

## Supporting Statement

### Occupational Requirements Survey (ORS) Production

#### **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 technical reports and American Statistical Association (ASA) and Federal Committee on Statistical Methodology (FCSM) papers listed in the references section. The following is a brief summary of the primary statistical features of the Occupational Requirements Survey.**

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 new 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. These tests examined the feasibility of gathering the basic information desired and the availability of data, the efficiency of alternative collection procedures, and the probable degree of cooperation from respondents. The BLS is currently collecting ORS data for a Pre-Production test designed to closely mirror production procedures and protocols. After the Pre-Production test is completed, collection of the first ORS production sample will begin. Sections 1-3 of this document describe the selection of the ORS production samples, the collection process for the ORS data elements, and planned estimates to be produced. Data from the samples will be used to produce outputs, such as the indicator of "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 production of the ORS.

In FY 2015, ORS production will begin by selecting samples using the methodology described in this document. Current plans call for collection of the first ORS production sample to begin in September 2015 and continue for approximately nine consecutive months.

## **1. Universe and Sample Size**

### **1a. Universe**

The ORS will measure information such as time to proficiency, mental-cognitive and physical demands, and environmental conditions by percentage, mean, percentile, and mode for national level estimates by occupation. The universe for this survey will consist of 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.

The potential respondent universe that will be used in the selection of the ORS production sample of establishments is derived from the QCEW and a supplementary file of railroads for each stratum in the sample. 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 (OES). The ORS universe will include all establishments in the 50 States and the District of Columbia with ownership of State and Local governments and private sector industries, excluding agriculture, forestry, and fishing (NAICS Sector 11) and private households (NAICS Subsector 814). Estimate of the current universe size, based on the 2<sup>nd</sup> quarter 2014 QCEW data, is about 9,000,000 establishments. Estimates of the current sample sizes are 4,250 establishments for the first year and 10,000 establishments for subsequent years. Data for the duties and responsibilities of a sample of jobs will be collected in all sampled establishments.

Data from all of the ORS production sampled establishments (4,250 in Year 1 and 10,000 in subsequent years) will be collected once to capture all of the needed ORS data elements. Data will be collected for no more than one sample in a three year period for any given establishment. For sampled establishments selected in multiple samples, data from the most recent collection will be used in estimation for the new sample unit with no additional contact with respondent.

### **1b. Sample Size**

*Scope* - The ORS production frame is as defined above, and an independent sample will be selected each year from this establishment sampling frame using the most recent quarter of data available and as defined below.

*Sample Stratification* – All ORS production sample units will be selected using a 2-stage stratified design with probability proportional to employment sampling at each stage. The first stage of sample selection will be a probability sample of establishments and the second stage will be a probability sample of jobs within sampled establishments. Each sample of establishments will be drawn by first stratifying the establishment sampling frame by defined industry and ownership. The 23 industry strata for private industry and 10 for government shown below are based on the 2012 North American Industry Classification System (NAICS). The industry strata

are defined in the chart below. Frame (Establishments in Universe) counts are based on data from the 2<sup>nd</sup> quarter of 2014.

<b>ORS Production Stratification</b>		<i>Private Industry</i>		<i>State &amp; Local Govt.</i>	
<b>Detailed Industry</b>	<b>Included NAICS Codes</b>	<b>Establishments in Universe</b>	<b>Expected Sample Size</b>	<b>Establishments in Universe</b>	<b>Expected Sample Size</b>
Educational Services (Rest of)	61 (excl. 6111-6113)	80,655	22	1,251	2
Elementary and Secondary Schools	6111	17,249	23	61,828	247
Junior Colleges, Colleges and Universities	6112, 6113	8,671	36	7,550	78
Mining	21	35,471	27	6,339	7
Construction	23	746,906	197		
Manufacturing	31-33	336,416	385		
Healthcare, Social Assistance (Rest of)	62 (excl. 622, 623)	1,277,003	311	8,470	11
Hospitals	622	8,829	149	2,644	34
Nursing and Residential Care Facilities	623	74,363	103	2,081	7
Utilities	22	17,382	18	12,863	20
Wholesale Trade	42	618,652	184		
Retail Trade	44-45	1,036,005	483		
Transportation and Warehousing	48-49	228,101	138		
Information	51	149,018	86	20,678	29
Finance (Rest of)	52 (excl. 524)	282,821	109		
Insurance	524	187,344	69		
Real Estate, Renting, Leasing	53	357,544	65		
Professional, Scientific, Technical	54	1,104,197	262		
Management of Companies and Enterprises	55	59,734	68		
Admin., Support, Waste Management	56	494,313	273		
Arts, Entertainment, Recreation	71	131,361	74		
Accommodation and Food Services	72	656,131	405		
Other Services (excl. Public Administration)	81 (excl. 814)	572,205	127		
Public Administration	92 excl. 928			109,075	202

Within the industry stratification, the ORS production sample will be implicitly stratified by the 24 geographic areas listed below. This will ensure that the sample includes establishments from all parts of the country.

