Validation of Estimates in the Occupational Requirements Survey: Analysis of Approaches

Kristin N. Smyth

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

Abstract

The Occupational Requirements Survey (ORS) is an establishment survey conducted by the Bureau of Labor Statistics (BLS) for the Social Security Administration (SSA). The survey collects 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 those occupations are performed. This paper examines the process for validating estimates from ORS. It describes the purpose, procedures, and systems of validation and how they contribute to increasing the precision of the ORS measurements and properly gauging the quality of the final product. This paper also takes a look at estimates in the broader context of occupational research and other sources of information on occupational requirements. This paper examines what useful comparisons can be made between these sources and ORS and what further research would help to refine ORS validation procedures.

Key Words: occupational requirements, validation, data review

1. Introduction

In the summer of 2012, the Social Security Administration (SSA) and the Bureau of Labor Statistics (BLS) signed an interagency agreement to begin the process of testing the collection of data on occupational requirements. As a result, the Occupational Requirements Survey (ORS) began testing in late 2012. The goal of ORS is to collect and publish occupational information that will replace the outdated data currently used by SSA.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. More information on the background of ORS can be found in previous JSM papers or in the next section.

ORS 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 process of the estimates for ORS data. Section 2 provides background information on the Occupational Requirements Survey. Section 3 explains the context of the validation environment. Section 4 explores the tools of validation. 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 (SSA) administers two large disability programs which 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 restrictions (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. SSA has incorporated many DOT conventions into their disability regulations. However, the DOT was last updated in its entirety in the late 1970's, and 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 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.

SSA has investigated numerous alternative data sources for the DOT such as adapting the Employment and Training Administration's Occupational Information Network (O*NET) [3], using the BLS Occupational Employment Statistics program (OES) [4], and developing their own survey. But they were not successful with any of those potential data sources and turned to the National Compensation Survey program at the Bureau of Labor Statistics.

NCS is a national survey of business establishments conducted by the BLS [5]. 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 in the Federal Pay Setting process.

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) [6]. Each job selected for study is classified using the Standard

Occupational Classification system (SOC) [7]. In addition, each job is classified by work level – from entry level to expert, nonsupervisory employee to manager, etc. [8]. 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 training 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 infrastructure to collect data on occupational requirements.

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 infrastructure. Field testing to date has focused on developing procedures, protocols, and collection aids using the NCS infrastructure. These testing phases were analyzed primarily using qualitative techniques but have shown that this survey is operationally feasible.

The pre-production test might better be described as a "dress rehearsal" as the collection procedures, data capture systems, and review processes were structured to be as close as possible to those that will be used in production. The sample design for the pre-production test was similar to that which will be used in production, but was altered to meet test goals. While 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, in FY 2015 BLS integrated the prior work into a large-scale nationally representative pre-production test. For more information on the pre-production test there is a BLS website [9].

3. What is the Purpose and Context of Validation?

Validation of the ORS Estimates refers to the review of the 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 to confirm that our methodological processes (estimation, imputation, publication and confidentiality criteria, and weighting) are working as intended.

Both reliability and validity are concepts that are important to the ORS validation efforts. Reliability describes a measure's ability to be reproduced under the same conditions whereas validity is a measure of the relationship between the data and its intended purpose. In order to examine the reliability and validity of the ORS data, a description of the data and its purpose is first required.

The ORS data (see Table 1) consists of occupational information on experience and educational requirements, cognitive demands, physical demands, and exposures to environmental conditions. Data elements that are categorical measures, such as whether or not the worker is exposed to wetness, have a set of predetermined values, one of which will be selected as a response. Continuous measurement data, such as number of hours required to stand, may be limited by a minimum, such as zero hours, or a maximum, such as 100 percent. Some measures, like the physical demands and environmental conditions, are concrete conditions that likely have precise measurements. Some other measures, like cognitive ability, are concepts for which the measurement will necessarily be more indirect.

