sectors	Number of Industries*	Number of States
Certainty	2 or more	2 or more	Not considered
	1	3 or more	Not considered
	1	2	2+
	1	2	1
Noncertainty	1	2**	1
	1	1	1+

* 4-digit NAICS industry

**Sectors 44-45, 62 (selected pairs of industries)

To determine the *initial* sector allocations of companies in each sector frame, let

N_{NC} = total number of companies on the AIES frame that are eligible for sampling (noncertainty)

N_s = total number of companies in sector s on the AIES frame U_s that are eligible for sampling

N_{11} = total number of eligible agriculture sector companies in U_{11}

N_{21} = total number of eligible mining, quarrying, & oil & gas extraction sector companies in U_{21}

N_{23} = total number of eligible construction sector companies in U_{23}

N_{31-33} = total number of eligible manufacturing sector companies in U_{31-33}

N_{55} = total number of eligible management of companies & enterprises sector companies in U_{55}

n = initial total number of companies to be sampled in the AIES excluding those in sectors 11, 21, 23, 31-33, and 55.

The initial allocations of companies in the agriculture (11), mining, quarrying, & oil & gas extraction (21), construction (23), manufacturing (31-33), and management (55) sectors are

$$n_{11} = 250$$

$$n_{21} = 850$$

$$n_{23} = 3,500$$

$$n_{31-33} = 30,000$$

$$n_{55} = 100$$

The values for allocations in agriculture (11), mining (21), construction (23), and management of companies & enterprises (55) sectors are their respective sample sizes from the Annual Capital Expenditures Survey (ACES). Historically, the Annual Survey of Manufactures (ASM) has published state level estimates for manufacturing (31-33) sector. The relatively large allocation for this sector should yield domain estimates of comparable – or higher – reliability for the AIES than their ASM counterparts, while also accounting for the elimination of the previous survey's nonmail component.

The initial sector level allocations for the other sectors (n_s) are obtained with proportional allocation, with $n_s = n \left(\frac{N_s}{\ddot{N}} \right)$, where $\ddot{N} = N_{NC} - N_{11} - N_{21} - N_{23} - N_{31-33} - N_{55}$. Initial allocations may be further increased as needed to attain targeted coefficient of variations (CV) of 2% or lower on the frame analysis variables for each national tabulation industry level and targeted CV's of 15% or lower for geographic industry estimates.

Let U_{IC} be the set of companies on the AIES frame that are included in the initial certainty complexity stratum, so that U_{NC} is comprised of the remaining companies on the AIES frame. The remaining companies are split into 19 separate sector (2-digit NAICS) frames (U_s) from the companies in U_{NC} .

The within-sector stratification procedure is applied to all sectors except agriculture (11), mining (21), construction (23), and management of companies and enterprises (55). As these sectors do not have geographic publication requirements, their initial strata are defined by 3-digit NAICS.

For the remaining sectors, the noncertainty USUs (companies) are stratified into:

- A single diverse company stratum (1) comprising companies that operate in two different 4-digit NAICS industries and one state (sectors 44-45 and 62 only)
- One multi-state stratum per 3-digit NAICS industry¹ comprising companies that operate in one 3-digit NAICS industry and 2 or more states
- Twenty-seven geographic substrata per 3-digit NAICS industry comprising companies that solely operate in the identified 3-digit NAICS industry and the geographic substratum: see table 2.

Table 2: Geographic Strata for the AIES Sampling Design^{1, 2, 3}

Region	Total States	Direct Use States	States Included in Balance of Region
Northeast	9	Massachusetts, New York, New Jersey, Pennsylvania	Connecticut, Maine, New Hampshire, Rhode Island, Vermont
Midwest	12	Illinois, Indiana, Michigan, Minnesota, Missouri, Ohio, Wisconsin	Iowa, Kansas, Nebraska, North Dakota, South Dakota
South	16	Florida, Georgia, Maryland, North Carolina, Tennessee, Texas, Virginia	Alabama, Arkansas, Delaware, District of Columbia, Kentucky, Louisiana, Mississippi,

¹ Services industries incorporate taxable and tax-exempt categories into the strata definitions, essentially doubling the number of strata for these industries. Wholesale trade industries incorporate type of operation category into the strata definitions, essentially tripling the number of strata for these industries.