## 24 Geographic Areas

Atlanta-Sandy Springs-Gainesville, GA-AL CSA
Boston-Worcester-Manchester, MA-NH CSA
Chicago-Naperville-Michigan City, IL-IN-WI CSA
Dallas-Fort Worth, TX CSA
Detroit-Warren-Flint, MI CSA
Houston-Baytown-Huntsville, TX CSA
Los Angeles-Long Beach-Riverside, CA CSA
Minneapolis-St. Paul-St. Cloud, MN-WI CSA
New York-Newark-Bridgeport, NY-NJ-CT-PA CSA
Philadelphia-Camden-Vineland, PA-NJ-DE-MD CSA
San Jose-San Francisco-Oakland, CA CSA
Seattle-Tacoma-Olympia, WA CSA
Washington-Baltimore-No. Virginia, DC-MD-VA-WV CSA
Miami-Fort Lauderdale-Miami Beach, FL MSA
Phoenix-Mesa-Scottsdale, AZ MSA
Rest of New England Census Division (excl. areas above)
Rest of Middle Atlantic Census Division (excl. areas above)
Rest of East South Central Census Division (excl. areas above)
Rest of South Atlantic Census Division (excl. areas above)
Rest of North Central Census Division (excl. areas above)
Rest of West North Central Census Division (excl. areas above)
Rest of West South Central Census Division (excl. areas above)
Rest of Mountain Census Division (excl. areas above)
Rest of Pacific Census Division (excl. areas above)

*Sample Allocation* - The total ORS production sample will consist of approximately 4,250 establishments for Year 1 and 10,000 establishments for subsequent years. The private portion of this sample will be approximately 85% (3,613 for Year 1, and projected 8,500 for the subsequent years). The State and Local government sample will be approximately 15% of the total sample (637 for Year 1, and projected 1,500 for the subsequent years). The establishment sample allocation to the industry strata will be proportional to stratum employment resulting in the expected sample sizes for the first production sample shown in the ORS Production Stratification table earlier. Sample sizes in subsequent years will be larger than those shown in the prior table.

Collection of the ORS production sample for Year 1 will take approximately 9 months, beginning in the fall of 2015 after the conclusion of the Pre-Production test. Collection of the samples after Year 1 will take approximately 12 months each after an initial three month refinement period that overlaps collection of the prior sample. Therefore, in order to complete the Year 1 collection and produce estimates within a condensed time schedule by the end of FY 2016, the Year 1 sample will be reduced to 4,250 establishments.

*Sample Selection* – The ORS production sample will use systematic sampling with the probability of selection proportionate to the measure of size. The measure of size (MOS) will be the frame unit (establishment) employment.

After the sample of establishments is drawn, jobs will be selected in each sampled establishment. The number of jobs selected in an establishment will range from 4 to 8 depending on the total number of employees in the establishment, except for government, aircraft manufacturing units, and units with less than 4 workers. In government, the number of jobs selected will range from 4 to 20. In aircraft manufacturing, the number of jobs selected will range from 4 for establishments with less than 50 workers to 32 for establishments with 10,000 or more workers. In establishments with less than 4 workers, all jobs will be selected. The probability of a job being selected will be proportionate to its employment within the establishment.

Sample weights will be assigned to each of the selected establishments and jobs in the sample to represent the entire frame. Units selected as certainty will be self-representing and will carry a sample weight of one. The sample weight for the non-certainty units will be the inverse of the probability of selection.

In addition to the 4,250 units selected in year one and totally independent from the design described above, an supplemental 2,250 units from the current National Compensation Survey (NCS) private sample will be included for collection of the ORS data elements. These units are all active for the NCS but will be rotating out of the NCS sample after October of 2015. The NCS design at the time that these units were selected was a three-stage stratified design with Probability Proportionate to Size used at each level. The first stage was the selection of 152 areas with the second stage being the selection of establishments within each selected area. The final stage was the selection of occupations within sampled establishments. For more details on this design see “Evaluating Sample Design Issues in the National Compensation Survey” by Ferguson et al (see Attachment 1).

