

Validation in the Occupational Requirements Survey: Analysis of Approaches

Kristin N. Smyth¹

¹U.S. Bureau of Labor Statistics, 2 Massachusetts Ave., NE, Room 4160,
Washington, DC 20212

Abstract

The Bureau of Labor Statistics (BLS) is working with the Social Security Administration (SSA) to carry out a series of tests to determine the feasibility of using the National Compensation Survey (NCS) platform to accurately and reliably capture data that are relevant to the SSA's disability program. The proposed new Occupational Requirements Survey (ORS) is envisioned to be an establishment survey that provides information on the vocational preparation and the cognitive and physical requirements of occupations in the U.S. economy as well as the environmental conditions in which the work is performed. This paper presents various methods for reviewing and validating the reasonableness of survey estimates. Interactive graphics in visualization software packages have been developed as new tools to complement traditional validation processes. These methods are being designed to properly gauge the reasonableness of the estimates as there is little historical precedent to which these data can be compared.

Key Words: data editing, visualization, data review, data graphics

1. Introduction

The Social Security Administration (SSA) approached the Bureau of Labor Statistics (BLS), specifically the National Compensation Survey (NCS), because NCS collects data on work characteristics of occupations in the U.S. economy. SSA is interested in occupational information for use in their disability programs. The SSA and the BLS have entered into annual agreements for collecting new data on occupational information for use in the disability programs. The goal of the Occupational Requirements Survey (ORS) is to collect and eventually publish information that will replace the outdated occupational data currently used by SSA. All ORS products will be made public for use by non-profits, employment agencies, state or federal agencies, the disability community, and other stakeholders.

An ORS interviewer attempts to collect close to 70 data elements related to the occupational requirements of a job. The following four groups of information will be collected:

- Physical demand characteristics/factors of occupations (e.g. crawling, hearing, or stooping)
- Educational Requirements
- Cognitive elements required to perform work

- Environmental conditions in which the work is completed

This paper explores the validation of the estimates for ORS data. Section 2 provides background information on the Occupational Requirements Survey. Section 3 explains the typical validation process and the challenges posed by the ORS data elements. Section 4 explores possible validation strategies. The paper ends with a conclusion and description of further research to be completed.

2. Background Information on ORS

In addition to providing Social Security benefits to retirees and survivors, the Social Security Administration administers two large disability programs that provide benefit payments to millions of beneficiaries each year. Determinations for adult disability applicants are based on a five step process that evaluates the capabilities of the worker, the requirements of their past work, and their ability to perform other work in the U.S. economy. In some cases, if an applicant is denied disability benefits, SSA policy requires adjudicators to document the decision by citing examples of jobs the claimant can still perform despite their limitations (such as limited ability to balance, stand, or carry objects) [1].

For over 50 years, the Social Security Administration has turned to the Department of Labor's Dictionary of Occupational Titles (DOT) [2] as its primary source of occupational information to process the disability claims [3]. SSA has incorporated many DOT conventions into their disability regulations. However, the DOT was last updated in its entirety in the 1970's, although a partial update was completed in 1991. Consequently, the SSA adjudicators who make the disability decisions must continue to refer to an increasingly outdated resource because it remains the most compatible with their statutory mandate and is the best source of available data at this time.

When an applicant is denied SSA benefits, SSA must sometimes document the decision by citing examples of jobs that the claimant can still perform, despite their functional limitations. However, since the DOT has not been updated for so long, there are some jobs in the American economy that are not even represented in the DOT, and other jobs, in fact many often-cited jobs, no longer exist in large numbers in the American economy. For example, a job that is often cited is "envelope addressor," because it is an example of a low-skilled job from the DOT with very low physical demands. There are serious doubts about whether or not this job still exists in the economy.

SSA has investigated the numerous alternative data sources for the DOT such as adapting the Employment and Training Administration's Occupational Information Network (O*NET) [4], using the BLS Occupational Employment Statistics program (OES) [5], and developing their own survey. But SSA was not successful with any of these potential data sources and turned to the National Compensation Survey (NCS) at the Bureau of Labor Statistics (BLS).

