Occupational Requirements Survey Pre-Production Report
September 28, 2015

Executive Summary

In fiscal year 2015, the Bureau of Labor Statistics (BLS) completed data collection for the Occupational Requirements Survey (ORS) pre-production test. The pre-production test might better be described as a "dress rehearsal" as the collection procedures, data capture systems, and review were structured to be as close as possible to those that will be used in production.[1] The feasibility tests in FY 2014 and earlier were intended to gauge the viability of collecting occupational data elements and to test modes of collection and procedures and BLS integrated the results of this prior work into the large-scale nationally representative pre-production test.

This report is a compilation of two individual reports, the "Occupational Requirements Survey Pre-Production Collection Report" and "Occupational Requirements Survey Pre-Production Estimation and Validation Report" posted to the BLS website in June and September, 2015, respectively.[2] Because these reports were issued so close to the end of the pre-production test, some numbers may have changed.

Preliminary indications show that BLS is able to successfully collect data on occupational requirements that will that will meet the needs of the Social Security Administration (SSA) as inputs into their Occupational Information System (OIS). Overall and item-level response rates, evidence from the quality assurance program, and feedback from data collection debriefs provide support that survey procedures are working well and data on job requirements can be collected by BLS field economists.

Initial work on estimation and validation of the data indicate that BLS is able to successfully produce estimates of occupational requirements. Though the pre-production sample was relatively small, over 11,000 estimates were calculated and validated. These estimates span 21 2-digit SOCs and 48 8-digit SOCs and provide information about the requirements of civilian jobs. We continue to refine collection procedures, definitions, and training to ensure that collection of the ORS data elements will lead to high quality estimates as we transition into full-scale production of ORS.

As BLS moves into production, however, there will be continued refinements and testing, particularly for the mental and cognitive elements, as described in this report.

Background and Pre-production Test Overview

In the summer of 2012, the SSA and BLS signed an interagency agreement, which has been updated annually, to begin the process of testing the collection of data on occupational requirements. As a result, BLS established ORS as a test survey in late 2012. The goal of ORS is to collect and publish occupational information that meets the needs of SSA at the level of the eight-digit standard occupational classification (SOC) that is used by the Occupational Information Network (O*NET).[3]

---

[1] The sample design was similar that which will be used in production, but altered to meet test goals.

[2] Please see www.bls.gov/ncs/ors/ for copies of these earlier reports.

[3] The occupational classification system most typically used by BLS is the six-digit SOC (www.bls.gov/soc/), generally referred to as "detailed occupations". O*NET uses a more detailed occupational taxonomy (www.onetcenter.org/taxonomy.html), classifying occupations at eight-digits and referring to these as "O*NET-SOC 2010 occupations". There are 840 six-digit SOCs and 1,110 eight-digit SOCs.

The ORS data are collected under the umbrella of the National Compensation Survey (NCS), which uses Field Economists (FEs) to collect data. FEs generally collect data elements through either a personal visit to the establishment or remotely via telephone, email, mail, or a combination of modes.

For ORS, FEs are collecting occupationally-specific data elements to meet SSA's needs in the following categories:
- Physical demands
- Specific vocational preparation
- Mental and cognitive demands
- Environmental conditions in which the work is performed.

There were roughly 70 ORS-specific data elements in all, with the majority of these falling in the category of physical demands.

In fiscal years 2013 and 2014, several feasibility tests were performed to assess the viability of collecting data on occupational requirements using the platform currently used by the NCS. These tests provided evidence that the NCS platform could be adapted to ORS data collection, which led to the pre-production test in FY 2015.[4] Unlike the earlier tests, which were small-scale and tested a subset of data elements or the viability of different collection methods, the pre-production test was designed as a relatively large-scale, nationally representative test of ORS data collection. ORS pre-production data collection began in October 2014 and continued until May 2015. The sampling, data collection, procedures, and review were designed to mimic what will occur during ORS production.

This report summarizes the pre-production test from sampling through validation, describes additional testing, and discusses next steps for ORS data collection as BLS moves into production.

