

SUPPORTING STATEMENT FOR THE OCCUPATIONAL REQUIREMENTS SURVEY

OMB CONTROL NO. 1220-0189

Part B. Collection of Information Employing Statistical Methods

For detailed technical materials on the sample allocation, selection, and estimation methods as well as other related statistical procedures see the BLS Handbook of Methods, BLS technical reports, and American Statistical Association (ASA) papers listed in the references section. The following is a brief summary of the primary statistical features of the Occupational Requirements Survey (ORS).

The Occupational Requirements Survey (ORS) is an establishment survey that the Bureau of Labor Statistics (BLS) is conducting to collect information about the requirements of occupations in the U.S. economy, including the vocational preparation, the cognitive and physical requirements, and the environmental conditions in which the work is performed. The Social Security Administration (SSA), one of several users of this occupational information, is funding the survey through an Interagency Agreement (IAA). Prior planning for ORS involved several feasibility tests throughout Fiscal Years 2013 and 2014 and a three-year production wave beginning in Fiscal Year 2015. The BLS collected ORS data for the three-year production wave using a two-stage stratified design with probability proportional to employment sampling at each stage. Under that design, occupations with low employment in the current economy had a smaller probability of selection resulting in an insufficient number of observations to publish ORS estimates for these low employment occupations.

In late FY 2018, ORS production began a new five-year wave by selecting samples using the methodology described in this document. Collection of the first ORS production sample under this design began in September 2018 and continued for approximately twelve consecutive months. As of November 2020, BLS completed collection of two samples under this design. The full five-year sample design is expected to be complete in July 2023 with final estimates published by December 2023.

Sections 1-3 of this document describe the current selection process of the ORS production samples, the collection process for the ORS data elements, and planned estimates to be produced. Data from the samples are used to produce outputs, such as the "time to proficiency" of occupations, the mental-cognitive and physical demands of work, and the environmental conditions in which work is performed. Section 4 of this document describes the efforts conducted by the BLS to prepare for this second wave of production of the ORS.

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.

1a. Universe

The ORS will measure constructs such as time to proficiency, mental-cognitive and physical demands, and environmental conditions and produce national-level estimates by occupation of percentages, means, percentiles, and modes of variables derived from measurements capturing information about ORS constructs.

The frame for the ORS sample under this design is developed from several sources:

- The Occupational Employment and Wage Statistics (OEWS) sample [previously known as the Occupational Employment Statistics (OES) program] of establishments and occupations. The OEWS sample contains over 1 million establishments from private industry and state and local government.
- A modeled occupation frame created by the OEWS program. The OEWS uses the private industry portion of their sample to predict occupational distributions for not sampled for or non-responding to the OEWS private industry establishments.
- The Quarterly Contribution Reports (QCR) filed by employers subject to State Unemployment Insurance (UI) laws. The BLS receives the QCR for the Quarterly Census of Employment and Wages (QCEW) Program from the 50 States, the District of Columbia, Puerto Rico, and the U.S. Virgin Islands. The QCEW data, which are compiled each calendar quarter, provide a comprehensive business name and address file with employment, wage, detailed geography (i.e., county), and industry information at the six-digit North American Industry Classification System (NAICS) level.
- In many states railroad establishments are not required to report to the State UI. BLS obtains railroad establishment information from State partners that work directly with staff in the office of Occupational Employment Statistics (OEWS).

The ORS universe includes all establishments in the 50 States and the District of Columbia with ownerships of State and Local governments and private sector industries, excluding agriculture, forestry, and fishing (NAICS Sector 11) and private households (NAICS Subsector 814). The estimate of the current universe size, based on the most recent QCEW data, is about 9,500,000 establishments.

Since the OEWS modeled frame only includes establishments in the private industry, separate sampling frames are created for private industry versus government (state and local combined).

- To create the private industry frame, data from the OEWS modeled frame are combined with the establishment-level data from the private-industry QCEW and railroad files to

create lists of occupations at the establishment level. The modeled information is then supplemented with the collected OEWS sample of establishments and occupations to create the full private industry frame of occupations at an establishment level.

- The frame for state and local government is the government establishments from the QCEW.

All ORS production sampled establishments (approximately 10,000 per year) are interviewed once in an attempt to capture all of the needed ORS data elements.

1b. Sample Size

Scope - The ORS production frame is as defined above. The sampling design for the five-year sample is a two-stage stratified random sample of establishments and occupations within selected establishments.

Sample Stratification – Both the private industry and government sector samples are stratified, however, the sample cells (i.e., strata) are defined differently for each sector and the input frames separate.

For private industry, based on data found in the modeled OEWS frame file, all private industry establishments are first identified as either having a “Rare Occupation” or not. For the purposes of sample selection, in most cases a “rare occupation” is defined as one of the 6-digit Standard Occupational Classifications (SOCs) with the lowest May 2017 OEWS employment, across all ownerships.

Strata are then formed by the cross-classification of the predicted presence/absence of a “rare occupation” in the establishment, Census Region (Northeast, Southeast, Midwest, West), and aggregate industry (Education, Goods Producing, Health Care, Financial Activities, Service Providing), leading to forty strata.

Since the OEWS modeled frame does not include State and local government entities, Rare Occupation Status does not apply to the government sector for the purposes of stratification. Thus, for the state and local government frame, strata are formed by the cross-classification of Census Region (Northeast, Southeast, Midwest, West), and detailed industry, also leading to forty strata.