Since the two parts to the year one sample will be selected independent of each other and both have weights that represent the entire frame, weight adjustment factors will be applied to all sample units in order to properly represent the frame. The data from this full sample will be used in the estimates for a minimum of three years as described below.

Following the year one sample, all subsequent samples will be selected independent of the NCS sample and will not be supplemented by the NCS.

## **2. Sample Design**

### **2a. Sample Rotation**

The current plan for the ORS is to use a three-year rotation with the most recent three years of sample data used in estimation. The data for establishments selected in each yearly sample will be collected for one sample and will not be collected again for the ORS until at least three sample years later. For establishments selected more than once in a three year period, data will

be copied from the existing sample unit. This number of years is subject to change based on ongoing research exploring how often the requirements of work change.

## **2b. Estimation Procedure**

The ORS production plan is to produce estimates as described in the formulas below. Computation of these estimates will include weighting the data at both the unit (establishment and occupation/job) and item (individual ORS data element) level. The final weights will 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 will be a product of the inverse of the probability of selecting a particular establishment within detailed industry and the inverse of the probability of selecting a particular job within the selected establishment. Adjustments to the initial weights will be done when data are collected for more or less than the sampled establishment. This may be due to establishment mergers, splits, or the inability of respondents to provide the requested data for the sampled establishment. The two types of adjustments for non-response will include adjustment for establishment refusal to participate in the survey and 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 will update that weight based on current employment.

ORS will calculate percentages, means, percentiles, and modes for ORS data elements for the nation as a whole by occupation, defined by the Standard Occupation Code, SOC. ORS will use an 8-digit SOC code defined by O\*NET, resulting in the potential of data for 1,110 SOC codes. Before estimates of characteristics are released, they will first be 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 which would allow others to identify the establishment, firm, or enterprise.

### Calculate Estimates

ORS estimates will be 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 must be calculated for each domain (alternatively, each domain must be calculated for each characteristic). 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 record 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).

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

**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
- $\text{OccFW}_{ig}$  = 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.

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

**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, amount of stooping) among the appropriate categories (e.g., bachelor's degree, master's degree or seldom, occasionally).

## 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. Response bias can also occur when data are collected for a unit other than the sampled unit. For example, the respondent's focus on the question may be altered from what is required of the employee in that position to what the current employee is doing in that position. This misunderstanding would alter the sampled unit to a particular person rather than the sampled occupation. 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. 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 elements 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 initial 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 fluctuation between estimates from different samples using the same sample design. Sampling variance for the ORS data will be 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).

Variance estimation also identifies industries and occupations that contribute substantial portions of the sampling variance. Allocating more sample units to these domains often improves the efficiency of the sample. These variances will be considered in allocation and selection of future samples.

For the ORS production, the goal is to generate estimates for as many 8-digit SOCs as maintained by O\*NET 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 of levels should be accurate with a relative standard error less than 33% 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**

The ORS production data collection will be an ongoing process that will begin after the conclusion of the Pre-Production test and upon receipt of OMB approval. The BLS will conduct ORS as a nationwide survey composed of approximately 4,250 establishments for Year 1 and a projected 10,000 establishments for the subsequent years. Approximately fifteen percent of these establishments will be selected from State and Local government and the remainder of the sample will be selected from private industry. If an establishment selected in the current year was previously collected within the past 2 years, a collection will not be carried out and the establishment's data will simply be duplicated in our database.

## **3. Non-Response**

There are two types of non-response for ORS: total establishment non-response and partial non-response. The non-responses can occur at the establishment level, occupation level, or ORS 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 to be similar will be adjusted appropriately. Establishments will be considered similar if they are in the same ownership and 2-digit NAICS. If there are no sufficient data at this level, then a broader level of aggregation will be considered.

For partial non-response at the ORS element level, ORS will either produce estimates showing the amount of item non-response or will compute estimates that include a replacement value imputed based on information provided by establishments with similar characteristics. If imputation is used, it will be done separately for each ORS element using processes that are currently under evaluation. In some cases, BLS may compute estimates based solely on provided data and also report the percentage of non-provided data.