Sampling

The ORS pre-production sample was drawn from the same frame as the NCS – the Quarterly Census of Employment and Wages, which includes all establishments covered by State unemployment insurance laws, and a supplementary file of railroads. The frame contains virtually all establishments in the 50 United States and the District of Columbia in the private sector (excluding agriculture, forestry and fishing and private households) and in state and local governments. The pre-production ORS sample contained 2,549 establishments.

Roughly one-third of the ORS pre-production sample consisted of establishments that are also in the NCS sample ("NCS-ORS overlap") and the remainder are ORS-only.[5] Across all establishments (NCS-ORS overlap and ORS-only), approximately 15 percent are government owned and 85percent privately owned.

---

[4] Copies of earlier test reports can be found at: www.bls.gov/ncs/ors/home.htm, in the text box titled, "Fiscal Year 2014 Occupational Requirements Testing Results."

[5] Additional information on the survey design can be found as part of the Federal Register Notice, www.reginfo.gov/public/do/PRAViewDocument?ref_nbr=201403-1220-002.

For each establishment in the ORS sample (establishments are also referred to as "schedules"), jobs are selected for data collection through probability selection of occupations; these jobs are referred to as "quotes".[6] The number of jobs selected within a private establishment typically varies from 4 to 8, based on establishment size,[7] and in government, the number of jobs varies from 4 to 20. It is common for multiple individuals within an establishment to have the same job (e.g. elementary school teachers within a school/school district), which can result in fewer individual quotes for that establishment. Because the quote-level information is tied to the job, not the individual, sampling a certain number of jobs within an establishment is not equivalent to sampling a certain number of workers within an establishment. Sample weights will be assigned to each quote and establishment to represent the entire frame.

Data collection

At the close of the data review process, there were 7,109 quotes collected from 1,851 establishments, slightly less than four jobs per establishment. These jobs spanned 22 2-digit SOCs and 704 unique 8-digit SOCs. Though there are 23 2-digit SOCs, one of these, military-specific occupations (55-0000), is out of scope for ORS. The 704 8-digit SOCs represent 63 percent of the 1,090 unique 8-digit SOCs in scope for ORS.[8]

Of the 2,549 establishments contacted by field economists, 168 were either out of business, out of scope, or had no jobs in scope for ORS. Of the remaining 2,381 establishments, 1,851 of them provided usable data, indicating a usable establishment response rate of 78 percent.[9]

The quote-level response rate was 85 percent, with 15 percent refusals. Response rates on individual elements varies. Among physical demand elements, the response rate was 84 percent; the response rate was 85 percent for mental and cognitive demands, 86 percent for specific vocational preparation, and 92 percent for environmental conditions. Within these aggregate groupings, response rates for individual elements varied substantially.

Interview duration was, on average, 63 minutes. Roughly half (49 percent) of schedules were conducted by personal visit and the balance by some other mode (telephone, email, mail, or a combination of approaches).

Duration tended to be longer, on average, for personal visit appointments (76 minutes) than for other modes (50 minutes), which is expected since personal visit appointments were more likely to occur for establishments with a higher number of sampled jobs.

The establishment respondents in the ORS pre-production test included human resource professionals/workforce planners, workers compensation staff, safety/occupational health/risk

---

[6] Probability selection of occupations (PSO) is also used to select jobs in the NCS. Additional information on probability selection of occupations can be found in the BLS Handbook of Methods, www.bls.gov/opub/hom/ncs/home.htm

[7] There are two exceptions. 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.

[8] Eight-digit SOCs within military-specific occupations are out of scope for the survey.

[9] Response rate is calculated as the ratio of the number of usable establishments divided by the sum of usable and refusals.

management staff, and subject matter experts (supervisors, incumbents, department heads). These respondents were similar to those who provided data during previous feasibility tests. The feedback from FEs in pre-production is that finding the right respondent is key to data collection and that respondents with knowledge of job descriptions tended to be most helpful.

The HR departments were often the starting point for finding respondents, for two reasons. First, for establishments that are currently part of the NCS sample, the FEs already have established contacts who are often in HR. Second, in order to perform the job sampling within an establishment, FEs first need information that is typically maintained in the HR department. The FEs reported that often the primary respondent would consult with others on staff for responses to data elements.