Sample Allocation – The ORS must determine the number of units intended for each sample cell before it selects its establishment sample. The allocation process is run separately for each ownership sector (Private industry or Government).

The total ORS production sample will consist of approximately 50,000 establishments for each five-year production wave. The private portion of this sample is approximately 85% (42,500) and State and Local government portion is approximately 15% of the total sample (7,500). In order to accommodate the goal of ORS, to produce estimates of occupational requirements for as

many Standard Occupational Classification (SOC) codes as possible, a higher proportion of the total private industry sample size are allocated to the twenty “rare-occupation” strata than to the twenty “non-rare occupation” strata. Establishment allocation to the cells within the “rare/non-rare” strata is proportional to total employment within the cell. Establishment allocation to the sample cells for the state and local government sector are proportional to the total employment within the cell.

Note: For the FYE2022 sample year only (sample 724), the number of ORS sampled units is increasing from approximately 10,000 units to 15,000 units in order to mitigate data loss due to pandemic non-response. The private portion of this sample is approximately 85% (12,750) and State and Local government portion is approximately 15% of the total sample (2,250).

Collection of the first ORS production sample under this five-year design spanned a consecutive twelve month period, beginning in the September of 2018.

Sample Selection – Sample selection involves two stages: establishment selection and occupation selection. For private industry, both stages are completed before the sample is fielded, with the exception of establishments that are also in the National Compensation Survey (NCS). For private industry sample units that are also sampled, by chance, for the NCS and for all sampled state and local government establishments, occupational selection is done after establishment contact.

At the first stage of sample selection, all establishments are selected with probability proportional to employment size of the establishment.

For each private industry establishment that is not in the NCS, an occupational quote allocation is assigned based on establishment size, noting that there is one quote per SOC and the quote allocation can only be as large as the total number of distinct SOCs.

BLS acknowledges that some of the allocated quotes will not exist in the sampled establishment. This is because the occupational distribution information for each establishment is a prediction, or a best guess of the occupations that exist in the establishment. Due to imperfections in SOC quote information, BLS samples twice as many occupations as needed for each of the establishment size classes in the following manner:

- Up to 4 employees: Total number of distinct SOCs
- 5 – 49 employees: Up to 8 SOCs/quotes
- 50 – 249 employees: Up to 12 SOCs/quotes
- 250+ employees: Up to 16 SOCs/quotes

Within each selected establishment, the allocated occupations are selected based on the predicted occupational distribution in the following manner.

- In the twenty “non-rare” strata, BLS selects all SOCs/quotes using a systematic sampling strategy.

- In the twenty “rare” strata, if the selected establishment has only “rare” SOCs, BLS selects SOCs/quotes with certainty or using a systematic sampling strategy in accordance with the quote allocation.
- In the twenty “rare” strata, if the selected establishment has a mix of “rare” and “non-rare” SOCs, BLS selects no more than one less the quote allocation from the “rare” SOCs either with certainty or using a systematic sampling strategy, depending on the total number of “rare” SOCs. BLS selects the remaining quotes from the “non-rare” SOCs using a systematic sampling strategy.

The selected occupations are ordered for each establishment. The field economist follows the order until the total number of quotes needed for the establishment size is identified and collected.

- Up to 4 employees: Total number of distinct SOCs
- 5 – 49 employees: Up to 4 SOCs/quotes
- 50 – 249 employees: Up to 6 SOCs/quotes
- 250+ employees: Up to 8 SOCs/quotes

For the private portion of the sample that overlaps with the NCS and the government portion of the sample, jobs are selected in each sampled establishment during the collection phase. The probability of a job being selected within this segment of the ORS sample is proportionate to its employment within the establishment. The number of jobs selected in an establishment ranges as follows:

- Up to 4 employees: Total number of distinct SOCs
- 5 – 49 employees: 4 SOCs/quotes
- 50 – 249 employees: 6 SOCs/quotes
- 250+ employees: 8 SOCs/quotes

Sample weights are assigned to each of the selected establishments and jobs in the sample to represent the entire frame. Units selected as certainty are self-representing and carry a sample weight of one. The sample weight for the non-certainty units is the inverse of the probability of selection. For additional details on the sample design, please see “Evaluation of a Sample Design Based on Predicted Occupational Frame Data” paper by McNulty and Yu (Attachment 1).

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. Sample Design

The design plan for the ORS uses a five-year rotation with complete estimates published after a full five-year sample has been fielded and collected. Limited interim results will be produced on an annual basis for estimates that meet all BLS confidentiality and SSA interagency agreement guidelines. The full five-year sample is split evenly and collected over a five-year period with approximately one-fifth collected each year. The data for establishments in each sample year are collected once and are not be collected again for the ORS until a new production wave is fielded.

2b. Estimation Procedure

The ORS production plan is to produce estimates as described in the formulas below. Computation of these estimates includes weighting the data at both the unit (establishment and occupation/job) and item (individual ORS data element) level. The final weights include the initial sample weights, adjustments to the initial sample weights, two types of adjustments for non-response, and benchmarking. The initial sample weight for a job in a particular establishment is a product of the inverse of the probability of selecting a particular establishment within its stratum and the inverse of the probability of selecting a particular job within the selected establishment. Adjustments to the initial weights are done when data are collected for more or less than the sampled establishment. This may be due to establishment mergers, splits, the inability of respondents to provide the requested data for the sampled establishment, or inaccuracies in the predicted occupational distribution information for the sampled establishment which results in Probability Sampling of Occupations (PSO) being used. The two types of adjustments for non-response include an adjustment for establishment refusal to participate in the survey and an adjustment for respondent refusal to provide data for a particular job.