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

To maximize the response rate for this survey, field economists will initially refine addresses ensuring contact with the appropriate employer. Then, employers will be mailed a letter explaining the importance of the survey and the need for voluntary cooperation. The letter will also include the Bureau's pledge of confidentiality. A field economist will call the establishment after the package is sent and attempt to enroll them into the survey. Non-respondents and establishments that are reluctant to participate will be re-contacted by a field economist specially trained in refusal aversion and conversion. Additionally, respondents will be offered a variety of methods, including personal visit, telephone, fax, and email, through which they can provide data.

The Alternative Modes Test was conducted as part of the FY 2014 Feasibility Tests approved by OMB under control number 1220-0164 to determine how to collect high quality ORS data via phone, email, or fax. This led to the identification of best practices for multiple-mode collection and modifications needed to tools and procedures for future research and testing. There is a continuous effort to maximize response rates, including the use of alternative modes of data collection to obtain quality data from the ORS respondents. Details of the Alternative Modes test can be found in the "Occupational Requirement Survey, Consolidated Feasibility Tests Summary Report, Fiscal Year 2014" (see Attachment 6).

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

As with other surveys, ORS experiences a certain level of non-response. To adjust for the non-respondents, ORS plans to divide 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 will be treated using a Non-Response Adjustment Factor (NRAF). Within each sampling cell, NRAFs will be calculated based on the weighted ratio of the number of viable establishments to the number of usable respondents in the sample cell. Item non-response will be adjusted using item imputation.

### **3c. Non-Response Bias Research**

Extensive 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 ORS; ORS data will be collected from typical NCS respondents and there are many similarities between the data elements collected for both NCS and ORS.

BLS will analyze survey response rates from the ORS production at the establishment, occupational quote, and item (i.e., individual data element) levels. The data will be analyzed using un-weighted response rates and response rates weighted by the sample weight at each level

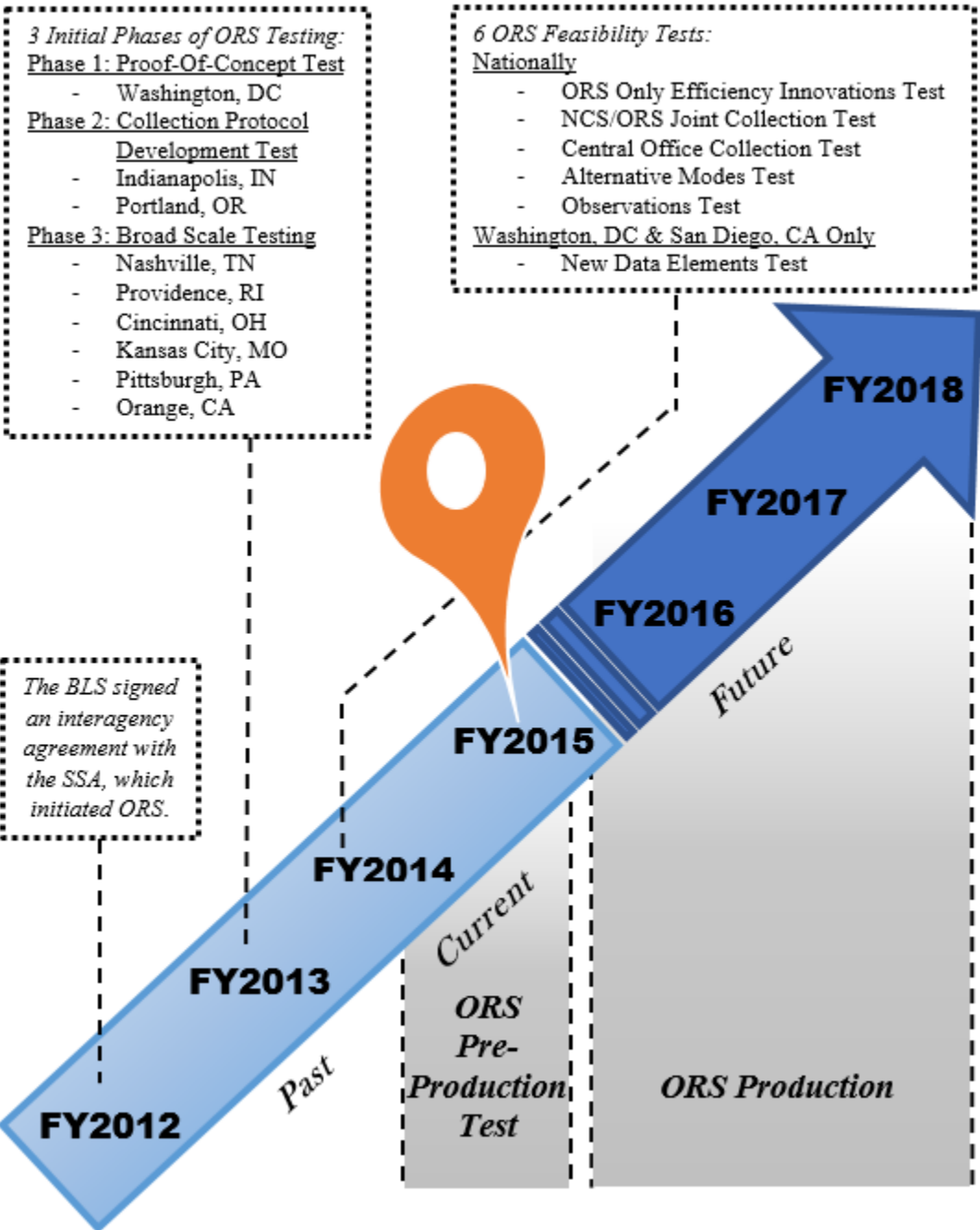
of detail. We plan to review the response rates in aggregate and by available auxiliary variables such as industry, occupation, geography (i.e. Census regions and BLS data collection regions), and establishment size. BLS will use the results from the analysis to identify the auxiliary variables that are most likely to contribute significantly to bias reduction. Once these variables are identified they will be used in the data processing system to reduce potential nonresponse bias. Other methods for assessing bias, such as re-contact of a subsample of refusals may be considered in the future but will not be conducted as part of the work planned for the first production sample.

