

SUPPORTING STATEMENT FOR BLS CURRENT EMPLOYMENT STATISTICS PROGRAM

OMB CONTROL NO. 1220-0011

B. COLLECTION OF DATA EMPLOYING STATISTICAL METHODS

For the CES program's detailed descriptions of statistical methods and techniques, see the CES Quick Guide (<https://www.bls.gov/bls/empsitquickguide.htm>). The guide also provides links to the BLS Handbook of Methods and Technical Notes.

1. Describe (including a numerical estimate) the potential respondent universe and any sampling or other respondent selection methods to be used. Data on the number of entities (e.g., establishments, State and local government units, households, or persons) in the universe covered by the collection and in the corresponding sample are to be provided in tabular form for the universe as a whole and for each of the strata in the proposed sample. Indicate expected response rates for the collection as a whole. If the collection had been conducted previously, include the actual response rate achieved during the last collection.

The Current Employment Statistics (CES) sample for the private sector is drawn from a private sector universe comprised of about 10.7 million in-scope U.S. private business establishments (worksites). Private households and all agriculture except logging are excluded from the frame. The universe frame, known as the Longitudinal Database (LDB), serves as the sampling frame and research database of the Bureau of Labor Statistics (BLS). The primary data source for the LDB is the Quarterly Contribution Reports filed by employers with their State's Unemployment Insurance Agency. The BLS has cooperative agreements with each State to access these data through the Quarterly Census of Employment and Wages (QCEW) program (OMB Control No. 1220-0012). The LDB contains microdata records with the name, location, North American Industry Classification System (NAICS) industrial classification, employment, and wages of nearly all nonfarm establishments in the United States. Each quarter, the LDB is updated with the most recent universe data available.

The CES probability sample includes approximately 210,000 Unemployment Insurance (UI) accounts selected from the approximately 9.2 million UI accounts in the private sector. The sampled UI accounts cover approximately 1.3 million individual reporting units or worksites. These UI accounts are selected on a random probability basis as described in section 2a below.

In addition, the CES program collects data from Federal, State, and local governments. The governments' sample is not selected on a probability basis, because data are collected for a large percentage of the population employment. Data are collected for: approximately 95.8 percent (2.83 million employees) of total Federal civilian employment; approximately 77.2 percent (3.67 million employees) of total State employment; and approximately 67.9 percent (9.76 million employees) of total local government employment. The government sample units are selected from the 20,811 State, and 77,189 local government UI accounts on the LDB. Federal government employment is largely collected from the National Finance Center.

Data from sample members are collected each month on employment, hours, and earnings. The survey is a Federal-State cooperative survey, with the national sample being a composite of the State samples.

	Preliminary Collection Rate[1] (1st Closing)	Preliminary Response Rate[2] (1st Closing, private only)	Final Collection Rate (3rd Closing)	Final Response Rate (3rd Closing, private only)
01-2020	76%	52%	92%	60%
02-2020	77%	52%	92%	58%
03-2020	66%	37%	93%	54%
04-2020	75%	42%	91%	56%
05-2020	69%	42%	91%	55%
06-2020	63%	43%	90%	55%
07-2020	78%	49%	92%	55%
08-2020	77%	46%	91%	54%
09-2020	70%	43%	90%	54%
10-2020	79%	46%	90%	53%
11-2020	74%	42%	90%	52%
12-2020	76%	45%	95%	52%
01-2021	73%	43%	94%	52%
02-2021	72%	41%	93%	50%
03-2021	67%	41%	93%	51%
04-2021	72%	43%	93%	50%
05-2021	66%	39%	93%	49%
06-2021	64%	37%	93%	48%
07-2021	72%	39%	93%	48%
08-2021	71%	39%	93%	41%
09-2021	74%	42%	93%	48%
10-2021	71%	39%	91%	47%
11-2021	65%	34%	93%	44%
12-2021	71%	38%	93%	46%
01-2022	65%	36%	91%	46%
02-2022	64%	35%	92%	46%
03-2022	62%	35%	95%	46%
04-2022	70%	39%	95%	46%
05-2022	63%	35%	94%	45%
06-2022	73%	38%	95%	45%
07-2022	68%	38%	96%	46%
08-2022	67%	36%	95%	45%
09-2022	73%	39%	95%	45%
10-2022	67%	37%	92%	45%
11-2022	49%	32%	94%	44%
12-2022	64%	35%	94%	44%
01-2023	64%	34%	95%	43%
02-2023	75%	37%	95%	43%
03-2023	74%	37%	95%	43%
04-2023	71%	35%	95%	42%
05-2023	55%	32%	93%	42%
06-2023	66%	34%	95%	42%
07-2023	68%	34%	95%	42%

Collection rate is the number of reporting locations that responded, divided by the total number of locations selected for the sample, and minus those that are out of business, prior refusals, or tacit refusals of more than 6 months.