Benchmarking, or post-stratification, is the process of adjusting the weight of each establishment in the survey to match the distribution of employment by detailed industry at the reference period. Because the sample of establishments is selected from a frame that is approximately two years old by the time the data is used in estimation and sample weights reflect employment when selected, the benchmark process updates that weight based on current employment.

ORS calculates percentages, means, percentiles, and modes for ORS data elements for the nation as a whole by occupation, defined by SOC. ORS uses a 6-digit SOC code resulting in the potential of data for 867 SOC codes. Before estimates of characteristics are released, they are first screened to ensure that they do not violate the BLS confidentiality pledge. A promise is made to each private industry respondent and those government sector respondents who request confidentiality that BLS will not release its reported data to the public in a manner that would allow others to identify the establishment, firm, or enterprise.

Calculate Estimates

ORS estimates are defined in two dimensions. A set of conditions describes the domains and a separate set of conditions describes the characteristics. Domain conditions may include specific occupations, occupational groups, worker characteristics, and geographic region. Characteristic conditions depend on the ORS data elements, such as previous experience or the required number of hours an employee must stand in a typical day. Each characteristic is calculated for

each domain. If a quote meets the domain condition for a particular estimate, the X_{ig} value in the formulas below is 1; otherwise, it is 0. Likewise, if a quote meets the characteristic condition for a particular estimate, the Z_{ig} value in the formulas below is 1; otherwise, it is 0. The final quote weight ensures that each quote used in estimation represents the appropriate number of employees from the sampling frame.

Estimates that use the mean or percentile formulas require an additional quantity for estimation, Q_{ig} , the value of the variable corresponding to this quantity. For more information, see “Estimation Considerations for the Occupational Requirements Survey” by Rhein (see Attachment 2).

Estimation Formulas (All estimates use quote-level records, where quote represents the selected workers within a sampled establishment job.)

1. Percent of employees with characteristic: Percent of employees with a given characteristic out of all employees in the domain. These percentages would be calculated for categorical elements (e.g., type of degree required) and for element durations within SSA categories (e.g., Seldom, Frequently).

$$\frac{\left[\sum_{i=1}^I \sum_{g=1}^{G_i} \text{OccFW}_{ig} X_{ig} Z_{ig} \right]}{\left[\sum_{i=1}^I \sum_{g=1}^{G_i} \text{OccFW}_{ig} X_{ig} \right]} \times 100$$

Estimation Formula Notation

- i = Establishment
- g = Occupation within establishment i
- I = Total number of establishments
- G_i = Total number of quotes selected in establishment i
- X_{ig} = 1 if quote ig meets the condition set in the domain (denominator) condition
= 0 otherwise
- Z_{ig} = 1 if quote ig meets the condition set in the characteristic condition
= 0 otherwise
- OccFW_{ig} = Final quote weight for occupation g in establishment i

To calculate the percent of employees with a given characteristic out of all employees in the domain, add the final quote weights across only those quotes that meet the domain (denominator) condition and characteristic condition. Then divide that number by the sum of the final quote weights across quotes that meet the domain (denominator) condition. Multiply the final quotient by 100 to yield a percentage.

2. Mean: Average value of a quantity for a characteristic. These estimates would be calculated for element durations and other numeric elements.

$$\frac{\left[\sum_{i=1}^I \sum_{g=1}^{G_i} \text{OccFW}_{ig} \times X_{ig} \times Z_{ig} \times Q_{ig} \right]}{\left[\sum_{i=1}^I \sum_{g=1}^{G_i} \text{OccFW}_{ig} \times X_{ig} \times Z_{ig} \right]}$$

Estimation Formula Notation

- i = Establishment
- g = Occupation within establishment i
- I = Total number of establishments in the survey
- G_i = Total number of quotes in establishment i
- X_{ig} = 1 if quote ig meets the condition set in the domain condition
= 0 otherwise
- Z_{ig} = 1 if quote ig meets the condition set in the characteristic condition
= 0 otherwise
- OccFW_{ig} = Final quote weight for occupation g in establishment i
- Q_{ig} = Value of a quantity for a quote g in establishment i

To calculate the average value of a quantity for a characteristic, multiply the final quote weight and the value of the quantity for those quotes that meet the domain (denominator) condition and characteristic condition; add these values across all contributing quotes to create the numerator. Divide this number by the sum of the final quote weights across only those quotes that meet the domain (denominator) condition and characteristic condition.

3. Percentiles: Value of a quantity at given percentile. These estimates would be calculated for element durations and other numeric elements.

The p -th percentile is the value Q_{ig} such that

- the sum of final quote weights (OccFW_{ig}) across quotes with a value less than Q_{ig} is less than p percent of all final quote weights, and
- the sum of final quote weights (OccFW_{ig}) across quotes with a value more than Q_{ig} is less than $(100-p)$ percent of all final quote weights.

It is possible that there are no specific quotes ig for which *both* of these properties hold. This occurs when there exists a quote for which the OccFW_{ig} of records whose value is less than Q_{ig} equals p percent of all final quote weights. In this situation, the p -th percentile is the average of Q_{ig} and the value on the record with the next *lowest* value. The Q_{ig} values must be sorted in ascending order.