² Each Direct Use state defines a separate geographic substrata. The combined *Balance of Region* states for each region define one geographic substrata.

³Substrata determined after analysis of (1) Business register establishment counts, (2) 2020 decennial census population counts, and (3) Annual Business Survey cell suppression counts.

			Oklahoma, South Carolina, West Virginia
West	13	Arizona, California, Colorado, Oregon, Washington	Alaska, Hawaii, Idaho, Montana, Nevada, New Mexico, Utah, Wyoming

Figure 1 illustrates the initial stratification procedure.

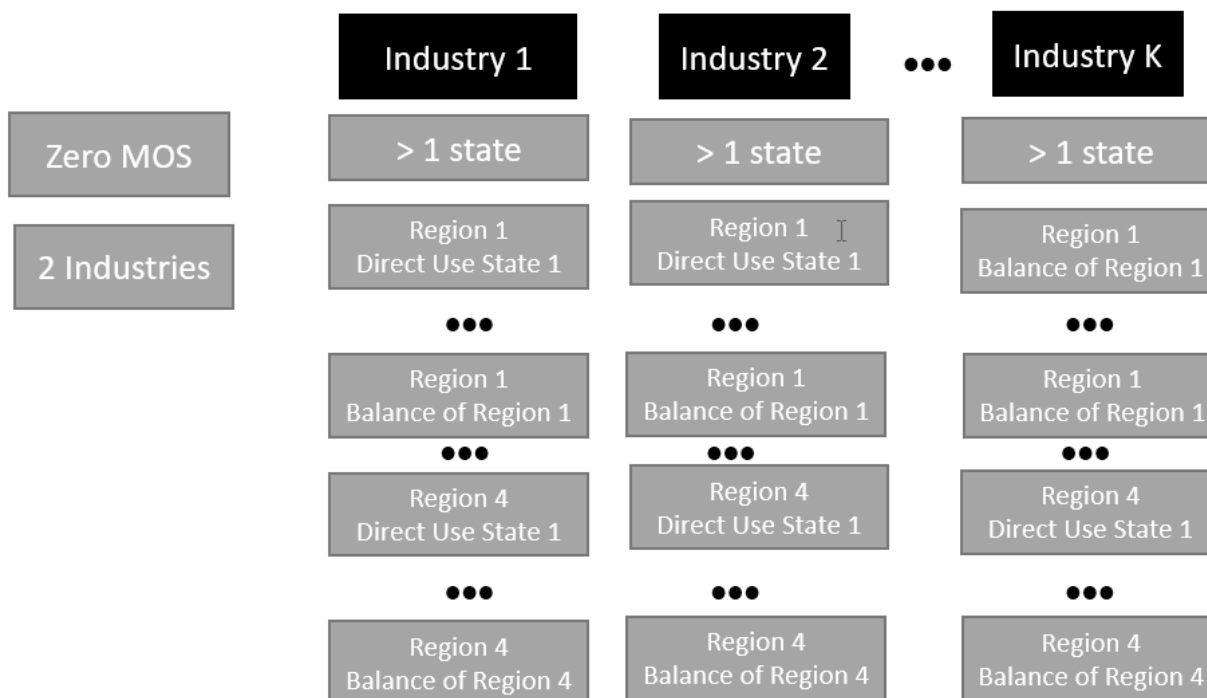


Figure 1: AIES Stratification Within Sector for all sectors but agriculture (11), mining (21), construction (23), and management (55). The Zero MOS contains all companies in the sector with a MOS value of zero. There are 28 geographic strata corresponding to each broad stratum: a single multi-state stratum (labeled >1 state) and 27 single geographic area strata. The 2 industries stratum is used only in retail (44-45) and health care (62).

