Survey Design

Target population:

The scope of ACES is to capture capital investment by all domestic, private, non-farm businesses, including agricultural non-farm business and businesses with and without employees. Investment made after applying for an Employer Identification Number (EIN) from the Internal Revenue Service (IRS) but before having any payroll or receipts is also included.

Major exclusions are foreign operations of U.S. businesses, businesses in the U.S. territories, government operations (including the U.S. Postal Service), agricultural production companies and private households.

Sampling unit:

The sampling unit is the company, which is compiled from the Census Bureau's Business Register. The Business Register is at the establishment level, and all establishments linked to a given company are compiled into a sampling unit. For purposes of sampling and adjustment, a company is given a NAICS based business code derived from the NAICS coding of its establishments. The survey samples, collects, and tabulates information at the company level. The records of how large companies invest are often maintained at the headquarters level, and not at the location of each physical operating location.

Sampling design and sampling frame:

ACES is best described as a stratified random sample, which collects information at the company level. There are two distinct samples. The employer sample has basic strata composed of about 131 different industry groupings, then divided into several substrata by a measure of size. The second sample is also a stratified simple random sample, but done only by one of four convenience groupings based on nature of administrative data.

The value of having ACES collect at the company level is that all the company's various business activities, even if unknown to the Census Bureau previously, can be captured. The records of how the company invests are maintained at the headquarters level, and not at the location of each physical operating location. Companies may elect to have divisions within the company report, but the sampling unit and tabulation unit will be the company. A company's importance to the survey depends on their employment, payroll, and business activity. The greater the number of employees or the larger the payroll, the more likely, in general, a company is to be selected in the sample. The influence that the amount of payroll has on the likelihood of selection is adjusted by the business activity such that two companies with similar payroll but in different business activities will not have the same likelihood of selection.

The estimates are from two distinct samples drawn using different sampling designs from two distinct frames. The first frame collects in-scope companies with employees. Companies sampled from this frame will receive an ACE-1 instrument. The frame and sample are referred to as the ACE-1 frame and sample, respectively. The second frame collects companies without employees, also called the nonemployer companies. Companies sampled from this nonemployer company frame will receive an

abbreviated ACE-2 instrument. The frame and sample are referred to as the ACE-2 frame and sample, respectively.

Administrative records data in the Census Bureau's establishment-based database, the Business Register or BR, is used to create the 2022 ACE-1 frame. Administrative data about the previous year, in this case 2021, concerning companies with employees is not of sufficient quality to use for sampling until about 9 months after the end of the target year. The final 2021 administrative data was available in October 2022 and was used to create the 2022 ACE-1 sampling frame. The administrative records are for each physical business entity, called 'the establishment', located within the United States. This information includes company ownership and contact information, and data such as employment, payroll and a single NAICS code associated with each establishment that describes its major business activity. Establishments that are out of scope due to location, ownership, or major business activity based on their assigned NAICS codes, are removed from the sampling frame.

In the ACE-1 frame, company level records are created by aggregating establishments with the same ownership. Most companies, about 6.1 million in 2021, are only a single establishment and no aggregation is needed. These are called single-unit companies (SU). A company that has more than a single establishment, which number is more than 174,000 in 2021, are called multi-unit companies (MU).

For single unit companies, the business activity classification is the NAICS classification already assigned to its sole establishment in the BR. For multi-unit companies, business activity classification is assigned based on an examination of its constituent establishments. The payroll data for each establishment is collected by its assigned NAICS industry. The multi-unit company is assigned a code based on an algorithm that first assigns the company to the trade area in which it had an active establishment with the most payrolls (e.g., manufacturing, construction, etc.). Within that trade area, the three-digit NAICS with the highest payroll is selected. The subsector within that sector with the highest payroll is selected, with the process continuing until the multi-unit company has a six-digit NAICS code. Each company, multi-unit and single unit will have a 2017 NAICS code. This NAICS code is then recoded to an Annual Capital Expenditures Survey specific industry code (ACE codes). These ACE codes are composed primarily of three digits and select four-digit NAICS codes.

All the companies in the 2022 ACE-1 sampling frame, both single unit and multiunit companies, are then partitioned into two major portions: the certainties and noncertainties. Certainties are those companies that will be selected in the sample based solely on their number of employees being at least 500. The data collected from certainties will only represent themselves, aside from further adjustments for nonresponse, while noncertainties have their data magnified by a sampling weight to represent other similar companies that were selected.

The certainty portion of the ACE-1 sampling frame includes 3834 single unit companies and 15,475 multi-unit companies, for a total of 19,309 certainty companies. Many strata will have far more certainties selected than others. The number of certainties selected does not necessarily reflect how many businesses are in that that ACE code, nor the relative amount of investment that companies in that ACE code account for.

The remaining ACE-1 frame companies are partitioned by their payroll into four substrata, then assigned a probability of selection based on a measure of size. This portion is called the noncertainty portion of the frame and is about 6.1 million cases. Those that are selected are the noncertainty companies in the

sample. The size of the noncertainty portion is managed to keep the total ACE-1 sample size around 50,000. Allocation across the noncertainties by ACE code and substratum uses the Lavallee Hidiroglou method.

One important deficiency in the sample design ACE-1 frame is the weak relationship of the administrative data such as the number of employees or payroll to the amount of capital expenditures. While many multiunit companies will be in successive cycles, even their value of capital expenditures in the previous cycle is not a reliable predictor of investment in the current cycle. Knowledge of business activity and size is still helpful in reducing the error due to sampling, and still allows for a reasonable sampling design.

During the creation of the ACE-1 frame, significant parts of the ACE-2 frame are also identified as being associated with an employer identification number, an EIN, but not associated with any employees at that time. These include the 3C, which were known to have had employees in the past, and 3F stratum, which have only recently applied for an EIN. The companies in these two ACE-2 strata are included because, while they are not considered active businesses in many economic surveys, they have been shown to make some capital investments. However, response rates for these companies are often lower than other parts of the survey since many of these are found to be inactive, or to not have even begun operations yet. A filter is used to remove companies that had a higher probability of being either out of business or having failed to start up, while retaining other companies in the hope of retaining the bulk of the investment of these groups.

The other two strata of the ACE-2 come from a distinct nonemployer data set of sole proprietorships, the 3E cases, and the partnerships and corporations, the 3D stratum. Collectively, these four strata usually account for another 20,000 companies getting the abbreviated ACE-2 form. There was an ACE-2 component in 2022, however, in ACES 2021, no nonemployers were mailed. The tables will reflect the absence of the ACE-2 data in 2021.

Frequency of sample redesign:

The ACES sample is reselected annually. The two frames are recreated each year from the best available administrative data several months before mailing out the collection instrument. The analysis on where to shift sample to and from and redone each cycle based on the best available survey and administration data. After these processes, the samples are redrawn each year for both the ACE-1 and the ACE-2 portions. No units are deliberately held over consecutive cycles, but several thousand of the larger units may meet the selection criteria in successive cycles. This is particularly (but not exclusively) true in the ACE-1 certainty portion, where larger companies may be in multiple ACES cycles. No units are deliberately excluded from the sample due to having been selected in a previous cycle.

The Annual Integrated Economic Survey, AIES, will replace ACES beginning with reference period 2023.

Sample maintenance:

ACES resamples each year based on rebuilding the sampling frames and sampling analysis. There is no need for sample maintenance performed during the year to adjust the sample or sampling units. Some maintenance is done for those companies that were sampled but request to report as several separate units. In addition, a sampling unit is not considered a respondent unless all its reporting units have sufficiently responded.