Include only quotes that meet the domain condition and the characteristic condition – i.e., where:
 $X_{ig} \times Z_{ig} = 1$.

Estimation Formula Notation

- i = Establishment
- g = Occupation within establishment i
- X_{ig} = 1 if quote ig meets the condition set in the domain condition
= 0 otherwise
- Z_{ig} = 1 if quote ig meets the condition set in the characteristic condition
= 0 otherwise
- $OccFW_{ig}$ = Final quote weight for occupation g in establishment i
- Q_{ig} = Value of a quantity for a specific characteristic for occupation g in establishment i

4. **Modes:** The category with the largest weighted employment from among all possible categories of a characteristic. These estimates will be calculated for all categorical elements (e.g., type of degree required) among the appropriate categories (e.g., bachelor's degree, master's degree).

2c. Reliability

Measuring the Quality of the Estimates

The two basic sources of error in the survey estimates are bias and variance. Bias is the amount by which estimates systematically do not reflect the characteristics of the entire population. Many of the components of bias can be categorized as either response or non-response bias.

Response bias occurs when respondents' answers systematically differ in the same direction from the correct values. For example, this occurs when respondents incorrectly indicate "no" to a certain ORS element's presence when that ORS element actually existed. Another example may occur when, in providing the requested ORS data elements, the respondent focuses only how the selected employee performs the duties in his position, rather than what is required by the position. Response bias can be measured by using a re-interview survey. Properly designed and implemented, this can also indicate where improvements are needed and how to make these improvements. For production, the ORS data will be reviewed for adherence to ORS collection procedures using a multi-stage review strategy. Approximately five percent of the sampled establishments will be re-contacted to confirm the accuracy of coding for selected data elements. The remaining ORS units will either be reviewed in total or for selected data elements by an independent reviewer in the Regional or National Offices. All schedules in the sample will be eligible for one and only one type of non-statistical review, in other words a responding establishment may be re-contacted at most once for an additional review. Additionally, all schedules will be reviewed for statistical validity to ensure the accuracy of the sample weight with the data that was collected.

Non-response bias is the amount by which estimates obtained do not properly reflect the characteristics of non-respondents. This bias occurs when non-responding establishments have ORS element data that are different from those of responding establishments. Non-response bias is being addressed by efforts to reduce the amount of non-response. Another BLS establishment based program, the National Compensation Survey (NCS), has analyzed the extent of non-

response bias using administrative data from the survey frame. The results from this analysis are documented in the 2006 ASA Proceedings of Survey Research Methods Section (See Attachment 3). A follow-up study from 2008 is also listed in the references (See Attachment 4). Details regarding adjustment for non-response are provided in Section 3 below. These studies provide knowledge that can be incorporated into ORS. See Section 3c for more information about non-response studies.

Another source of error in the estimates is sampling variance. Sampling variance is a measure of the variation among estimates from different samples using the same sample design. Sampling variance for the ORS data is calculated using a technique called balanced half-sample replication. For national estimates, this is done by forming different re-groupings of half of the sample units. For each half-sample, a "replicate" estimate is computed with the same formula as the regular or "full-sample" estimate, except that the final weights are adjusted. If a unit is in the half-sample, its weight is multiplied by $(2-k)$; if not, its weight is multiplied by k . For all ORS estimates, $k = 0.5$, so the multipliers will be 1.5 and 0.5. Sampling variance computed using this approach is the sum of the squared difference between each replicate estimate and the full sample estimate averaged over the number of replicates and adjusted by the factor of $1/(1-k)^2$ to account for the adjustment to the final weights. This approach is similar to that used in the NCS. For more details, see the NCS Chapter of the BLS Handbook of Methods (See Attachment 5).

For ORS production, the goal is to generate estimates for as many 6-digit SOC codes as possible, given the sample size and BLS requirement to protect respondent confidentiality and produce accurate estimates. Additional estimates for aggregate SOC codes will be generated if they are supported by the data. Estimates should be accurate with a relative standard error less than 50% on average and the percent estimates are expected to be within 5 percent of the true (population) percent at the 90 percent confidence level.

2d. Data Collection Cycles

ORS production data collection under this design began in September 2018. Collection will span approximately 60 months (12 months for each one-fifth portion of total sample assigned each year) with complete estimates produced at the conclusion of the total five-year design. Limited interim results will be produced on an annual basis for estimates that meet all BLS confidentiality and SSA interagency agreement guidelines. The BLS will conduct ORS as a national survey composed of no more than 50,000 establishments. Approximately 15 percent of these establishments will be selected from State and Local government and the remainder of the sample will be selected from private industry.

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.

There are two types of non-response for ORS: total establishment non-response and partial non-response with the latter occurring at the occupation or data element level. The assumption for all non-response adjustments is that non-respondents are similar to respondents.

To adjust for establishment or occupation non-response, weights of responding units or occupations that are deemed similar will be adjusted appropriately. Establishments are considered similar if they are in the same “rare/non-rare” strata, ownership, and 2-digit NAICS. If there are not sufficient data at this level, then a broader level of aggregation is considered. For partial non-response at the ORS element level, ORS computes estimates that include a replacement value imputed based on information provided by establishments with similar characteristics.

For ORS, the un-weighted establishment response rates for the first two sample groups under the second wave that ended collection in July 2020 was 65%. At the occupation level, the un-weighted response rate for the same sample groups was 77%. Please note that decline in establishment level response rates is attributed to collection during the COVID-19 pandemic.

