

## **Bureau of Labor Statistics**

### **United States Business Response Survey and JOLTS Supplemental Survey**

#### **Part B. Statistical Methods**

BLS will include all of the statistical design information on the website developed for disseminating the results to the public. This information will afford additional transparency with the public about the underlying methods used to produce the information.

#### **1. Respondent/Sample Universe**

##### **1a. Universe**

The Bureau of Labor Statistics (BLS) plans to conduct a new information collection, the Business Response Survey (BRS) questionnaire on two samples of establishments.

- The large sample will be drawn from the establishments included in the BLS Business Register, built from the Quarterly Census of Employment and Wages (QCEW). There are currently 9 million in-scope establishments on the BLS Business Register. The universe of respondents to the QCEW are the 50 States, the District of Columbia, Puerto Rico, and the U.S. Virgin Islands. The primary source of data for these 53 entities are the Quarterly Contribution Reports (QCRs) submitted to State Workforce Agencies (SWAs) by employers subject to State Unemployment Insurance (UI) laws. The QCEW data, which are compiled for each calendar quarter, provide a comprehensive business name and address file with employment and wage information by industry, at the six-digit North American Industry Classification System (NAICS) level, and at the national, State, Metropolitan Statistical Area (MSA), and county levels for employers subject to State UI laws. Similar data for Federal Government employees covered by the Unemployment Compensation for Federal Employees program (UCFE) are also included.
- The JOLTS Supplemental Survey sample will be the sample used for the March 2020 JOLTS collection. This sample was also drawn from the BLS Business Register, and the methods for selection are described in the OMB package for the JOLTS information collection (OMB Control No. 0170).

The scope of the large sample BRS is:

- Private sector establishments on the frame in all 50 states, the District of Columbia, and Puerto Rico
- Excluding Private Households (NAICS 814110); Services for the Elderly and Disabled persons (NAICS 624120 with employment < 2); U.S. Postal Service (NAICS 491110); and unclassified accounts (NAICS 999999).

The BRS sampling frame consists of approximately 9 million in-scope establishments.

##### **1b. Sample size**

For the large sample BRS, BLS plans to select a sample of approximately 597,000 establishments. The objective of the large sample is to produce statistics at detailed levels including by size class, state, industry, and some state-industry, state-size combinations. Production of statistics at any detailed levels will depend on the ability to pass disclosure requirements to ensure confidentiality of the establishments responding to the survey.

In addition to the large sample BRS described above, BLS plans to survey the March 2020 JOLTS sample of establishments, OMB Package (1220-0170). Collecting this information for these JOLTS sample units will be critical in analyzing and understanding the job openings and closings data reported in the JOLTS survey for the same timeframe.

	Sample	Estimated Responses (25%)
Large Sample	596,884	149,388
Supplemental Sample using March 2020 JOLTS sample	13,237	3,310

The JOLTS supplemental sample will be kept separate from the primary sample, with the goal for the supplemental sample being to produce analysis that builds upon and adds value to the data already collected in the JOLTS. The longitudinal aspect of the JOLTS will allow for an analysis of the collected information by looking at establishment staffing patterns prior to and after completing the BRS. This is a unique analysis from the primary sample, and adds a valuable dimension to understanding business responses to the Coronavirus pandemic.

## 2. Data Collection

### 2a. Sample Design and Selection Procedures

The survey analysis breakouts requested by researchers are show below, noting that all 50 states plus Washington DC and Puerto Rico are included:

- State by Special Interest NAICS Sector Categories  
(6 special interest NAICS categories: 23, 31-33, 44-45, 62Alt, 72, Others)
- State by Size  
(2 size categories: big, small)
- NAICS Sector Categories  
(22 NAICS sector categories; 2 size categories (big, small))
- Industry Size Class  
(9 classes: 1-4, 5-9, 10-19, 20-49, 50-99, 100-249, 250-499, 500-999, 1000+)

Sufficiency is determined for each survey analysis breakout listed above. The four sets of sufficiency counts are then meshed together to create a single unified sample design. The unified design specifies a targeted number of responders for each State by NAICS Sector Category by Industry Size Class combination. Sample sizes are then derived by multiplying each combination’s targeted number of responders by an estimate of the survey response rate. The overall sample size is the sum of these individual cell sample sizes.