#### **4. Testing Procedures**

Various tests have been completed prior to the start of the ORS production samples. Field testing has focused on developing procedures, protocols, and collection aids. These testing phases were analyzed primarily using qualitative techniques but have shown that this survey is operationally feasible. Survey design testing was 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, current, and future 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. Respondents agreed to participate in the test. BLS field economists were able to capture the required data from traditional NCS respondents, and individual data element response rates were very high. 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 that had been updated following Phase 1 testing (e.g., streamlined collection tools; implementation of a probability-based establishment selection method; refined frequency questions; limited phone collection). This test was also developed to assess ORS collection outside the DC metropolitan area using an expanded number of BLS field economists. 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. Respondents agreed to participate in the test. BLS field economists were able to capture the required data from typical NCS respondents, and individual data element response rates were very high.

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. Phase 3 testing also included supplemental tests to assess the feasibility of Central Office Collection (COC), joint collection of ORS and Employment Cost Index (ECI) elements, and conducting "efficiency" interviews. 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). As in the prior two tests, respondents agreed to participate in the test. BLS field economists were able to capture the required data from traditional NCS respondents, and individual data element response rates were very high.

- Fiscal Year 2014

The BLS completed six feasibility tests to refine the ORS methodology. Five tests were conducted nationally, across all six BLS regions; and one test, the New Data Elements Test, was conducted only in two Metropolitan Areas: Washington, D.C. and San Diego, CA.

The six feasibility tests were designed to:

- refine the methods to develop more efficient approaches for data collection as identified during fiscal year 2013 testing (ORS Only Efficiency Innovations Test);
- determine how best to collect occupational requirements data elements and NCS data elements from the same establishment (NCS/ORS Joint Collection Test);
- determine how best to collect the new mental and cognitive demands for work data elements, and evaluate the use of occupational task lists as developed by ETA's O\*NET program during data collection (New Data Elements Test);
- determine how best to collect occupational requirements data elements from America's largest firms and State governments (Central Office Collection Test); and
- determine how best to collect occupational requirements data elements when a personal visit is not optimal due to respondent resistance, collection costs, or other factors (Alternative Modes Test).

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. All test objectives were successfully met and these activities established a strong foundation for the Pre-Production Test. 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 6).

#### Current

- Fiscal Year 2015

The goal of the Pre-Production test is to test all survey activities by mirroring production procedures, processes and protocols as closely as possible. Pre-Production data collection will run for approximately six consecutive months. ORS collection is occurring on live NCS schedules as well as a set of ORS-only schedules. The field economists are following the standard non-response follow-up protocols and making as much of an attempt to collect every assigned schedule as possible. Collection, data capture, and data review milestones have been set. All ORS data elements planned for Production are being collected from all schedules. For NCS establishments that have already been initiated, NCS data elements are being extracted from the NCS database and will not be collected again.