Because the AIES is ultimately a multipurpose sample, it is not possible to optimize for a single data item. Instead, the AIES utilizes stratified sequential sampling, also known as Chromy’s method, which is available from PROC SURVEYSELECT in SAS®.

Stratified sequential sampling is a without-replacement PPS sampling method that allows for a within-strata sort. Within each AIES sampling stratum, the USUs (companies) are sorted by national industry tabulation NAICS (4-digit, 5-digit, or 6-digit) and state, providing implicit stratification. Because the USU probabilities of selection are a monotonically increasing

function, the AIES sampling procedure includes a second-level sort on the USU annual payroll, which should help insure a representative sample in terms of unit size within stratum.

An approximation of the AIES sample size is presented in Table 3:

Table 3 AIES Allocation

Sector	Total
Complex Certainties	36,500
11/21/23/55	5,140
22	950
31-33	33,100
42	17,800
44-45	38,700
48-49	13,000
51	5,200
52	15,000
53	20,500
54	50,300
56	21,000
61	5,650
62	40,300
71	8,200
72	34,200
81	39,400
Total	384,940

In addition to the AIES sample, we plan to add approximately 1,000 large, multiunit companies (primarily covering out-of-scope enterprises) to collection activities to update and maintain a centralized, multipurpose Business Register (BR).

Response Rates

This survey is designed to replace a suite of seven independently designed surveys. The survey unit response rate is expected to fall in the range of those surveys of approximately 50 – 70%. The reporting unit response rate (URR) is the ratio of the number of active, in-scope companies that provided sufficient data to be considered a respondent to the total number of active, in-scope businesses in the sample, times 100 percent. The URR assesses the risk of nonresponse bias. Similar metrics about nonresponse bias are the total quantity response rate (TQRR) and/or the imputation rate (IR). These metrics measure the proportion of the total estimate coming from reported or imputed, sample-weight-adjusted data values. A separate metric can be computed for each item collected in AIES, whereas there is only one unit response rate for the survey. TQRR/IR can be used to gain insight into the quality of item-level estimates versus the overall URR.

Non-Response and Estimation

Data are imputed using survey data and administrative data as input for unit non-response, item non-response, and for responses that fail computer or analyst edits.

Totals are computed as the sum of weighted data (reported and imputed) for all selected sampling units that meet the sample canvass and tabulation criteria given above. The sampling weight for a given sampling unit is the reciprocal of its probability of inclusion into the sample. Measures of variation will be provided in the form of coefficients of variation and standard errors.

Benchmarking

Published estimates from the AIES will be benchmarked to the Economic Census.

2. Procedures for Collecting Information

In the past, the Census Bureau has conducted numerous annual surveys that were often differentiated by type of business (i.e., retail, service, wholesale). In Survey Year 2023 (SY23/Year 1), we will be integrating seven existing annual business surveys into one streamlined single survey called the Annual Integrated Economic Survey (AIES). In Survey Year 2022 (SY22/Year 0), we will conduct a Dress Rehearsal to test and refine our methods prior to the full implementation in SY23. The AIES survey request will replace requests that respondents would have been asked to participate in for the following surveys:

- Annual Retail Trade Survey (ARTS)
- Annual Wholesale Trade Survey (AWTS)
- Service Annual Survey (SAS)
- Annual Survey of Manufactures (ASM)
- Annual Capital Expenditures Survey (ACES)
- Manufacturers' Unfilled Orders Survey (M3UFO)
- Report of Organization

Survey Year 2021 (SY21) was the final collection for the Report of Organization and ASM. SY22 (with data collection during calendar year 2023) will be the final collection for the other five surveys (ARTS, AWTS, SAS, ACES, and M3UFO). During calendar year 2023, we will also be conducting the 2022 Economic Census and the AIES Dress Rehearsal. Coordinating SY22 collection activities with messaging to increase awareness of the AIES will be vital to successfully communicating this significant change to how we collect our data without negatively affecting the collections for the legacy surveys or the 2022 Economic Census.