3a. Maximize Response Rates

To maximize the response rate for this survey, field economists initially refine addresses ensuring contact with the appropriate employer. Then, employers are mailed or emailed a letter explaining the importance of the survey and the need for voluntary cooperation. The letter also includes the Bureau’s pledge of confidentiality. A field economist calls the establishment after the package is sent to attempt to enroll them into the survey. Non-respondents and establishments that are reluctant to participate are re-contacted by a field economist specially trained in refusal aversion and conversion. Additionally, respondents are offered a variety of methods, including personal visit, telephone, fax, and email, through which they can provide data. As a result of collection under the COVID-19 pandemic, the BLS recently added video collection as an option for the respondents to provide ORS data.

3b. Non-Response Adjustment

As with other surveys, ORS experiences a certain level of non-response. To adjust for the non-respondents, ORS divides the non-response into two groups, 1) unit non-respondents and 2) item non-respondents. Unit non-respondents are the establishments (or occupations) for which no ORS data was collected, whereas item non-respondents are the establishments that report only a portion of the requested ORS data elements for the selected occupations.

The unit (establishment or occupation) non-response is treated using a Non-Response Adjustment Factor (NRAF). Within each sampling cell, NRAFs are calculated based on the weighted ratio of the number of viable, i.e., in-scope and sampled, establishments to the number of usable, i.e., provided any data, respondents in the sample cell. Item non-response is adjusted using item imputation.

3c. Non-Response Bias Research

Prior research was done to assess whether non-respondents to the NCS survey differ systematically in some important respect from respondents and would thus bias NCS estimates. Details of this study are described in the two papers by Ponikowski, McNulty, and Crockett referenced in Section 2c (See Attachments 3 and 4). These studies provided knowledge that can be incorporated into future ORS non-response bias research.

BLS also analyzed survey response rates from the Pre-production test of the ORS sample at the establishment, occupational quote, and item (i.e., individual data element) levels. The data was analyzed using un-weighted response rates and response rates weighted by the sample weight at each level of detail. Results from the Pre-production test are detailed in the paper by Yu, Ponikowski, and McNulty (see Attachment 6). In a continued effort to monitor response rates at the establishment, occupation, and item levels, the BLS will run the same non-response analysis at the conclusion of each production sample.

BLS plans to review the response rates in aggregate and by available auxiliary variables such as industry, occupation, geography, e.g., Census and BLS data collection regions, and establishment size. BLS will use the results from the analysis to identify the auxiliary variables that are most likely to contribute significantly to bias reduction. This research is expected to continue through 2022. Once these variables are identified they will be used in the data processing system to reduce potential nonresponse bias.

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.

Various tests have been completed both prior to the start and during the first wave of the ORS production samples. Field testing focused on developing procedures, protocols, and collection aids. These testing phases were analyzed primarily using qualitative techniques but showed that this survey was operationally feasible. Survey design testing was also conducted to ensure that we have the best possible sample design to meet the needs of the ORS. Data review processes and validation techniques were also analyzed to ensure quality data can be produced.

4a. Tests of Collection Procedures

The timeline below is an overview of past and current testing of collection procedures.

Past

- Fiscal Year 2012

The BLS signed an interagency agreement with SSA to design, develop, and conduct a series of tests using the NCS platform. The purpose was to assess the feasibility of using the NCS to accurately and reliably capture data relevant to SSA's disability program. The resulting initiative—the ORS—was launched to capture data elements new to NCS using the NCS survey platform.

- Fiscal Year 2013

BLS completed three initial phases of ORS testing: a Proof-Of-Concept Test, a Collection Protocol Development Test, and a Broad Scale Test of various protocols. Results from this testing indicated that it is feasible for BLS to collect data relevant to the SSA's disability program using the NCS platform. Details on these collection tests are documented in the "Testing the Collection of Occupational Requirements Data" report found in the 2013 ASA Papers and Proceedings (see Attachment 7).

The results of Phase 1's Proof-Of-Concept Test suggested that BLS' approach is viable. Additional information on this test and the lessons learned are available in the "Occupational Requirement Survey, Phase 1 Summary Report, Fiscal Year 2013" (see Attachment 8).

Phase 2's Collection Protocol Development Test evaluated ORS collection protocols and aids. The results of Phase 2 testing, which can be found in the "Occupational Requirement Survey, Phase 2 Summary Report, Fiscal Year 2013" (see Attachment 9), demonstrated the effectiveness of the revised materials and procedures and the continued viability of BLS collection of data relevant to the SSA's disability program.

Phase 3's Broad Scale Testing was designed to show whether ORS field economists from across the country could collect all of the ORS data elements in addition to wages and leveling information in a uniform and efficient manner. The Phase 3 testing demonstrated the effectiveness of the revised materials and procedures and the continued viability of BLS collection of data relevant to the SSA's disability program. The details of this test and the results are further documented in the "Occupational Requirement Survey, Phase 3 Summary Report, Fiscal Year 2013" (see Attachment 10).

- Fiscal Year 2014

The BLS completed six feasibility tests to refine the ORS methodology. In general, the results from these tests confirmed BLS' viability at collecting data relevant to ORS and demonstrated the effectiveness of the revised materials and procedures tested. More

detailed information on these feasibility tests as well as key findings can be found in the “Occupational Requirement Survey, Consolidated Feasibility Tests Summary Report, Fiscal Year 2014” (see Attachment 11).