Sufficiency is determined based on estimating proportions to a certain degree of precision, where precision is based on researcher needs weighed versus survey burden and cost. The formula for the sample sufficiency of an estimation cell is based on the deconstruction of the formula for the variance of a proportion (using simple random sampling within the cell):

$$n_h = \frac{s_h^2 + p_h(1 - p_h)}{s_h^2 + \left( \frac{p_h(1 - p_h)}{N_h} \right)}$$

Where:

h is the stratum or analysis breakout cell

$n_h$  is the sample size sufficient to estimate to the desired precision in stratum h

$N_h$  is the stratum h population

$p_h$  is a guess at the eventual proportion estimate in stratum h (always set to 0.5)

$s_h$  is the standard error chosen by the researcher to set the precision level for stratum h

The following table provides overall sample sizes resulting from applying the design using different combinations of standard error and response rate settings:

<b>P</b>	<b>Standard Error</b>	<b>95% Confidence Interval Error Bound (+/-)</b>	<b>90% Confidence Interval Error Bound (+/-)</b>	<b>Est. Resp. Rate</b>	<b>Total Targets</b>	<b>Total Sample Size</b>
0.5	0.05	0.098	0.0825	25%	42665	167643
<b>0.5</b>	<b>0.025</b>	<b>0.049</b>	<b>0.04125</b>	<b>25%</b>	<b>149388</b>	<b>596884</b>
0.5	0.05	0.098	0.0825	20%	42595	208461

The sample size is based on the criteria described in the table above, specifically the second row, in addition to the number of state-industry-size class cells we wish to estimate. Given government mandates related to the virus will differ from state to state, and industries and size classes had different programs targeted towards them, it was desired to produce estimates for specific industry-state groups and size class

– state groups, and given the burden estimate and the fact that this sample size is much smaller than other online surveys approved asking about business responses to the Coronavirus pandemic.

Once sample sizes have been finalized for each state, NAICS sector, and size class stratum, samples will be selected within each stratum using simple random sampling.

The supplemental JOLTS sample is based directly on the random sample of respondents in March 2020 and thus relies on the same sampling procedure as JOLTS. For the sake of clarity in understanding the sample designs and sample weights calculated for the BRS primary sample, the two samples will be analyzed separately. This also emphasizes that the analyses for the samples are complementary.

## 2b. Estimation Procedure

The primary measure of interest will be an estimated proportion possessing an attribute being assessed by a survey question. Proportions will be estimated for each of the analysis breakout cells described in section 2a. The proportion estimate formula is:

$$\hat{p} = \sum_{h=1}^H W_h \hat{p}_h$$

Where:

$\hat{p}$  is the est. weighted sample prop. (for a particular survey question) for a breakout cell

$\hat{p}_h$  is the est. sample prop. (for a particular survey question) for stratum h within the breakout cell

$W_h$  is the stratum weight within the breakout cell

H is the number of strata in the breakout cell

Each stratum weight is the population proportion of each stratum relative to the composite population of interest:

$$W_h = N_h / N$$

Where:

$N_h$  is the population of stratum h within the analysis breakout cell

$N$  is the population of the entire analysis breakout cell

The formula for the estimated sample proportion for some stratum (h), generalized for non-response adjustment, is:

$$\hat{p}_h = w_h^{resp} \hat{p}_h^{resp} + w_h^{nonresp} \hat{p}_h^{nonresp}$$

Where:

$\hat{p}_h$  is the non-response-adjusted estimated proportion possessing the attribute of interest in stratum h

$\hat{p}_h^{resp}$  is the estimated proportion of responders possessing the attribute of interest in stratum h

$\hat{p}_h^{nonresp}$  is the estimated proportion of non-responders possessing the attribute of interest in stratum h

$w_h^{resp}$  is the proportion of the stratum h sample that responded to the survey question

$w_h^{nonresp}$  is the proportion of the stratum h sample that did not respond to the survey question

The proportion of non-responders possessing the attribute of interest is generally unknowable. Therefore, the assumption is made that, within each stratum, responders and non-responders possess the attribute of interest in the same proportion, and therefore the formula reduces as follows:

$$\hat{p}_h = \hat{p}_h^{resp}$$

At a later date, survey sponsors may decide to request composite estimates for new analysis breakouts that were not determined ahead of time. The definitions above can be generalized to composite estimators over various strata groupings by redefining the universe to the one of interest. However, the sample was not designed with the intention of being sufficient for these new analyses.

Estimation procedure for the supplemental JOLTS sample will be based on sampling weights derived from standard JOLTS procedures.