The table below shows the number of firms for the AIES Dress Rehearsal in SY22 (Year 0) and the full implementation in SY23 (Year 1).

SY22/ Year 0	SY23/ Year 1
8,470	384,940

The initial letter to firms identifies the survey name, due date, instructions for accessing the electronic survey instrument, authority for collection, and burden estimate. It also provides a website with information for respondents and a telephone number for those needing assistance. When we have an email address for the respondent, we will also send out the initial survey request by email. We will periodically follow-up with nonrespondents via mail, email, and phone.

Electronic reporting will be the only advertised reporting option. Nevertheless, the Census Bureau will accept data submitted through other methods. For example, if a firm does not have access to the internet, we can arrange for the business to provide data to an analyst via phone.

3. Methods to Maximize Response

Our intention is for the redesigned survey to be easier to complete (reducing respondent burden), resulting in better and more timely data, and allowing the Census Bureau to operate more efficiently to reduce costs.

To increase awareness of the AIES both during and shortly after the legacy surveys conclude collection for SY22, we will be contacting respondents and non-respondents for each of the legacy surveys to inform them that the legacy survey will be discontinued and communicate the upcoming AIES via email (respondents only) and paper (both respondents and nonrespondents). For the AIES Dress Rehearsal, an advance notice will be sent to the contacts from the legacy surveys to introduce changes in our data collection and inform firms that they will receive the new AIES survey instead of the SY22 legacy annual surveys.

After the AIES initial letter, firms are given at least 30 business days to respond and are given extension dates upon request. Approximately a week before the AIES due date, respondents will receive a due date reminder. Shortly after the due date, we will periodically follow-up with delinquent cases via mail, email, or phone.

We provide a website with information for respondents and a telephone number and secure messaging for those needing assistance. For selected large and complex companies, account managers provide tailored customer service and work with companies to ensure timely and accurate reporting.

4. Tests of Procedures or Methods

Record Keeping Study

A first step in respondent-centered research in support of the development of the AIES was a series of 29 interviews with medium-sized companies across various industries exploring their record keeping practices. These interviews explored the link between financial records and company organizational and management practices. First, researchers asked respondents to describe in their own words what their business does or makes, and then to indicate the North American Industry Classification System (NAICS) most appropriate for their business. Next, researchers explored companies' reporting of four key variables (sales/receipts/revenues; inventory; expenses, focusing on payroll and employment; and capital expenditures); in some cases, respondents were shown specific questionnaires from the annual surveys they had received – and to which they had responded. They were asked to explain their general response process, how they gathered data from multiple sources, and whether they needed to manipulate the data in order to provide answers to what they thought the questions were asking. Researchers probed apparent discrepancies between respondents' reporting practices and the question's intent.

Major findings from the Record Keeping Study shed light on how companies' practices translate into response processes for Census Bureau surveys. This includes three major findings: first, industry classification is challenging and unnatural for respondents, and companies can struggle to fit within the NAICS classification system. Second, businesses varied in their operating units, and this impacted the level and detail at which they keep their records, which can negatively or positively impact response. Finally, consolidated financial records are a mainstay for businesses, and act as an 'anchor' for most detailed information that respondents provide.

Data Accessibility Study

Following the Record Keeping Study, Census Bureau researchers engaged in an additional 30 in-depth interviews with medium sized businesses across various industries about the accessibility of their data at various levels of the business. These interviews first established definitions and equivalencies of response units (for example, what is the company term that means the same as 'establishment'?). Then, respondents were asked about 'general' and 'specific' industry classifications (akin to the four- and six-digit NAICS classifications). Finally, respondents were asked to use a four-point color coded scale ranging from green (very accessible) to red (not at all accessible) to categorize data for specific topics (e.g., revenue, expenses, and payroll) at specific levels (e.g., company-wide, by establishment, or at the general and specific industry levels).