NCS is a national survey of business establishments conducted by the BLS [6]. Initial data from each sampled establishment are collected during a one year sample initiation period. Many collected data elements are then updated each quarter while other data elements are updated annually for at least three years. The data from the NCS are used to produce the

Employer Cost Index (ECI), Employer Costs for Employee Compensation (ECEC), and various estimates about employer provided benefits. Additionally, data from the NCS are combined with data from the OES to produce statistics that are used to help the President's Pay Agent and the Federal Salary Council recommend changes in how certain Federal employees are paid.

In order to produce these measures, the NCS collects information about the sampled business or governmental operation and about the occupations that are selected for detailed study. Each sample unit is classified using the North American Industry Classification System (NAICS) [7]. Each job selected for study is classified using the Standard Occupational Classification system (SOC) [8]. In addition, each job is classified by work level – from entry level to expert, nonsupervisory employee to manager, etc. [9]. These distinctions are made by collecting information on the knowledge required to do the job, the job controls provided, the complexity of the tasks, the contacts made by the workers, and the physical environment where the work is performed. Many of these data elements are very similar to the types of data needed by SSA for the disability determination process.

All NCS data collection is performed by professional economists or statisticians, generically called field economists. Each field economist must have a college diploma and is required to complete a rigorous training and certification program before collecting data independently. As part of this training program, each field economist must complete several calibration exercises to ensure that collected data are coded the same way no matter which field economist collects the data. NCS uses processes like the field economist training to help ensure that the data collected in all sectors of the economy in all parts of the country are coded uniformly.

SSA asked the NCS to partner with them under an annual interagency reimbursable agreement to test the NCS ability to use the NCS platform to collect data on four groups of information related to employer requirements for an occupation:

- Physical demand (PD) characteristics/factors of occupations (e.g. crawling, hearing, or stooping)
- Educational Requirements
- Cognitive elements required to perform work
- Environmental conditions in which the work is completed

If BLS is able to collect these data about work demands, SSA would have new and better data to use in its disability programs. SSA cited three key advantages of using NCS to provide this updated data:

- Reputation - SSA was impressed with the BLS reputation for producing high quality, statistically accurate data that are trusted by our data users and follow statistically accepted methods and principles.
- Trained Workforce – SSA was also impressed that NCS Field Economists have experience collecting information about occupations in America's work force and collecting data similar to that needed by SSA.
- Survey Infrastructure - After attempting to develop their own survey, SSA was also impressed with the fact that NCS has infrastructure in place across the country to manage and implement a new survey to meet their data needs as well as systems and processes to support all the steps of the survey.

Since 2012, NCS has been testing our ability to collect these new data elements using the NCS survey platform. Field testing to date has focused on developing procedures, protocols, and collection aids using the NCS platform. These testing phases were analyzed primarily using qualitative techniques but have shown that this survey is operationally feasible. Now it is time to turn our attention to ensuring the high quality of the estimates produced. This paper presents the validation process that is being developed for the estimates in the ORS program.

3. Validation Processes

ORS captures occupational information on educational requirements, cognitive demands, physical demands, and exposures to environmental conditions. Each of the data elements falls into two data types: categorical and continuous. Data elements that are categorical have a set of predetermined values, one of which will be selected as a response. Continuous data may be limited by a minimum, such as zero hours, or a maximum, such as 100 percent.