## 2c. Reliability

Variance estimation will involve (i) the application of the general formula for the variance of a composite proportion estimator drawn from a stratified random sample and (ii) the application of the basic formula for the variance of a proportion drawn from a simple random sample. Specifically, the variance of the proportion estimator for some particular analysis breakout cell is:

$$Var(\hat{p}) = \sum_{h=1}^H W_h^2 Var(\hat{p}_h), \text{ where } W_h = \frac{N_h}{N}.$$

Under the assumption that, within each particular stratum, non-responders possess the attribute of interest in the same proportion as responders, the formula for the within stratum variance of a proportion calculated from a simple random sample is:

$$Var(\hat{p}_h) = \left(1 - \frac{n_h}{N_h}\right) \left(\frac{\hat{p}_h(1 - \hat{p}_h)}{n_h - 1}\right)$$

In the formula above, note that  $n_h$  is the number of establishments in stratum h that *responded* to the survey question of interest. It is *not* the stratum h sample size.

The formulas above can be tailored to the desired composite estimator by applying it across only the set of strata that are relevant to that particular composite.

Variance calculation for the supplemental JOLTS sample will be based on the standard calculations done in monthly JOLTS samples.

## 2d. Data Collection Cycles

Data collection will be coordinated through the Office of Field Operations (OFO) for selected establishments that are also in other BLS surveys to reduce respondent burden as much as possible.

The infrastructure to support the BRS is in place, as BLS will make use of existing technology and data collection infrastructure from the Annual Refiling Survey (ARS) and the Quick Business Survey (QBS). This will allow BLS to begin to field the BRS within a week of receipt of clearance.

BLS will use both email and physically printed and mailed letters for solicitation purposes, following the existing, proven structure of the ARS. The BRS will make use of ARS-collected email addresses and UI Tax email addresses from 40 participating states for use in email solicitation. The QCEW program maintains a list of 5.5 million email addresses for the ARS. To reduce respondent burden, reduce survey costs, and increase response rates, the BRS will utilize these email addresses to solicit businesses whenever an email address is available.

When an email address is not available, the respondent will receive a printed solicitation letter. Where an email address was unusable or unresponsive, the respondent will also receive a printed solicitation letter.

In the BRS, respondents will be asked to confirm their mailing address and physical location address and provide contact information, confirm their business activity (NAICS code) or indicate if it has changed as a result of the pandemic, and answer seven questions through the online application. Non-response follow-up will follow the well-established procedures developed for the ARS, and prior Quick Business Survey (QBS) testing over the last three years.

Emailed solicitations will go out concurrently with mailed solicitations covering 100 percent of the BRS sample. After about two weeks, a follow-up email will be sent to the email-sample non-responders to maximize the email response rate. A second printed solicitation that also includes any email non-respondents and undeliverable emails will be mailed about 5 weeks after the initial mailout. The QBS pilots conducted in 2018 and 2020 have been instrumental in developing “quick” data collection procedures allowing for an expedited analysis. The QBS made use of existing technology, data collection infrastructure and plans and result analysis from the ARS, just as we propose to do here for the BRS. BLS will be able to make use of much of what has been learned piloting and testing the QBS for this BRS collection and review. The first two QBS pilots provide a roadmap for the BRS and offer insights in the expected response rate to the BRS. The first QBS pilot tested several contact methods and resulted in an overall response rate of 25 percent after four months of data collection, with variations between 44

percent and 18 percent based on solicitation methodology. Pilot 1 was fielded mid-way through the ARS collection cycle, concurrent with the second ARS mailout, and so did not have the benefit of the initial ARS response. The second QBS pilot reached a 22 percent response rate after two months of data collection and 38 percent after 8 months. Pilot 2 was conducted using the ARS timeline, concurrent with the first ARS solicitation, but was limited to a single mailout within the first 2 months, and so did not take advantage of the accelerated solicitation schedule planned for the BRS. Clearly, the anticipated response rate for the BRS is in line with the QBS testing results.

Overall data collection is expected to be conducted over an eight week period, giving time for response rate review and analysis to determine the optimal data collection outreach methods to maximize response rates and reduce burden. BLS will rely on review and analysis tools developed for the ARS and QBS to assist in decision making.

### **3. Response Rates**

BLS expects a response rate of 25% for both the large sample and the supplemental JOLTS sample. This expectation is based on responses to prior test estimates of the QBS and results observed from the Census Pulse Survey of establishments..