Procedures

There are a number of resources provided for FEs to ensure consistent procedures, collection, and coding of data elements. FEs are provided with a collection manual (and training on the contents of the manual) which covers such topics as securing cooperation from respondents, fundamentals/definitions of ORS pre-production data elements (along with practical examples), guidance on how elements might be related, and instructions on coding in the data capture system.

During collection, FEs may submit inquiries to resolve ambiguities that arise. Often these are one-off topics, such as whether walkie-talkies would fall under the category of phone communication. The procedures staff provides a response to the inquiry, which is then distributed to all FEs. There were 23 such responses provided during pre-production.

If there is evidence that additional clarification is needed on data elements (i.e. there exists a more general source of confusion), then the procedures staff highlights the correct procedure through a "procedures alert," that is distributed to all FEs. An example of this during pre-production was an alert that provided additional information on "deviations from the norm" for mental and cognitive elements. During pre-production there were five procedures alerts issued.

Debriefs with the FEs were also used during the ORS pre-production test. These were an opportunity to share best practices, get feedback on respondent comprehension of data elements, and for FEs to highlight any problems they were having with collection. Feedback from these debriefs will inform FE training and the materials provided in the collection manual as preparations are made for ORS data collection for FY 2016.

ORS Review Program

The ORS pre-production data were subject to a review program modeled on the NCS that will also be used in production. The data review program has components that occur in the field as well as components that occur after the data are submitted to BLS headquarters.

The quality assurance process starts with the classroom and on-the-job training of newly hired Field Economists, which takes place over 9-12 months. The classroom training includes survey procedures, collection protocols, data capture and review systems, and interviewing techniques. Once basic training is completed, the certification process begins, which pairs a new FE with a certified FE who will observe data collection and review collected data.

Additional review of collected data includes: full schedule review, where all data elements for an establishment (also referred to as a "schedule") are reviewed by experienced field staff; technical re-interview, where respondents are re-contacted and a subset of data elements are re-collected; and targeted schedule review, where a subset of elements are reviewed by staff at BLS headquarters. Roughly 15 percent of schedules are subject to full schedule review, 5percent to technical re-interview, and 20percent to targeted review.

The remaining 60 percent of schedules are subject only to secondary review, in which data elements flagged outside of the primary data capture system are reviewed. This secondary review focuses on verifying data elements that are unexpected or inconsistent with other data elements (e.g. physical elements that are unexpected given the occupation). This also encompasses review across schedules to identify schedules and elements with anomalies.

The review staff work with the FEs to ensure that all data are correctly coded and that any flagged elements have documentation verifying that the elements are correctly coded. Reviewers questioned approximately one percent of the data during ORS pre-production. When a field economist receives a question from a reviewer, the FE reviews the values coded in the data capture system, reviews their notes from the interview with the respondent, may review the procedures and collection manual, and may call back the respondent to clarify any information from the interview. Roughly 61 percent of the time these questions resulted in data being changed.

Estimation and Validation

While there were roughly 70 ORS-specific elements collected in pre-production, there are many more estimates that correspond to these elements.[10] For categorical elements, the percentage of workers associated with each category are estimated. For example, the responses for encountering noise are quiet, moderate, loud, and very loud and percentage of workers will be calculated for each of these. For continuous variables, mean, mode and percentiles (10th, 25th, 50th, 75th, and 90th) are being estimated. An example of this would be time spent standing (measured as a duration in hours). Additionally, some ORS estimates are calculated from multiple ORS elements. Specific vocational preparation (SVP; the amount of time required to learn the techniques, acquire the information, and develop the facility needed for average job performance), for example, is calculated based on education, experience, certification/licensing, and post-employment training.

If we wished to produce only one set of estimates for the roughly 70 ORS-specific elements, for example, national estimates for civilian workers, then 689 estimates would be produced.[11] However, the goal is to produce not just "top line" estimates, but estimates on more detailed subgroups, including by 2-digit and 8-digit O*NET-SOC. These series of estimates are:

1. Civilian, all workers
2. Civilian, 2-digit SOC groups (22)
3. Civilian, 8-digit SOC (1,090)

---

[10] A table of ORS elements is presented in Appendix A at the end of this document.
[11] Civilian workers includes private industry and state and local government workers.