Response rate is the number of Unemployment Insurance (UI) numbers that responded, divided by the total number of UIs selected for the sample (excluding UIs determined to be out of business).

2. Describe the procedures for the collection of information including:

- **Statistical methodology for stratification and sample selection,**
- **Estimation procedure,**
- **Degree of accuracy needed for the purpose described in the justification,**
- **Unusual problems requiring specialized sampling procedures, and**
- **Any use of periodic (less frequent than annual) data collection cycles to reduce burden.**

2a. Statistical methodology for stratification and sample design

The CES probability sample is a stratified, simple random sample, where the strata are defined by State, major (NAICS) industry division, and employment size (BLS Handbook of Methods, Chapter 2, page 5, Bureau of Labor Statistics, 2011). Stratification groups population members together for the purpose of sample allocation and selection. With 13 industries and 8 size classes, there are 104 total strata per State (Erkens, Huff, and Gershunskaya, 2005). The sampling rates for each stratum are determined through a method known as optimum allocation, which distributes a fixed number of sample units across a set of strata to minimize the overall variance, or sampling error, on the employment estimates at a State level.

The total probability sample size is fixed for each State, and we allocate that sample to minimize the standard error of each State's Total Private employment estimate. The sample size was fixed during the research for the conversion from the quota sample to the probability sample in the 1990s. During the initial research for that conversion to a probability-based sample, CES fixed the sample sizes for all states to the same size as the quota sample for all states. CES originally allocated sample to a State's strata with a consideration for strata cost, but CES eliminated that cost aspect around 2007 or 2008.

Every 5 years the sample is reallocated across states to allow for a better distribution of sample to states. The process is roughly equivalent to a reallocation of sample that's proportional to the Total Private employment in a state, but constraints are placed on the reallocation so that no states lose too much sample.

The UI account is the basic sampling unit of the CES survey. UI account numbers are unique within a State and include all individual establishments within a firm. Though the LDB is updated on a quarterly basis, the sample is redrawn annually due to budget and operational constraints. The annual sample is redrawn to reflect the changes in the population or the sampling frame. These changes include but are not limited to: removal of establishments that have gone out of business; inclusion of new establishments (i.e., births); growth or decline in the employment of each establishment and consequently the UI account; and updated industrial classification for some establishments. The annual sample of UI accounts is drawn using a permanent random numbers technique in order to achieve the desired sample overlap with the UIs from previous samples (Butani, Robertson, and Mueller, 1998). The use of this technique

keeps solicitation costs within budget and keeps the sample statistically valid in terms of updated probabilities of selection that reflect all recent changes (Crankshaw, Kujawa, and Stamas, 2002). In addition to the annual redraw, the CES sample is updated on a semi-annual basis, as more recent universe data becomes available. The semi-annual update provides the opportunity to sample birth units that were not previously on the sampling frame during the annual redraw.

Large UI accounts are sampled with virtual certainty. In addition, all units reporting through the Electronic Data Interchange (EDI) center are sampled with certainty. EDI units are collected via direct file transmission from large, multi-unit employers for whom conventional data collection methods would be less cost effective.

The size of the currently selected probability sample and sample coverage is shown in the following table.