The full list of data elements is shown below:

Physical Demands

Hearing
 One on one hearing
 Group hearing
 Telephone
 Other Sounds
 Noise Intensity
 Keyboarding (10-key)
 Keyboarding (Other)
 Keyboarding (Touch Screen)
 Keyboarding (Traditional)
 Far Visual Acuity
 Near Visual Acuity
 Peripheral Vision
 Sitting
 Sitting vs Standing at Will
 Standing/Walking
 Stooping
 Kneeling
 Crawling
 Crouching
 Pushing/Pulling with Feet Only
 Pushing/Pulling with Foot/Leg
 Pushing/Pulling with Hand/Arm
 Reaching at/below Shoulder
 Reaching Overhead
 Strength

Educational Requirements

Certifications, license, and training
 Degree
 Literacy
 Post-employment training
 Previous Experience
 SVP
 Time to Average Performance
 Skill Level
 Job Zone

Environmental Condition

Extreme Cold
 Extreme Heat
 Fumes, Noxious Odors, Dusts, Gases
 Heavy Vibration
 High, Exposed Places
 Humidity
 Outdoors
 Proximity to Moving Parts
 Toxic, Caustic Chemicals
 Wetness

Cognitive Elements

Frequency of verbal interaction with regular contacts
 Frequency of verbal interaction with other contacts
 Measure of Job Control

Most Weight Lifted	Frequency of Job Deviations in Tasks
Amount of Weight Lifted Constantly	Frequency of Job Deviations in Work Schedule
Amount of Weight Lifted Frequently	Frequency of Job Deviations in Work Location
Amount of Weight Lifted Occasionally	Job Complexity
Amount of Weight Lifted Seldomly	Type of Contact with Regular Contacts
Climbing Ladders/Ropes/Scaffolds	Type of Contact with Other Contacts
Climbing Ramps/Stairs, Job-related	
Climbing Ramps/Stairs, structural	
Fine Manipulation	
Gross Manipulation	
Using One Hand or Both	
Driving	
Driving Vehicle Type	

Once the field economists collect and code the data, they undergo a series of quality reviews. An ORS review program is in development to create the processes, procedures, tools, and systems that will be used to check the micro-data as they come in from the field. The ORS Review Program is a separate process that focuses on the micro-data and includes the goal of micro-data review. 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 has been 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.

The process for validating the estimates constitutes a separate but related set of activities. Estimate validation focuses on aggregated tabulations of weighted data as opposed to individual data. 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.

Since validation and data review have interrelated goals, lessons learned from one may influence the other in ORS development, and some tools may be of use in accomplishing both goals. This paper aims to explain the validation side of these activities, but as it necessarily relates to micro-data review, certain aspects of review are referred to as well.

The validation tools in development are intended to bring attention to the estimates that are unusual or unreliable and assist in determining what went into the construction of those estimates. Estimates or tabulations that do not conform to expectations are considered anomalous. Expectations cover internal expectations and those found from outside sources (if any). Part of developing the validation process is searching for and identifying patterns

in the micro-data and recognizing relationships for use in building future systems both in validation and also in data review, like systematic review edits, reviewing parameters, training reviewers, etc. These patterns in the micro-data help to not only form our expectations for the estimates but also to provide information that may help in the development of procedures, data collection, and review parameters.

Validation in the NCS platform is done by comparing the current estimate to a set of expectations, primarily its historical counterpart. If the current estimate has changed by a factor that is greater than the expected threshold, the underlying data for that estimate are checked. NCS also assesses the effects of estimation, sampling, weighting, and methodological or procedural changes to see if these factors have had an influence on the tabulations or estimates.

In ORS, however, there exists no perfect historical match for the information being collected. Previous occupational research studies are out of date or in a form that make comparisons unwieldy. With few third party or historical sources to draw from for validation, there exists a new challenge within ORS to identify unexpected tabulations. NCS also makes use of unstructured review of estimates by senior economists based on salient changes (such as heavily weighted observations) in the estimates during review. This may be of use in the ORS program as well.

The data collected during testing for ORS are of limited use for developing estimate validation. First, the test data are not numerous enough to produce estimates that can be used as a testing ground for how the estimates should relate to one another. They also are not weighted. Additionally, through the various phases of testing, the ORS questions and collection procedures have evolved, making it difficult to compare data collected in the last phase of testing to data collected in the first.