This set of interviewing further detailed what data are available at what levels. Respondents still struggled with their NAICS classification, noting that the categories could be too broad or, conversely, not encompassing enough to accurately describe the company as a whole. At the same time, researchers noted that company-level data are the most accessible regardless of size and sector of the business, and that more granular levels of data can be increasingly challenging

to report, with respondents struggling most to report information by industry. Note that this interviewing also asked about the availability of data at the state level, and that many respondents indicated that to provide response by state, they would first have to output data by establishment and sum up to the state level.

Coordinated Collection Respondent and Non-Respondent Debriefings

In preparation for the integrated survey, staff at the Census Bureau began a program of consolidating company contacts across the disparate current annual surveys to identify one primary contact for the AIES. To better understand how contact consolidation might impact response processes, researchers conducted two rounds of debriefing interviews. The first was 35 interviews with responding companies across various sizes and industries, during which Census Bureau researchers asked questions about communication materials and challenges. The second round was 19 interviews with non-responding companies of various sizes and industries to understand where response may have broken down.

From the respondent debriefing interviews, several major findings informed the contact strategy for the AIES. Researchers noted that many respondents did not receive – or do not remember receiving – many of the communications that the Census Bureau sent in support of consolidating contacts. Some of this may be related to the COVID-19 global pandemic, both because respondents may have been working remotely and not received the contact, or because of the general inconsistencies in mail delivery during the pandemic. Researchers also noted that respondents generally received Census Bureau communications positively (especially the online Respondent Portal that they use to complete surveys), but that respondents' experiences with reaching out for support from the Census Bureau can be mixed. Non-respondent debriefings identified barriers to response, including company-related issues like staffing shortages and data being uncentralized, surveys that do not reflect current company structure, and communication shortfalls from the Census Bureau.

Survey Structure Interviewing

To start to build out the survey flow for the AIES, Census Bureau researchers conducted 39 interviews with companies across various industries focused on preliminary 'mock ups' of survey screens designed to give respondents an overview of each module and the survey at large. Respondents were asked about their record keeping practices and preferences for reporting (e.g., using a spreadsheet or the traditional page-by-page response). Finally, respondents were asked to reflect on the concept of the AIES as a whole, their thoughts about the overall design, and their thoughts on its effect on their burden.

For this interviewing, Census Bureau researchers noted that the AIES introduces complexities for response both at the units of collection (company, establishment, and industry) and by the various topics mixed into one survey. Again, the issue of NAICS classification 'fitting' a business generally – and more specifically the classification of specific industries – was noted as

a space where respondents sometimes struggled. Most respondents indicated wanting to use the downloadable spreadsheet option, to gather their data first and then enter it into the survey, and to submit data across all survey sections at one time, instead of section by section. Respondents also stressed their reliance on others within the company to complete the survey, and that they were open to using survey previews (as PDFs or otherwise) to make a plan for data collection and/or to email to other departments or locations for response.

AIES Pilot Phase I

With this qualitative research in mind, Census Bureau researchers determined that the next research step was to induce independent response from the field for the integrated survey. 78 companies from across the country in various industries were invited to participate in Phase I of the AIES Pilot, and from that, 62 companies (79 percent) provided at least some response data to the survey. In addition to the online survey, 10 companies participated in debriefing interviews asking about response processes for the integrated instrument, 15 companies completed a Response Analysis Survey (a short survey with questions centered on real and perceived response burden), and 11 companies called while 52 companies wrote emails requesting additional information, asking questions, or otherwise reaching out for support.