**CES Survey Universe and Sample Size Comparison on NAICS basis, Private Sector
(in thousands)**

	Universe (Q1 2023)			Sample (Q1 2023)		
Industry	UI Accounts	Reporting Units	Employment	UI Accounts	Reporting Units	Employment
Mining and Logging	33.5	38.5	625.3	2.1	5.6	291.1
Construction	874.5	905.8	7,657.2	18.7	44.9	1,371.9
Manufacturing	350.6	385.4	12,847.2	14.8	38.0	5,182.7
Services	7,749.3	9,235.6	103,567.1	173.6	1,163.7	45,311.8
Education	146.7	156.1	3,172.3	5.0	11.3	1,472.4
Total	9,007.9	10,565.2	124,696.8	209.2	1,252.2	52,157.5

2b. Estimation Procedure

The estimation technique used in estimating All Employees (AE) is a weighted link-relative estimator, which is a form of ratio estimation (for detailed mathematical formulae of this estimator, see the BLS Handbook of Methods, Chapter 2, pages 5-7, Bureau of Labor Statistics, 2011). From a sample composed of establishments reporting for both the previous and current months, the ratio of current month weighted employment to that of the previous month, weighted employment is computed. The weights are defined to be the inverse of the probability of selection in the sample. The weight is calculated based on the number of UI accounts actually selected within each allocation cell. The CES sample alone is not sufficient for estimating the total change in employment because of the births and deaths of firms (firm births are restricted because these firms are not present in the CES sample frame due to timing lags, and firm deaths may not be reliably reported by respondents). CES accounts for the net contributions to employment change from births and deaths in its AE estimates by adding a net birth-death residual forecasted from historical population residuals to the current month's weighted employment (NOTE: a recent methodology change due to the pandemic's effects on net birth-death residuals includes the use reported deaths and births in estimation by a factor that represents by how much each exceeded their normal proportion historically). Estimates are

calculated within each estimation cell and then summed across appropriate cells to form estimates for aggregate levels.

The weighted difference link and taper estimator used for non-AE datatypes accounts for the over-the-month change of the sampled units, but also includes a tapering feature used to move the estimate towards the sample over time. Sample averages for non-AE datatypes are known to vary widely from month to month and the taper portion of the formula helps alleviate this issue as it uses a portion of the prior month's estimate and a portion of the prior month's sample average (for detailed information on this estimator, including a discussion on the weights used for the link and taper components, see the BLS Handbook of Methods, Chapter 2, pages 10-11, Bureau of Labor Statistics, 2011). Like the estimator for AE, only the matched sample is used to reduce the variance on the over-the-month change.

2c. Degree of accuracy needed for the purpose described in the justification

Like other sample surveys, the CES survey is subject to two types of error, sampling and non-sampling error. The magnitude of sampling error, or variance, is directly related to the size of the sample and the percentage of universe coverage achieved by the sample. Because the CES sample covers over one-third of total universe employment, the sample error on the national, first closing, total nonfarm employment estimate is small. The relative standard error (2023) for the major industry divisions at the national level under the probability sample are given in the table below.

Industry (form type)	Major Industry Division	Average Relative Standard Error for All Employment (in percent)
Mining and Logging	Mining and Logging	3.4
Construction	Construction	0.7
Manufacturing	Manufacturing	0.4
Services	Wholesale Trade	0.6
	Retail Trade	0.4
	Transportation and Warehousing	0.7
	Utilities	1.8
	Information	2.3
	Financial Activities	0.5
	Professional and Business Services	0.5
	Leisure and Hospitality	0.8
	Health Care	0.4
	Other Services	0.8
Education	Education	1.1

For more information on relative standard error and a comparable table can be found in section 1C of the Technical Notes for the Current Employment Statistics Survey at <https://www.bls.gov/web/empsit/cestn.htm#section1c>.

The estimation of sample variances for the CES survey is accomplished through the method of Balanced Half Samples (BHS). This replication technique uses half samples of the original sample and calculates estimates using those subsamples. The sample variance is calculated by measuring the variability of the estimates made from these subsamples. For a detailed mathematical presentation of this method, see the BLS CES Handbook of Methods at www.bls.gov/opub/hom/ces/calculation.htm#reliability.

2d. Unusual problems requiring specialized sampling procedures

Benchmark Revisions

The sum of sampling and nonsampling error can be considered the total survey error. Most sample surveys are only able to publish the sampling error as their only measure of error. The CES program has the ability to produce an approximation of the total error, on a lagged basis, because of the availability of the independently derived universe data from the QCEW program. While the benchmark error is used as a measure of the total error for the CES survey estimate, it actually represents the difference between two independent estimates derived from separate processes, and, thus, reflects the errors present in each program. Historically, the benchmark revision has been very small for total nonfarm employment. Over the past decade, the absolute benchmark error has averaged 0.2 percent, with an absolute range from <0.05 percent to 0.3 percent. For more information of benchmark revisions, see <https://www.bls.gov/web/empst/cesbmart.htm>.