Producing estimates at this level of detail results in hundreds of thousands of potential estimates, as reflected in the table below.

Table 1: Potential estimates

Series	Potential Series	Potential Estimates
All workers	1	689
2-digit SOC	22	15,158
8-digit SOC	1,090	751,010

In order to produce any estimates, however, there are a number of steps that BLS uses. Sample weights are assigned to each establishment and occupation and then adjusted for non-response. The weights are then benchmarked to account for current employment. Then estimates are produced and reviewed for reliability and confidentiality.

The failure to meet BLS criteria for reliability and confidentiality often is attributable to too few observations available for a particular estimate, which may be due to too few quotes in a particular SOC group, how the data are clustered within a SOC, or to item-level non-response to particular data elements. For example, although there are five categories of responses for encountering noise, it may be the case that workers in a given occupation are never exposed to noise at all five levels.
The loss of estimates due to item non-response (where we have some, but not all, elements for a quote), can commonly be addressed through imputation. The estimates for ORS pre-production are not based on imputation, as we have not finalized an imputation approach. Research in this area is on-going.

Once estimates are produced, they go through a validation process, by which they are deemed "fit for use." There are multiple approaches for validating ORS estimates. Visualization tools play a large part in the validation process. First, estimates are reviewed to see whether they conform to expectations. In other BLS programs, such as the National Compensation Survey, expectations may be formed based on past values of the estimates; however, as ORS is a new program, this is not an option for pre-production validation. For pre-production, staff involved in validation examined estimates, using visualization tools and other reports, for outliers and apparent inconsistencies. For example, one should find that landscapers or dishwashers have higher estimates for exposure to wetness than occupations that more typically work indoors and away from sinks. SVP should be relatively higher for occupations typically classified as "white collar" versus "blue collar."

Other approaches to validation involve comparing ORS pre-production estimates to other data sources with similar elements. For example, ORS estimates of SVP can be compared to SVP estimates from O*NET and the Dictionary of Occupational Titles (DOT). Also, BLS collects information on physical environment for the NCS which can be compared to the ORS elements for physical demands.

When the validation team identifies potential inconsistencies or weaknesses in the estimates, a decision may be made to suppress a subset of estimates. In the case of ORS, estimates suppression occurred most frequently due to high relative frequency of "unknown" as a response to data elements. Additionally, if a suppression is made in a set of estimates that are part of an additive group (a set of estimates that sums to 100 percent) then a secondary suppression is necessary as well. For example, task complexity has five categories: very complex, complex, moderate, simple, and very simple. If an

estimate was suppressed for "very simple" then an estimate for another category would also be suppressed.

Table 2 builds on Table 1 to give a sense of which estimates have passed these criteria from the pre-production test.

Table 2: Potential estimates and pre-production estimates

Series	Potential Series	Potential Estimates	Series with no collected data in pre-production	Series available for calculation	Number of estimates passing all criteria
All workers	1	689	0	1	504
2-digit SOC	22	15,158	0	21	5,081
8-digit SOC	1,090	751,010	386	48	5,714

Over three-quarters of a million estimates would be possible with full coverage of 8-digit SOCs. The pre-production test, with 7,109 unique quotes collected from 1,851 establishments, had a relatively small sample, so it is not surprising that the actual number of estimates passing BLS's criteria is considerably fewer than the potential number of estimates.

Turning first to the 2-digit SOC series, column 4 of Table 2 shows that there were no series without collected data in pre-production. This means at least one quote was obtained in each 22 2-digit SOC group. However, column 5, "series available for calculation," shows that only 21 of the 22 SOCs were used to generate estimates. While data were obtained for all 2-digit SOCs, there were not enough quotes obtained in one of the SOCs (farming, fishing, and forestry occupations) to produce even one estimate for this group. In theory, 15,158 estimates could have been produced for these 21 2-digit SOC groups, in actuality, 5,081 estimates passed all BLS criteria.