Thus, in order to validate the estimates from testing, some different methods have been considered. The estimates within an occupation, or SOC code, are expected to remain consistent between phases. When the estimate differs, the procedural changes that may have been enacted are reviewed to determine if they played a role in the unexpected estimate. Unexpected differences in the estimates between characteristics like full/part-time and union/non-union are also examined to make sure they are not having effects where we wouldn't expect. Over time we should also be able to uncover any strong correlations between certain variables, such as time spent outdoors and lifting, which can then be used as a predictor for those estimates.

There are three different types of estimates in the ORS dataset. Some questions have categorical data, for instance an answer of either "Yes" or "No." Some estimates are averages based on continuous data, like a number of hours that a task is performed. Some estimates are percentiles, showing how much activity is performed at a certain threshold. In addition, some tabulations are created by compiling these different kinds of data, like the "SVP" or Specific Vocational Preparation, which combines categorical data like the degree required with continuous information like the number of months of training required to arrive at a new measurement. Each of these types of estimates requires validation.

Determining whether an estimate is fit for use isn't about liking the data, it is purely ensuring that quality thresholds for things like response rate and imputation are met. It is not the purpose to invalidate correctly collected data, only to ensure the process is working as it should. The end result of the estimation process should be to validate that the

construction of the estimate is good, which means that the estimation processes are working as intended and able to support the publication of accurate estimates.

4. Validation Techniques

The objective for a validation plan in ORS is to be able to ensure that estimates are reasonable compared to expectations. The validation options available will depend on the data element or elements being estimated and validated. The assumption in validating the estimates is that the underlying data have been reviewed in the ORS data review process. However, it may still be advantageous in the beginning phases of this new survey to leverage validation tools to inform data review or vice versa. Validation may reveal specific areas where the initial review of the micro-data may need to be confirmed, and in some cases, where some additional review may be necessary. Several tools for validating certain types of elements are being developed, including validation reports, outlier search software, and relationship dashboards and graphs. Each of these tools is described below.

Validation Reports - The most straightforward tool in conducting validation will be automatically generated reports. These will run from a program that looks at every single individual estimate and flags estimates that should be reviewed by trained staff, either because an estimate differs from a set expectation or because it differs from a historical record. To review how consistent the data within a SOC (Standard Occupational Classification) code are, the data that underlies the estimate for elements such as time to average performance can be evaluated using standard deviation. These reports might cast light on the specific areas to investigate further to identify the drivers of the estimates. Such investigations may reveal that the micro-data underpinning that estimate had outliers that were reviewed and verified during the ORS review process. It might also reveal that while data were verified, the weight of that observation is notably large, causing the outlier to have a larger than expected impact. It might trigger that an underlying piece of micro-data is errant. All of the above would be investigated and explained where and when possible and documented so that the estimates can be declared fit-for-use.

Outlier Search - Software options can be used to expeditiously drill down from any anomalous estimates to the micro-data identified as contributing to “outliers” to verify the underlying information. The anomalous estimates will be noted in other reports, but the visual nature of the dashboards allows users to click down from the estimate to the underlying data. For example, for the physical elements, the estimates will be available as percentile estimates, and these can be lined up in a graphical interface to quickly see if anything pops out as an outlier. By eliminating some of the manual work of reviewing individual responses, more data can be reviewed. Related to the outlier search is the examination of the relationship between elements, discussed below.

Relationship Dashboards and Graphs – JMP and Tableau Software easily integrate with SAS to produce color coded, interactive graphical interfaces from data. With these tools, estimates can be looked at within a broader context in order to identify which estimates are not conforming to expectations. The dashboard may, for example, compare estimates on lifting/carrying to estimates on weather. It seems that an established relationship might be that jobs where the incumbent lifts/carries more than 33 percent of the day should take place outdoors and be exposed to weather. With this visualization, any unusual deviation from that expectation is easily seen and can be explored further.