Specialized Procedures

The BLS has conducted extensive research into various ways to more directly capture the impact of new business births. This research included obtaining early records of new UI accounts and a pilot program to solicit from this frame. Operationally, a sample-based approach did not yield satisfactory results. This was mainly due to the lack of a comprehensive sampling frame on a timely basis. While both employment gains and losses from new and failed businesses are large in terms of over the year change, research conducted by the BLS shows the net employment (employment gained minus employment lost) is small because the gains and losses offset each other (Mueller, 2006). The sample design accounts for the majority of the employment gain from new businesses by imputing for UI accounts that have gone out-of-business (Kropf, Strifas, and Traetow, 2002). On a semi-annual basis, the universe is reviewed to identify new births. A portion of the births are selected on a probability basis. Thus, only births (and deaths) since the semi-annual update (about a 15-month lag) must be imputed for. The BLS has researched models to account for the residual birth employment not accounted for by the death imputation model. Models are currently in use for all private-industry estimates. For more information on the birth/death model, see <https://www.bls.gov/web/empst/cesbd.htm>.

2e. Any use of periodic (less frequent than annual) data collection cycles to reduce burden.

The CES survey was mandated by Congress to be a monthly survey.

3. Describe methods to maximize response rates and to deal with issues of non-response. The accuracy and reliability of information collected must be shown to be adequate for intended uses. For collections based on sampling, a special justification must be provided for any collection that will not yield "reliable" data that can be generalized to the universe studied.

New firms are enrolled into the survey by interviewers working in BLS Data Collection Centers. The response rate for new enrollments was approximately 65% prior to the COVID-19 pandemic. Since the COVID-19 pandemic, response rates for new enrollments have ranged from 30 to 68 percent by the third month after enrollment. After enrollment, sample attrition averages about 1 percent per month. The CES survey rotates new units into the sample each quarter both to replace deaths and to re-align the sample by industry and size. Typically, about 25 to 30 percent of the units are replaced each year.

The response rate used in making final estimates (i.e., 3rd Closing) for the private sector probability sample averaged 60 percent for the year prior to the COVID-19 pandemic. Since the start of the COVID-19 pandemic, response rates for final estimates have ranged from 41 to 60 percent. Please refer to the response rates/collection rates table provided in the answer to question #1 for additional information. As indicated earlier, these sample respondents from the private sector are combined with government reports that cover 95.8 percent of Federal government employment, 77.2 percent of State government employment, and 67.9 percent of local government employment to make estimates for total nonfarm employment.

The link-relative estimating technique implicitly adjusts for nonrespondents using the respondents' relationship of current to previous month's employment within each estimation cell.

CES survey estimates are generated three times for each month to reflect additional sample received and error corrections. (Estimates are revised two more times to reflect updated universe counts.) Policy makers in both the private and public sectors rely extensively on the first estimate for the month. The BLS has implemented procedures to limit the size of revisions in these preliminary estimates. Automated collection methods, such as CATI and Web, have been identified as the best possible means of overcoming the revision problem in the first estimate. These methods have been found to consistently maintain collection rates for preliminary estimates between 68 and 80 percent.

The BLS and the cooperating Territories conduct an extensive and vigorous program of notification and non-response follow-up. These include:

- Targeted advance notice emails sent to all sample units; and
- Time specific nonresponse prompting emails and telephone calls.

The BLS conducts an average of 15,000 non-response prompt phone calls per month.

In addition, the BLS follows an aggressive refusal and delinquent conversion protocol. Each month the BLS Data Collection Centers target prior refusals and delinquents for re-contact. About one-half of these refusals and delinquents agree to resume reporting.

Growth of EDI, the direct transfer of data from the firm to the BLS, also provides a high level of response and stability. The BLS currently collects approximately 124,000 reports from 87 large firms via EDI. For final estimates, virtually all of these firms provide data. EDI also experiences very few refusals but unfortunately the EDI collection method is not applicable to smaller firms.