Every normal production activity associated with each of BLS' product lines are being conducted during Pre-Production testing. Production activities include selecting ORS samples, training staff, conducting calibration exercises, collecting the data, conducting all review activities, calculating estimates and standard errors, validating the estimates, and applying publication criteria to the computed estimates. All staff collecting data during the Pre-Production test are trained and are participating in calibration testing. Prior to test commencement, training was conducted for all collection staff. Staff have an ORS data capture system available for regional and national office use and ORS data review processes are being conducted. Data from this test that meets BLS publication criteria will be provided to SSA and released in a research report for the public. However, due to the sample size, the BLS only expects to be able to compute and release data for a very limited number of occupations or occupational groups, and these data will not be suitable for SSA disability determinations.



## Future

ORS production will begin by selecting samples using the methodology as described in this document. Current plans call for collection of the first schedules for the ORS production to start in September 2015 and to run for approximately nine consecutive months.

### **4b. Tests of Survey Design Processes**

#### Sample Design Options

To further ensure the BLS met the needs of the ORS by producing statistically valid data and of high quality, 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 11). This research continued into FY 2014 and expanded to look at other BLS surveys, including the Occupation Employment Statistics (OES) 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, which are documented in the ASA 2014 Papers and Proceedings titled “Occupational Requirement Survey, Sample Design Evaluation” by Ferguson et al (see Attachment 12) are:

1. Fully Integrated Survey Design – where the NCS establishment sample would be a subsample of the ORS establishment sample
2. Independent Survey Design – where the ORS establishment samples would be selected using a design appropriate for SSA’s needs, the NCS establishment samples would be selected using the current NCS sample design, and there would be no control on the amount of establishment sample overlap between the samples selected for the two surveys
3. Separated Survey Design – where the NCS establishment sample would be selected from the frame, the selected NCS establishments would be removed from the frame, and an independent ORS establishment sample would be selected from the rest of the frame
4. OES-ORS Integrated Design – where the ORS establishment sample would be selected as a subsample of the OES establishment sample

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. There are many things to consider when choosing a sample design for the ORS. Cost, individual respondent burden, overall respondent burden, response rates, data quality, the effect on the ECI, and whether the surveys could be integrated were all factors. After various testing on the four basic categories of ORS survey designs, the BLS determined that there are statistically viable designs under both the Fully Integrated and Independent Survey Design options. No integrated sample design that has been

studied fully meets the goals of both surveys (ORS and NCS) thus it was determined that the Independent Survey Design, out of the two statistically viable designs, was the best choice for implementation.

### 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 13).

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. The guide should include:
  - Background on the data needs, intended uses, and feasible collections.
  - A list of variables, levels of measurement, response options, and methods for calculating composite measures.
  - Variables and response options that are highest priority for testing based on the needs of SSA.
  - A list of occupations and data elements of highest priority for SSA.

- A clear statement of different data collection methods under considerations and rankings of cost feasibility.
  - A description of the format and content of data products.
- An evaluation on the existence of “gold standard” benchmarks for methods of data collection and for data elements. The evaluation should include:
  - Using field economists to observe occupations as the standard for physical demands and environmental conditions.
  - Using physical measuring devices for environmental conditions (such as noise)
  - Comparing alternative methods of data collection to determine their accuracy relative to the gold standard
- For data elements without any gold standards, multiple approaches may be used.
  - When ORS data elements have overlap with variables in existing microdata sets (e.g. education and training requirements), these databases should be used to measure agreement between ORS data and other data sources.
  - When there is little or no overlap between ORS data elements and existing databases (e.g. cognitive requirements), subject matter experts should be consulted to structure tests to determine validity. BLS should contract with an IO Psychologist to assist with this effort.
- 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 agrees with the recommendations provided by Dr. Handel. As a result, the BLS plans to begin a review initiative in FY 2015 including the development of a methodological guide, evaluation of “gold standard” benchmarks for data collection, and future testing of inter-rater reliability. 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 14). These recommendations, as well as the previous refinements of the ORS manual, the data review process, and the validation techniques developed to date will ensure ORS produces quality occupational data in the areas of vocational preparation, mental-cognitive and physical requirements, and environmental conditions as the BLS moves beyond testing and into full production.