There were 386 8-digit SOCs with no inputs from pre-production. Among the 704 SOCs for which there was at least one quote collected during pre-production, only 48 met the initial BLS estimate criteria. These 48, however, include occupations that we anticipate will be of high interest to potential ORS users, including:

- Heavy and tractor trailer truck drivers
- Food prep workers
- Janitors and cleaners
- Landscapers
- Laborers and freight, stock, and material movers
- Nursing assistants
- Waiters and waitresses
- Maids
- Cashiers

As with the 2-digit SOCs, it is important to note that we do not have a full set of 689 estimates for each of the 48 8-digit SOCs. If we had 689 estimates for each of the 48, there would be 33,072 estimates. Table 2 indicates that pre-production data collection generated 5,714 estimates.

Dissemination Plans

How will the ORS preproduction estimates be shared with the public? The current plan is to produce an article that highlights a subset of the preproduction estimates, which will be posted on the BLS ORS website before the end of the 2015 calendar year. Once ORS estimates are produced as part of official production, we envision disseminating results through channels that showcase the relationships between data elements. We anticipate that stakeholders will have multiple uses for the estimates and desire to see them presented in different ways. For example, one may want to understand the physical requirements associated with specific occupations. Alternatively, one may be interested in the set of occupations associated with a certain level of SVP. NCS hopes to develop dissemination tools that allow user interaction through dashboards and other technology. The overall goal is to display data visually in addition to access through the public database tools. These represent longer-term plans for dissemination.

Changes in Collection Anticipated for Production

Implementing a large-scale test of the ORS data elements has resulted in changes and refinements to several of the elements as the survey moves into production. These changes were made in consultation with the Social Security Administration and reflect the desire to collect data that meet the needs of SSA. What is most notable is that relatively few data elements are changing; the pre-production test has demonstrated that the definitions and wording of the majority of the data elements are working well in data collection.

There was only one change to the physical requirements elements. This involved rewording the question on climbing ramps or stairs to ensure that the climbing was work-related and not simply the result of someone, say, choosing to walk up stairs to a third floor office.

Among the environmental elements, there were two changes. The first consolidated the presence of hazardous contaminants into one data element rather than two. The second added an indicator for the presence of personal protective equipment for hazardous contaminants, moving mechanical parts, high exposed places, and noise intensity level.

All of the mental and cognitive elements are being revised for production. Highlights are provided below:
- The question on task complexity has been removed. A question has been added, "what type of decision-making is required to perform the tasks of the occupation?"
- The question on work controls has been removed. A question has been added, "what type of supervision does this occupation have?"
- A question has been added on pace of the work performed.
- A question has been added regarding what controls the pace of work (the work or the worker).
- The question on work deviating from the norm has been removed. A question has been added, "how often do (work/schedule/location) change in this occupation?"
- The questions on contacts have been revised to rescale the options for frequency of contacts.

Additional Testing

Two additional tests of data collection were conducted in June-August 2015. The first was the ORS job observation pre-test. The job observation pilot test was intended to assess whether the data collected through ORS interview collection methods are systematically different than data collected through direct observation. This pilot test involved re-contact of a subset of establishments that were interviewed as part of the pre-production test. FEs observed select jobs within the establishment and record data on physical and environmental data elements. This pilot test was in response to both Federal Register Notice public comments and an external subject matter expert's recommendations for testing and validation of ORS survey design.[12] A written report of this test will be available in fall 2015.

The second test was a test of the revised mental and cognitive elements, to understand respondents' perception of the wording of these questions. Data was collected from establishments who participated in ORS feasibility tests. If necessary, the results from this test will be used to further refine the mental and cognitive elements after the first year of ORS data is collected.

Lessons Learned and Overall Assessment

At this point, the results from the ORS pre-production test demonstrate that data on occupational requirements can be collected using the processes established by BLS, but final conclusions about the feasibility of these methods will be made when the analysis of pre-production data is complete. Training and procedures are designed for consistent data collection by field economists. The feedback channels between FEs and the procedures team ensure that questions are answered quickly, information is disseminated to all ORS staff, and any needed modifications to procedures are made and communicated effectively. The data review process will evolve as new procedures are implemented and the indicators used to identify inconsistencies in collected schedules are refined.