Each year the BLS conducts analyses of the survey estimates and decomposes the total survey error into its various components including sampling error, non-response error, reporting error, and errors from other sources. These analyses are possible since the LDB provides employment data for all units in the population. It is possible to use this employment information to calculate CES estimates using all of the units selected in the sample (100 percent response) and compare it with the CES estimates using only those units that responded to the CES survey. This provides a measure of the non-response error. Similar methods are used to measure other sources of survey error which can then be aggregated to the total survey error. See Gershunskaya, Eltinge, and Huff, 2002, for detailed mathematical formulae and numerical results on analysis of multiple error components in small domain estimates for the CES program. These analyses are useful in determining the source of errors for past estimates; however, they have not yet proven useful in predicting or limiting the errors in current estimates. Gershunskaya and Huff (2009) found that non-response error was not a principal factor in the total survey error and that the non-response error is not unidirectional or particularly related to the response rates in a research paper presented at the 2009 American Statistical Association Annual Meeting. The study covers the CES estimates from 2004 through 2008. Detailed tables of response rates for the private sector probability sample are available for earlier years and are updated on an annual basis. Details are available on request.

4. Describe any tests of procedures or methods to be undertaken. Testing is encouraged as an effective means of refining collections of information to minimize burden and improve utility. Tests must be approved if they call for answers to identical questions from 10 or more respondents. A proposed test or set of test may be submitted for approval separately or in combination with the main collection of information.

The CES program plans to conduct another phase of the CES Enrollment Package Test, initially conducted in 2020 with a sample of 100 firms (reduced from the original sample size of 400 due to the COVID pandemic). The goal is to streamline the survey's enrollment process and ultimately boost response rates. For the initial 2020 pilot, firms were contacted by regular mail using a modified, single-page letter and encouraged to report online. The pilot showed promise, with 12% of firms self-enrolling and providing data online. Given the moderate success of the initial test, an additional phase will be conducted using lessons learned from the pilot. In the additional phase we plan to send prospective respondents an enrollment email after a data collection specialist has confirmed their email address. The email will replace the survey's traditional CATI enrollment process and the paper letter from the initial test pilot. The email will ask the prospective respondent to self-enroll, as well as provide data, via the CES' existing web collection tool. The email has been developed in consultation with the BLS' Office of Survey Methods Research and is provided in **Attachment A**.

Test Plan

CES intends to conduct this test with a sample size of up to 1,000 respondents as soon as we have approval. The sample frame for the test will be the latest CES quarterly sampling panel. Units will qualify for the test when a data collection specialist confirms a valid email address with the prospective respondent. Once the email address is confirmed we will send the email to the respondent. The primary research objective of this experiment is to identify if the test procedure positively impacts enrollment rates. The results of the test will also determine if CES will expand availability of self-enrollment to all newly sampled respondents and if it is worthwhile to commit additional development resources to expand self-enrollment functionality on our current data collection website.

Assessment of Results

CES will compare the enrollment rate of the sampling panel used for the test against the prior quarter's sampling panel enrollment rate. An increase of at least 15% is the desired outcome. Additionally, analysis will be conducted comparing the delinquency and nonresponse of the two panels over time.

5. Provide the name and telephone number of individuals consulted on statistical aspects of the design and the name of the agency unit, contractor(s), grantee(s), or other person(s) who will actually collect and/or analyze person(s) who will actually collect and/or analyze the information for the agency.

Mr. Edwin Robison, Chief, Statistical Methods Division of the Office of Employment and Unemployment Statistics, is responsible for the statistical aspects of the CES survey.

6. References

American Statistical Association (1994) "A Research Agenda to Guide and Improve the Current Employment Statistics Survey." American Statistical Association Panel for the Bureau of Labor Statistics' Current Employment Statistics Survey, January, 1994. Alexandria, VA: American Statistical Association.