Major findings for the first phase of the pilot were focused on response process and resultant data. For response process, researchers noted that respondents used a hybrid response process whereby the most accessible data (company-wide) were reported using a traditional page-by-page online survey, but establishment and industry data were reported using a respond-by-spreadsheet survey. Many respondents mentioned the benefits to responding by spreadsheet, but that the spreadsheet must be designed to capture data holistically across the company. Finally, investigations into response burden for the collection instrument were mixed, with some respondents saying it took about the same amount of time or resources or less than the current annual surveys and others saying it took them more time or resources to complete the pilot survey.

AIES Pilot Phase II

With the richness of the data from Phase I of the AIES Pilot, Census Bureau researchers determined that the next step would be to reproduce the pilot, with more and larger businesses, and updated with findings from all of the preceding research. This research seeks to reengage the 62 businesses that provided any response data to the phase I pilot as well as an additional 827 companies from across the country, varying in size and industry. As in Phase I, the Phase II pilot features an updated respond-by-spreadsheet design that reflects the need for holistic collection across the company, as well as debriefing interviews, a repeat of the Response Analysis Survey, and cataloguing contact from the field as an additional source of feedback on the harmonized survey.

The AIES Pilot Phase II research field period began data collection in February 2023 and ran through April 17, 2023. Note that close out was late September, meaning that data will still be accepted through that date, but will not be considered for research findings.

Findings from this round of research are categorized into three topics. First, the Phase II pilot tested specific features of the response spreadsheet, finding that respondents reacted positively to having all locations displayed in one list, the inclusion of auto summing between units reduced item missingness, and giving a choice between location and industry reporting needs additional refinement. Second, the Phase II pilot provided additional information about respondent burden, finding that the scope and length of the survey continues to be challenging, and that burden will be negatively impacted by respondent learning curve as they adjust to the new survey. Finally, the Phase II pilot further explored responded messaging, finding that respondents are using the available response support materials and that contact strategies need additional refining.

Usability Testing

Once the online survey instrument has been programmed and internally tested, Census Bureau researchers will engage in up to three rounds of usability testing with companies varying in size, industry classification, and complexity. This research is being conducted under the Census Bureau's Generic Clearance for Census Bureau Questionnaire Pretesting Research (OMB Control Number 0607-0725) and is planned to include interviewing with up to 30 companies. This testing will focus on three aspects of survey response, including: response support materials like help screens and survey previews; interaction with the online respond-by-spreadsheet design; and the survey flow, including content and logic checks, submission, and other aspects of response. This research is scheduled to begin in **fall** 2023 and conclude with enough time to consider findings for the 2024 production survey design, as well as inform research considerations for post-launch instrument refinement.

Additional detail regarding the Census Bureau's respondent-centered formative research is provided in Attachment H and unit harmonization studies for the AIES in Attachment L

5. Contacts for Statistical Aspects and Data Collection

For questions regarding the planning and implementation of this survey, please contact Blynda Metcalf, Product Owner of the Annual Integrated Economic Survey (Associate Directorate for Economic Programs, U.S. Census Bureau) at 301-763-4781 or Blynda.K.Metcalf@census.gov

For inquiries about the methodology for this survey, please contact James Hunt, Methodology Director (Economic Statistical Methods Division, U.S. Census Bureau) at 301-763-6599 or James.W.Hunt@census.gov

LIST OF ATTACHMENTS

- A. Legal Citation – Title 13 U.S.C
- B. Electronic Instrument Selected Screen Shots
- C. Dress Rehearsal Debriefing Protocol
- D. Letters for the Dress Rehearsal
- E. Summary of the AIES Content
- F. 2023 AIES Draft Content
- G. Consultations with Key Federal Stakeholders
- H. Respondent-Centered Formative Research Overview
- I. Dress Rehearsal Sample Size and Burden Estimates
- J. AIES Sample Size and Burden Estimates
- K. 30-Day Notice Comments Received
- L. Unit Harmonization Studies for the Annual Integrated Economic Survey
- M. Letters for the Production AIES