- Fiscal Year 2015

The Pre-Production test was designed to test all survey activities by mirroring production procedures, processes and protocols as closely as possible. Every normal production activity associated with each of BLS’ product lines was conducted during Pre-Production testing. More detailed information on the Pre-Production test can be found in the “The Occupational Requirements Survey: estimates from preproduction testing” (see Attachment 12).

In Fiscal Year 2015, ORS production sample units were selected using a 2-stage stratified design with probability proportional to employment sampling at each stage. The total ORS production sample in the first year was 4,250 establishments and 10,000 establishments for the next two years. For more details on this design see paper by Ferguson and McNulty on “Occupational Requirements Survey Sample Design” (Attachment 13).

BLS also conducted a job observation test during the summer of 2015 to provide validation for the ORS physical elements by comparing the data collected during pre-production to those collected through direct job observation, which is more typical among small scale studies of job tasks. More details and results from this test can be found in the paper titled “Occupational Requirements Survey Job Observation Report” (see Attachment 14).

- Fiscal Year 2016 to Present

The ORS program completed collection of the three sample groups from the first production wave in the summer of 2018. During FY2017 and FY2018, three tests of data collection methods were conducted. The first test continued the BLS work to validate the ORS data and methodology. It was a larger scale version the FY2015 Job Observation test, and focused on ORS elements and occupations that are amenable to testing by observation. Research results from both observation tests are summarized in the “Observational collection compared to critical task threshold interview collection in the Occupational Requirements Survey” article on the www.bls.gov website.

Data for the second test was collected between September and November 2017 and focused on a comprehensive set of questions on the mental/cognitive demands for a job. Earlier cognitive data collection questions did not yield data that would meet SSA’s needs for adjudication and were discontinued in August 2017 after OMB approval was received on 4/28/2017. New questions were designed and tested on a limited basis through the BLS Office of Survey Methods Research (OSMR) generic Clearance 1220-0141 in the first half of 2017. The outcome of the test is a revised set of mental/cognitive questions and response answers incorporated into the ORS survey.

The final test was the FY2018 Occupational Selection Test. The primary goal of this test was to evaluate the new occupational quote selection described in section 1 above and the impact of this change on training and collection procedures. The test included a range of establishments in order to accurately refine data collection procedures and provide insights for field economist training. This test was conducted over a six week period in the Spring of 2018. The data and lessons learned from this test were used to develop comprehensive instructions and training tools for the BLS Field Economists to be used in the collection of all sample groups under wave two of the ORS.

Standardized Initial Contact Letter Test

ORS requested and received approval in June 2020 to conduct a standardized contact letter test. Introductory letters are a method used to maximize response rates. The current practice is for field economists and regional office management to determine when to send an introductory letter, and the language of such letters varies. BLS began testing a standardized letter to determine what impact the letter will have on response rates.

The test involves sending a standardized contact letter to 3,000 private establishments. No other aspect of information collection is modified. BLS does not expect using a standardized letter to impact respondent burden because the clearance hours already include time for introductory letters. The letter is intended to reduce the amount of time BLS staff spend providing this information in a non-standardized format and potentially increase response by providing timely standardized information. Response rates for the group that receives the standardized mailing will be compared to response rates from the remaining sample. From this testing and evaluation, BLS will determine if a standardized letter should be used. Due to coronavirus pandemic, the test could not progress as planned and evaluation is postponed to a future date.

4b. Tests of Survey Design Processes

Sample Design Options

To further ensure the BLS met the needs of the ORS by producing statistically valid and high quality data, testing on possible sample design options was also conducted. In FY 2013, the BLS began work to evaluate sample design options for ORS by reviewing the sample designs used for the NCS. More details on this initial sample design testing is available in the November 2013 FCSM Proceedings, “Sample Design Considerations for the Occupational Requirements Survey” (see Attachment 16). This research continued into FY 2014 and expanded to look at other BLS surveys, including the OEWS and Survey of Occupational Injuries and Illnesses (SOII). Since the ORS will be collected by trained field economists who also collect the NCS data, potential coordination with the NCS sample design was a key factor of consideration. As a result, four basic categories of ORS survey designs were identified to allow for different potential levels of coordination with NCS. These design options, are documented in the ASA 2014 Papers and

Proceedings titled “Occupational Requirement Survey, Sample Design Evaluation” by Ferguson et al (see Attachment 17).

While desirable for the ORS sample design to be integrated with NCS, it was unclear whether the NCS sample design would meet the goals of ORS. After various testing on the four basic categories of ORS survey designs, the BLS determined that the Independent Survey Design was the optimal design option for the first production wave. This design, as demonstrated through the first wave production, met the requirements of being able to produce reliable estimates for ORS data elements; however, it did not meet the needs of the SSA in terms of its ability to produce reliable estimates for ORS data elements for the vast majority of SOC codes.