Bureau of Labor Statistics. BLS Handbook of Methods Chapter 2: Employment, Hours, and Earnings from the Establishment Survey. Washington DC: Bureau of Labor Statistics, 2004, p.5. <https://www.bls.gov/opub/hom/pdf/ces-20110307.pdf>

Bureau of Labor Statistics. BLS Handbook of Methods Chapter 2: Employment, Hours, and Earnings from the Establishment Survey. Washington DC: Bureau of Labor Statistics, 2004, pp. 5-7. <https://www.bls.gov/opub/hom/pdf/ces-20110307.pdf>

Bureau of Labor Statistics. BLS Handbook of Methods Chapter 2: Employment, Hours, and Earnings from the Establishment Survey. Washington DC: Bureau of Labor Statistics, 2004, pp. 8-9. <https://www.bls.gov/opub/hom/pdf/ces-20110307.pdf>

Butani, Shail, Kenneth W. Robertson, and Kirk Mueller (1998) "Assigning Permanent Random Numbers to the Bureau of Labor Statistics Longitudinal (Universe) Data Base." Proceedings of the Survey Research Methods Section, American Statistical Association, 9-13 August, 1998. Dallas: American Statistical Association, 1998.
<https://www.bls.gov/osmr/research-papers/1998/pdf/st980080.pdf>

Crankshaw, Mark, Laurie Kujawa, and George Stamas (2002) "Recent Experiences in Survey Coordination and Sample Rotation within Monthly Business Establishment Surveys." Proceedings of the Survey Research Methods Section, American Statistical Association, 11-15 August, 2002. New York: American Statistical Association, 2002.

<http://www.bls.gov/osmr/research-papers/2002/pdf/st020290.pdf>

Erkens, Gregory, Larry L. Huff, and Julie B. Gershunskaya (2005) "Alternative Sample Allocations for the U.S. Current Employment Statistics Survey." Proceedings of the Survey Research Methods Section, American Statistical Association, 7-11 August, 2005. Minneapolis: American Statistical Association, 2005, pp. 1-4.
<https://www.bls.gov/osmr/research-papers/2005/pdf/st050240.pdf>

Gershunskaya, Julie, John L. Eltinge, and Larry L. Huff (2002) "Use of Auxiliary Information to Evaluate a Synthetic Estimator in the U.S. Current Employment Statistics Program." Proceedings of the Section on Survey Research Methods, American Statistical Association, 11-15 August, 2002. New York, NY: American Statistical Association, 2002.
<http://www.asasrms.org/Proceedings/y2002/Files/JSM2002-000844.pdf>

Gershunskaya, Julie, and Larry Huff (2009) "Components of Error Analysis in the Current Employment Statistics Survey." Proceedings of the Survey Research Methods Section, American Statistical Association, 1-6 August 2009. Washington, DC: American Statistical Association, 2009.
<https://www.bls.gov/osmr/research-papers/2009/st090050.htm>

Goldenberg, Karen L., Susan E. Moore, and Richard J. Rosen (1994) "Commercial Payroll Software and the Quality of Employment Data." Proceedings of the Survey Research Methods Section, American Statistical Association, 13-18 August, 1994. Toronto: American Statistical Association, 1994.
http://www.asasrms.org/Proceedings/papers/1994_178.pdf

Kristin Fairman, Margaret Applebaum, Chris Manning, and Polly Phipps (2009) "Response Analysis Survey: Examining Reasons for Employment Differences Between the QCEW and the CES Survey" Proceedings of the Survey Research Methods Section, American Statistical Association, 1-6 August 2009. Washington, DC: American Statistical Association, 2009.
<https://www.bls.gov/osmr/research-papers/2009/st090240.htm>

Kropf, Jurgen, Sharon Strifas, and Monica Traetow (2002) "Accounting for Business Births and Deaths in CES: Bias vs. Net Birth/Death Modeling." Washington DC: Bureau of Labor Statistics, 2002.

<https://www.bls.gov/osmr/research-papers/2002/pdf/st020090.pdf>

Mueller, Kirk (2006) "Impact of business births and deaths in the payroll survey." Monthly Labor Review, Vol. 129, No. 5, May 2006, pp. 28-34.

<https://www.bls.gov/opub/mlr/2006/05/art4full.pdf>

Werking, George S., Richard L. Clayton, and Richard J. Rosen (1995) "Studying the Causes of Employment Count Differences Reported in Two BLS Programs." Proceedings of the Survey Research Methods Section, American Statistical Association, 13-17 August, 1995. Orlando: American Statistical Association, 1995.

http://www.asasrms.org/Proceedings/papers/1995_137.pdf