The response rate for ORS is comparable to that for the NCS and item level response rates provide preliminary evidence that usable data on job requirements are being collected. The elements that address the physical, environmental, and vocational preparation requirements (tested through feasibility tests in prior years) appear to be working well in pre-production collection and few changes are being made to those elements as BLS moves into production.

There are some aspects to collection that need to be addressed going forward. Training and procedures for ORS production should address best practices for identifying the proper respondent as well as strategies for dealing with establishments for which FEs are collecting both ORS and NCS data elements. Item level response rates indicate that some respondents are having difficulty providing duration data for some job requirements and tools and procedures to assist field economists in obtaining measures for those data elements are needed.

The mental and cognitive elements have largely been rewritten for production, which will require additional training and procedures that assist FEs in collecting these new elements from respondents. The test of mental and cognitive elements scheduled for summer 2015 will help shape this. Minor refinements may be made during the first year of production collection to mental and cognitive

---

[12] A link to the subject matter expert's report, "Occupational Requirements Reliability and Validity Literature Research," can be found at www.bls.gov/ncs/ors/handel_report_feb15.pdf.

elements as part of the OMB clearance process for production collection. However, major changes cannot be made until the second year of collection. If the summer test and initial collection indicate further major changes to the mental and cognitive elements are needed, tests of the mental and cognitive elements will occur in FY2016.

Appendix A: List of ORS Elements

Specific Vocational Preparation -- 4 elements	Physical Demands - Exertion -- 14 elements
Minimum Formal Education or Literacy required	Most weight lifted/Carried ever
Pre-employment Training (license, certification, other)	Push/Pull with Feet Only: One or Both
Prior Work Experience	Push/Pull with Foot/Leg: One or Both
Post-employment training	Push/Pull with Hand/Arm: One or Both
	Pushing/Pulling with Feet Only
Mental and Cognitive Demands Elements -- 9 elements	Pushing/Pulling with Foot/Leg
Closeness of Job Control level	Pushing/Pulling with Hand/Arm
Complexity of Task level	Sitting
Frequency of Deviations from Normal Work Location	Sitting vs Standing at Will
Frequency of Deviations from Normal Work Schedule	Standing and Walking
Frequency of Deviations from Normal Work Tasks	Weight Lifted/Carried 2/3 of the time or more (range)
Frequency of verbal work related interaction with Other Contacts	Weight Lifted/Carried 1/3 up to 2/3 of the time (range)
Frequency of verbal work related interaction with Regular Contacts	Weight Lifted/Carried from 2percent up to 1/3 of the time (range)
Type of work related interactions with Other Contacts	Weight Lifted/Carried up to 2percent of the time (range)
Type of work related interactions with Regular Contacts	
	Physical Demands - Reaching/Manipulation - 14 elements
Auditory/Vision -- 10 elements	Overhead Reaching
Driving, Type of vehicle	Overhead Reaching: One or Both
Communicating Verbally	At/Below Shoulder Reaching
Hearing: One on one	At/Below Shoulder Reaching: One or Both
Hearing: Group	Fine Manipulation
Hearing: Telephone	Fine Manipulation: One Hand or Both
Hearing: Other Sounds	Gross Manipulation
Passage of Hearing Test	Gross Manipulation: One Hand or Both
Far Visual Acuity	Foot/Leg Controls
Near Visual Acuity	Foot/Leg Controls: One or Both
Peripheral Vision	Keyboarding: 10-key
	Keyboarding: Other
Environmental Conditions -- 11 elements	Keyboarding: Touch Screen
Extreme Cold	Keyboarding: Traditional
Extreme Heat	
Fumes, Noxious Odors, Dusts, Gases	**Physical Demands - Postural -- 7 elements**
Heavy Vibration	Climbing Ladders/Ropes/Scaffolds
High, Exposed Places	Climbing Ramps/Stairs: structural only
Humidity	Climbing Ramps/Stairs: work-related
Noise Intensity Level	Crawling
Outdoors	Crouching
Proximity to Moving Mechanical Parts	Kneeling
Toxic, Caustic Chemicals	Stooping
Wetness	