To improve the balance of the number of observations (quotes) sampled across all occupations and increase the publication rate across a greater number of occupations while maintaining current resource levels, in FY2017 the BLS began additional research into alternative sample design options for the ORS. Extensive research, including simulating hypothetical samples, analyzing sample allocations, and estimating the predicted number of observations per occupation per hypothetical sample, was completed for each option prior to coming to a final design. The options studied included:

1. Modify current ORS industry sample allocations but maintain the remaining design features.
2. Modify ORS industry sample allocations and PSO procedures but maintain the remaining design features.
3. Construct sample from subsamples that each target a specific group of occupations.
4. Target ORS sample to pairs of low employment occupations.
5. Two-phase stratified sampling to target specific occupations of interest.
6. Multiple frames stratified sampling to target specific occupations of interest.
7. Two-stage stratified sampling to target rare occupations
 - Description: The sampling design for the five-year private industry sample is a two-stage stratified sample of private industry establishments and occupations within selected establishments. Strata are formed by the cross-classification of the predicted presence/absence of a “rare occupation” in the establishment, Census Region (Northeast, Southeast, Midwest, West), and aggregate industry (Education, Goods Producing, Health Care, Financial Activities, Service Providing), leading to forty strata. For the purposes of sample selection, a “rare occupation” is defined as one of the 200 6-digit SOCs with the lowest May 2017 OEWS employment, across all ownerships.
 - Observations: This design limits oversampling of higher employment occupations by allocating more sample in occupations that would have a lower probability of selection under a probability proportional to occupational employment. Based on simulation results after five years almost 800 6-digit SOCs would meet publication criteria.

After reviewing the results of each of the above design approaches, option 7 yielded the most promising results to reach the goal of publishing estimates for a broader number of SOC codes. This design also has the potential to save time for both the Field Economist as well as the respondent by reducing the number of establishments for which the selection of occupations is completed during collection.

Data Review and Validation Processes

BLS has developed a variety of review methods to ensure data of quality are collected and coded. These methods include data review and validation processes and are available in more detail in the 2014 ASA Papers and Proceedings under the title “ Validation in the Occupational Requirements Survey: Analysis of Approaches” by Smyth (see Attachment 18).

The ORS Data Review Process is designed to create the processes, procedures, tools, and systems to check the micro-data as they come in from the field. This encompasses ensuring data integrity, furthering staff development, and ensuring high quality data for use in producing survey tabulations or estimates for validation. The review process is designed to increase the efficiency of review tools, build knowledge of patterns and relationships in the data, develop expectations for reviewing the micro-data, help refine procedures, aid in analysis of the data, and set expectations for validation of tabulations or future estimates.

To further ensure the accuracy of the data, the ORS Validation Process focuses on aggregated tabulations of weighted data as opposed to individual data. This entails a separate but related set of activities from data review. The goal of the validation process is to review the estimates and declare them Fit-For-Use (FFU), or ready for use in publication and dissemination, as well as confirming that our methodological processes (estimation, imputation, publication and confidentiality criteria, and weighting) are working as intended. Validation processes include investigating any anomalous estimates, handling them via suppressions or correction, explaining them, documenting the outcomes, and communicating the changes to inform any up/down-stream processes. All results of validation are documented and archived for future reference if necessary.

Overall, the ORS poses review and validation challenges for the BLS because of the unique nature of the data being collected. In order to better understand the occupational requirements data, the BLS engaged in a contract with Dr. Michael Handel, a subject matter expert. From the fall of 2014 through January 2015, Dr. Handel reviewed and analyzed literature related to the reliability and validity of occupational requirements data. At the conclusion of his work, Dr. Handel provided the BLS with the recommendations below with the understanding that the ORS is complex in nature and there is no “one size fits all” approach for testing reliability and validity of the data items:

- The development of a strategic documentation to guide methodological research.
- An evaluation on the existence of “gold standard” benchmarks for methods of data collection and for data elements.
- For data elements without any gold standards, multiple approaches may be used.

- Measures of agreement for ORS data should consist of assessing data agreement within method, as opposed to across methods. Because there are many characteristics of the interview that may cause variability (e.g. characteristics of the respondent, length of interview, characteristics of the job and establishment, identity of the field economist/field office), it would be significant to use debriefs with the field economists to identify the key characteristics of the interview to focus on for measures of reliability.
- Consideration should be given to variation caused by errors in coding occupations.

BLS management agreed with the recommendations provided by Dr. Handel. As a result, the BLS began a review initiative in FY 2015 including the development of a methodological guide, evaluation of “gold standard” benchmarks for data collection, and testing of inter-rater reliability (see “Occupational Requirements Survey Job Observation Report,” Attachment 14 for additional information). More detailed information on Dr. Handel’s proposals are explained in an Executive Summary paper titled “Methodological Issues Related to ORS Data Collection” by Dr. Handel (see Attachment 19). These recommendations, as well as refinements of the ORS manual, the data review process, and the validation techniques developed to date ensured ORS products are quality occupational data in the areas of vocational preparation, mental-cognitive and physical requirements, and environmental conditions.

Throughout the testing stages as well as the first wave of production for the ORS, BLS conducted various calibration activities. As stated in a paper by Mockovak, Yu & Earp, calibration training is a type of refresher training that compares interviewer performance against predetermined standards to assess rating accuracy, inter-rater reliability, and other measures of performance. In addition to those described in this paper, the BLS conducts calibration exercises to test staff understanding and adherence to problematic concepts. Most recently, in FY2018, calibration activities focused on SOC coding to better ensure the accuracy and consistency across all National and Regional Office staff that are involved in collection, coding and/or review of ORS microdata. Information obtained during the various calibration activities is used to enhance procedural guidance as well as training materials. Additional information is available in the “Using Calibration Training to Assess the Quality of Interviewer Performance,” ASA Papers and Proceedings (see attachment 20).

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.