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

Ms. Gwyn Ferguson, Chief, Statistical Methods Group of the Office of Compensation and Working Conditions, is responsible for the statistical aspects of the ORS production. Ms.

Ferguson can be reached on 202-691-6941. BLS seeks consultation with other outside experts on an as needed basis.

## 6. References

Gwyn R. Ferguson, Chester Ponikowski, and Joan Coleman, (October 2010), "Evaluating Sample Design Issues in the National Compensation Survey," ASA Papers and Proceedings, <http://www.bls.gov/osmr/pdf/st100220.pdf> (Attachment 1)

Bradley D. Rhein, Cheater H. Ponikowski, and Erin McNulty, (October 2014), "Estimation Considerations for the Occupational Requirements Survey," ASA Papers and Proceedings, [http://www.bls.gov/ncs/ors/estimation\\_considerations.pdf](http://www.bls.gov/ncs/ors/estimation_considerations.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, <http://www.bls.gov/ore/abstract/st/st060050.htm>, (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, <http://www.bls.gov/osmr/abstract/st/st080190.htm>, (Attachment 4)

Bureau of Labor Statistics' Handbook of Methods, Chapter 8, Bureau of Labor Statistics, 2010 <http://www.bls.gov/opub/hom/homch8.htm>, (Attachment 5)

The ORS Debrief Team, (November 2014) "Occupational Requirements Survey, Consolidated Feasibility Tests Summary Report, Fiscal Year 2014," Bureau of Labor Statistics, [http://ocwc.sp.bls.gov/ncsmanagers/392coord/fy14final/ORS\\_FY14\\_Feasibility\\_Test\\_Summary\\_Report.pdf](http://ocwc.sp.bls.gov/ncsmanagers/392coord/fy14final/ORS_FY14_Feasibility_Test_Summary_Report.pdf), (Attachment 6)

Gwyn R. Ferguson, (January 2014), "Testing the Collection of Occupational Requirements Data," ASA Papers and Proceedings, <http://www.bls.gov/osmr/abstract/st/st130220.htm>, (Attachment 7)

The ORS Debrief Team, (January 2013) "Occupational Requirements Survey, Phase 1 Summary Report, Fiscal Year 2013," Bureau of Labor Statistics, [http://www.bls.gov/ncs/ors/phase1\\_report.pdf](http://www.bls.gov/ncs/ors/phase1_report.pdf), (Attachment 8)

The ORS Debrief Team, (April 2013) "Occupational Requirements Survey, Phase 2 Summary Report, Fiscal Year 2013," Bureau of Labor Statistics, [http://www.bls.gov/ncs/ors/phase2\\_report.pdf](http://www.bls.gov/ncs/ors/phase2_report.pdf), (Attachment 9)

The ORS Debrief Team, (August 2013) "Occupational Requirements Survey, Phase 3 Summary Report, Fiscal Year 2013," Bureau of Labor Statistics, [http://www.bls.gov/ncs/ors/phase3\\_report.pdf](http://www.bls.gov/ncs/ors/phase3_report.pdf), (Attachment 10)

Bradley D. Rhein, Cheater H. Ponikowski, and Erin McNulty, (November 2013), "Sample Design Considerations for the Occupational Requirements Survey," FCSM Papers and Proceedings, [https://fcsm.sites.usa.gov/files/2014/05/H4\\_Rhein\\_2013FCSM.pdf](https://fcsm.sites.usa.gov/files/2014/05/H4_Rhein_2013FCSM.pdf), (Attachment 11)

Gwyn R. Ferguson, Erin McNulty, and Cheater H. Ponikowski (October 2014), "Occupational Requirements Survey, Sample Design Evaluation," ASA Papers and Proceedings [http://www.bls.gov/ncs/ors/sample\\_design.pdf](http://www.bls.gov/ncs/ors/sample_design.pdf) (Attachment 12)

Kristin N. Smyth, (October 2014), "Validation in the Occupational Requirements Survey," ASA Papers and Proceedings <http://www.bls.gov/ncs/ors/validation.pdf>, (Attachment 13)

Michael Handel, (February 2015), "Methodological Issues Related to ORS Data Collection," [www.bls.gov/ncs/ors/handel\\_exec\\_summ\\_feb15.pdf](http://www.bls.gov/ncs/ors/handel_exec_summ_feb15.pdf), (Attachment 14)