To plan for such an examination, a hierarchy of expected relationships has been established that indicates where a strong two or one way correlation should exist. These relationships are documented in a matrix, shown below, which makes it easy to visually comprehend the ways in which we expect the data to interact, and resultantly, show interrelationships between estimates. Across both axes of the matrix are the data elements and where the two elements meet in the middle, the box is shaded with a color to either indicate that there is a strong, weak, one-sided, or no relationship. As more data become available these relationships can become codified with evidence that they do have a statistical relationship to one another. The chart below shows an example of what such a matrix might look like. In this matrix, strong relationships are shown with various hues of green shading, weaker relationships are shown with various hues of yellow shading, and no relationship is shown with white boxes.

	Crawling	Crouching	Driving	Driving Vehicle Type	Far Visual Acuity	Fine Manipulation	Gross Manipulation	One on one hearing	Group hearing	Telephone	Other Sounds	Hearing Test	Keyboarding (10-key)	Keyboarding (Other)	Keyboarding (Touch Screen)	Keyboarding (Traditional)	Kneeling	Most Weight Lifted	Near Visual Acuity	
Climbing Ramps/Stairs, structural																				
Crawling	Green																	Green		
Crouching		Green																		
Driving			Yellow									Yellow								Yellow
Driving Vehicle Type				Yellow																
Far Visual Acuity					Yellow															
Fine Manipulation						Yellow						Yellow	Yellow	Yellow	Yellow					
Gross Manipulation							Yellow													
One on one hearing								Green	Green	Green										
Group hearing									Green	Green										
Telephone										Green										
Other Sounds											Green									
Hearing Test												Green								
Keyboarding (10-key)													Green							

We expect the estimates to have logical relationships between the elements, for example jobs with a higher measure of cognitive demand should have higher educational requirements. Other high level examples of relationships include expecting jobs with higher physical demands to often be found outdoors, similar physical activities to be grouped similarly, jobs that require contact with the public to include physical and cognitive elements that go along with such work. While these relationships exist at the micro-data level, we are learning and developing expectations about what they will look like at the aggregate, or estimate, level. No estimate would be expected to contradict any review edits. Factors such as region, industry, full/part-time, and union/non-union will be examined to ensure that they do not have effects on the data that would be unexpected. Our expectation is that they will not, but as we gather more data we hope to learn more about these issues. Over time, a library of expected relationships and variations will be built up from what is learned.

6. Conclusion and Next Steps

ORS poses unique validation challenges because of the unique nature of the data being collected. Little previous research has been done on collecting occupational cognitive requirements. Additionally, the physical and environmental demands are being documented in a new way for ORS. Validation plays a uniquely important role in the production of these estimates because of their newness. By pioneering new ways to identify unexpected patterns we can not only validate the data but gather information that can be used to improve the data review process, the future validation process, and occupational requirement research overall.

References

- [1] Social Security Administration, Occupational Information System Project, http://www.ssa.gov/disabilityresearch/occupational_info_systems.html.
- [2] U.S. Department of Labor, Employment and Training Administration (1991), “Dictionary of Occupational Titles, Fourth Edition, Revised 1991”
- [3] U.S. Department of Labor, O*Net Online, <http://www.onetonline.org/>
- [4] U.S. Bureau of Labor Statistics (2008) BLS Handbook of Methods, Occupational Employment Statistics, Chapter 3. <http://www.bls.gov/opub/hom/pdf/homch3.pdf>
- [5] U.S. Bureau of Labor Statistics (2013) BLS Handbook of Methods, National Compensation Measures, Chapter 8. <http://www.bls.gov/opub/hom/pdf/homch8.pdf>
- [6] See U.S. Bureau of Labor Statistics, Handbook of Methods, Chapter 8
- [7] See North American Industry Classification System website, <http://www.census.gov/eos/www/naics/>
- [8] See Standard Occupational Classification website, <http://www.bls.gov/soc/>
- [9] U.S. Bureau of Labor Statistics, National Compensation Survey: Guide for Evaluating Your Firms’ Jobs and Pay, May 2013 (Revised), <http://www.bls.gov/ncs/ocs/sp/ncbr0004.pdf>

Any opinions expressed in this paper are those of the author and do not constitute policy of the Bureau of Labor Statistics or the Social Security Administration.