#### **3a. Maximize Response Rates**

BLS will rely on established, tested data collection processes to ensure maximum response rates to the BRS. BLS will make use of existing processes from the ARS and QBS, along with consulting other established BLS survey programs to maximize efficiency and reduce burden. One way that BLS has found to maximize response rates through the ARS and QBS, is the utilization of email solicitation, which will be used for the BRS. To maximize response rates, email solicitation will be used where available. For establishments without an available email address and for those where the email address has proven unresponsive, a one-page solicitation letter will be used. This will allow for maximum reliance on electronic solicitation while still enabling BLS to reach respondents who are unresponsive to email solicitation. This approach has worked consistently for the ARS and QBS. BLS expects to have email addresses available for approximately 50 percent of the BRS sample. These email addresses cross all industries and size classes and are not concentrated in any discernable way. They have been collected via the ARS program through online collection since 2012. UI Tax email addresses are also provided by 40 participating states and territories for use by the QCEW program. A smaller number of email addresses will be available in twelve states that do not provide a UI Tax email address file. In these states, emailed solicitation will likely be possible for about 35 percent of the sampled units.

BLS will further coordinate with the Office of Field Operations (OFO) to ensure maximum response rates. This coordination will include sample review across other BLS surveys and data collection modes. Where BLS is collecting data for some other survey programs for a particular sample unit, OFO will coordinate with the other BLS survey collection and the sample unit to collect the BRS through the IDCF. By working through established, trusted BLS data collection contacts in other programs, respondents will be able to respond to the BRS. This will reduce respondent burden and maximize response rates, particularly among targeted respondent groups or certainty units.

All of the data collection will take place online via the BLS Internet Data Collection Facility (IDCF). This method of fully online data collection has been successfully employed and tested with the ARS and QBS and has been effective in reducing cost to the government, reducing respondent burden and maximizing response rates<sup>123</sup>.

### **3b. Non-Response Adjustment**

Because BLS expects a response rate of 25 percent, BLS will be performing a nonresponse bias analysis and determining appropriate nonresponse bias adjustments. Non-response bias adjustment methods will be assessed based on survey results. One method that will be explored will be to impute missing questionnaire responses using the response of the nearest responding neighbor. Candidate neighbors will be from within stratum if possible. Nearness will most likely be measured using physical distance rather than some other characteristics such as employment and/or wages. BLS will also explore other nonresponse bias adjustments, as appropriate, and final methods will be documented along with all of the other statistical design methods on the public webpage used for dissemination of the results.

### **3c. Confidentiality**

Before estimates are released to the public, they must first be screened to ensure that they do not violate the Bureau of Labor Statistics' (BLS) confidentiality pledge. A promise is made by the Bureau to each private sector sample unit that the BLS will not release its employment data to the public in a manner that would allow others to identify the unit. If an estimate fails a predetermined primary confidentiality threshold, such as the p% rule, then the cell can be protected. Whether this protection is suppression, rounding or other method is somewhat dependent on the objective of the survey. Secondary confidentiality protection is also implemented to protect respondent information at this level.

## **4. Testing Procedures**

The BRS is built on years of testing for the QBS and the ARS. This testing has informed the platform used, contact strategies, data processing methods, and expected data review time frames.

The content proposed for inclusion on the BRS questionnaire was cognitively tested by the Office of Survey Methods Research staff at BLS. Cognitive testing occurred with fewer than 10 respondents. A memo summarizing the cognitive testing methods and results is included as an attachment to this package. Overall, the testing indicated that the questions worked as intended and the likely BRS

---

<sup>1</sup> Stang and Thomas "Web Collection in the Quarterly Census of Employment and Wages Program", ICES-V, 2016. [http://ww2.amstat.org/meetings/ices/2016/proceedings/072\\_ices15Final00299.pdf](http://ww2.amstat.org/meetings/ices/2016/proceedings/072_ices15Final00299.pdf)

<sup>2</sup> Stang and Thomas "Email Solicitation for a Business Establishment Survey – Results from the 2015 Annual Refiling Survey", JSM 2016. <http://9004e5e16f4a25df17a0-290e28d0a6d5d71f78b4f59d5f323756.r86.cf1.rackcdn.com/ASA-JSM/pdf/389517.pdf>

<sup>3</sup> Stang and Thomas "Developing and Testing the Business Research Survey," JSM 2018. <https://ww2.amstat.org/meetings/jsm/2018/onlineprogram/AbstractDetails.cfm?abstractid=328621>



Bureau of Labor Statistics  
U.S. Business Response Survey and JOLTS Supplemental Survey  
1220-xxxx  
July 2020

respondents will have the information required to provide accurate answers without consulting records or other establishment employees.

### **5. Statistical and Analytical Responsibility**

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