Xingyou Zhang, Chief, Statistical Methods Group of the Office of Compensation and Working Conditions, is responsible for the statistical aspects of the ORS production. Xingyou Zhang can be reached on 202-691-6082. BLS seeks consultation with other outside experts on an as-needed basis.

6. References

- Erin McNulty and Alice Yu, (November 2019), "Evaluation of a Sample Design Based on Predicted Occupational Frame Data," ASA Papers and Proceedings, <https://www.bls.gov/osmr/research-papers/2019/pdf/st190050.pdf> (Attachment 1)
- Bradley D. Rhein, Chester H. Ponikowski, and Erin McNulty, (October 2014), "Estimation Considerations for the Occupational Requirements Survey," ASA Papers and Proceedings, <https://www.bls.gov/osmr/research-papers/2014/pdf/st140080.pdf> (Attachment 2)
- Chester H. Ponikowski and Erin E. McNulty, (December 2006), "Use of Administrative Data to Explore Effect of Establishment Nonresponse Adjustment on the National Compensation Survey", ASA Papers and Proceedings, <https://www.bls.gov/osmr/research-papers/2006/pdf/st060050.pdf> (Attachment 3)
- Chester H. Ponikowski, Erin McNulty and Jackson Crockett (October 2008) "Update on Use of Administrative Data To Explore Effect of Establishment Nonresponse Adjustment on the National Compensation Survey Estimates", ASA Papers and Proceedings, <https://www.bls.gov/osmr/research-papers/2008/pdf/st080190.pdf>, (Attachment 4)
- Bureau of Labor Statistics' Handbook of Methods, Bureau of Labor Statistics, <https://www.bls.gov/opub/hom/ors/pdf/ors.pdf>, (Attachment 5)
- Alice Yu, Chester H. Ponikowski, and Erin McNulty (October 2016), "Response Rates for the Pre-Production Test of the Occupational Requirements Survey," ASA Papers and Proceedings, <https://www.bls.gov/osmr/research-papers/2016/pdf/st160250.pdf> (Attachment 6)
- Gwyn R. Ferguson, (October 2013), "Testing the Collection of Occupational Requirements Data," ASA Papers and Proceedings, <https://www.bls.gov/osmr/research-papers/2013/pdf/st130220.pdf> (Attachment 7)
- The ORS Debrief Team, (January 2013) "Occupational Requirements Survey, Phase 1 Summary Report, Fiscal Year 2013," Bureau of Labor Statistics, <https://www.bls.gov/ors/research/collection/pdf/phase1-report.pdf>, (Attachment 8)
- The ORS Debrief Team, (April 2013) "Occupational Requirements Survey, Phase 2 Summary Report" Bureau of Labor Statistics, <https://www.bls.gov/ors/research/collection/pdf/phase2-report.pdf>, (Attachment 9)
- The ORS Debrief Team, (August 2013) "Occupational Requirements Survey, Phase 3 Summary Report" Bureau of Labor Statistics, <https://www.bls.gov/ors/research/collection/pdf/phase3-report.pdf>, (Attachment 10)

The ORS Debrief Team, (November 2014) "Occupational Requirements Survey, Consolidated Feasibility Tests Summary Report, Fiscal Year 2014," Bureau of Labor Statistics, <https://www.bls.gov/ors/research/collection/pdf/fy14-feasibility-test.pdf>, (Attachment 11)

Nicole Dangermond, (November 2015) "The Occupational Requirements Survey: estimates from preproduction testing," Monthly Labor Review, <https://www.bls.gov/opub/mlr/2015/article/the-occupational-requirements-survey.htm>, (Attachment 12)

Gwyn R. Ferguson and Erin McNulty, (October 2015), "Occupational Requirements Survey Sample Design," <https://www.bls.gov/osmr/research-papers/2015/pdf/st150060.pdf>, (Attachment 13)

The ORS Job Observation Test Team, (November 2015), "Occupational Requirements Survey Job Observation Report," <https://www.bls.gov/ors/research/collection/pdf/preproduction-job-observations-report-2015.pdf>, (Attachment 14)

Bradley D. Rhein, Chester H. Ponikowski, and Erin McNulty, (November 2013), "Sample Design Considerations for the Occupational Requirements Survey," FCSM Papers and Proceedings, https://nces.ed.gov/FCSM/pdf/H4_Rhein_2013FCSM_AC.pdf, (Attachment 15)

Gwyn R. Ferguson, Erin McNulty, and Chester H. Ponikowski (October 2014), "Occupational Requirements Survey, Sample Design Evaluation," ASA Papers and Proceedings, <https://www.bls.gov/osmr/research-papers/2014/pdf/st140130.pdf>, (Attachment 16)

Kristin N. Smyth, (October 2014), "Validation in the Occupational Requirements Survey: Analysis of Approaches," ASA Papers and Proceedings <https://www.bls.gov/osmr/research-papers/2014/pdf/st140210.pdf>, (Attachment 17)

Michael Handel, (February 2015), "Methodological Issues Related to ORS Data Collection," <https://www.bls.gov/ors/research/collection/pdf/handel-methodological-issues-data-collection-exec-summary-feb15.pdf>, (Attachment 18)

William Mockovak, Alice Yu & Morgan Earp, (September 2015), "Using Calibration Training to Assess the Quality of Interviewer Performance," ASA Papers and Proceedings, <https://www.bls.gov/osmr/research-papers/2015/pdf/st150010.pdf>, (Attachment 19)