There is a wide array of physical demands covered in the survey. Several different levels of hearing that might be required are investigated, such as the ability to hear on a telephone or with another person one-on-one. Postural requirements, like sitting and standing or moving low to the ground in a crouch or a kneel are recorded. So too are pushing and pulling with any of the extremities.

Educational requirements like the time it takes for any required schooling or licensure are also investigated. The minimum required previous work experience combines with these other areas to create a measure called specific vocational preparation, or SVP, which aims to give a sense of the time required to prepare for the job.

Environmental conditions reveal whether the job needs to be performed in extreme temperatures or exposed to certain kinds of risks like heights or chemicals. The cognitive elements of the survey ask about decision-making, supervision, and routine in an occupation.

Table 1 – ORS Pre-Production Test Data Elements

Educational Requirements -- 4 elements

Minimum Formal Education or Literacy required

Pre-employment Training (license, certification, other)

Prior Work Experience

Post-employment training

Cognitive Elements -- 9 elements

Closeness of Job Control level

Complexity of Task level

Frequency of Deviations from Normal Work Location

Frequency of Deviations from Normal Work Schedule

Frequency of Deviations from Normal Work Tasks

Frequency of verbal work related interaction with Other Contacts

Frequency of verbal work related interaction with Regular Contacts

Type of work related interactions with Other Contacts

Type of work related interactions with Regular Contacts

Auditory/Vision -- 10 elements

Driving, Type of vehicle

Communicating Verbally

Hearing: One on one

Hearing: Group

Hearing: Telephone

Hearing: Other Sounds

Passage of Hearing Test

Far Visual Acuity

Near Visual Acuity

Peripheral Vision

Environmental Conditions -- 11 elements

Extreme Cold

Extreme Heat

Fumes, Noxious Odors, Dusts, Gases

Heavy Vibration

High, Exposed Places

Humidity

Noise Intensity Level

Outdoors

Proximity to Moving Mechanical Parts

Toxic, Caustic Chemicals

Wetness

Exertion -- 14 elements

Most weight lifted/Carried ever

Push/Pull with Feet Only: One or Both

Push/Pull with Foot/Leg: One or Both

Push/Pull with Hand/Arm: One or Both

Pushing/Pulling with Feet Only

Pushing/Pulling with Foot/Leg

Pushing/Pulling with Hand/Arm

Sitting

Sitting vs Standing at Will

Standing and Walking

Weight Lifted/Carried 2/3 of the time or more (range)

Weight Lifted/Carried 1/3 up to 2/3 of the time (range)

Weight Lifted/Carried from 2% up to 1/3 of the time (range)

Weight Lifted/Carried up to 2% of the time (range)

Reaching/Manipulation - 14 elements

Overhead Reaching

Overhead Reaching: One or Both

At/Below Shoulder Reaching

At/Below Shoulder Reaching: One or Both

Fine Manipulation

Fine Manipulation: One Hand or Both

Gross Manipulation

Gross Manipulation: One Hand or Both

Foot/Leg Controls

Foot/Leg Controls: One or Both

Keyboarding: 10-key

Keyboarding: Other

Keyboarding: Touch Screen

Keyboarding: Traditional

Postural -- 7 elements

Climbing Ladders/Ropes/Scaffolds

Climbing Ramps/Stairs: structural only

Climbing Ramps/Stairs: work-related

Crawling

Crouching

Kneeling

Stooping

There are multiple purposes for the ORS data and many stakeholders including SSA, the general public, and the businesses and organizations that need occupational requirements data. Each of these stakeholders may have different uses in mind for ORS data. Primarily though, the ORS data hopes to document the functional requirements of occupations so that users can compare their functional capacities to that required for different occupations. Information about the requirements of work in the national economy is required for the SSA disability determination process, but it also may be of use to job seekers, researchers, insurance companies, and advocacy organizations. Whatever the final use of the data, ORS aims to be a reliable and accurate measure of the requirements for occupations in the United States.

There is no one survey or data source to which we can compare our estimates to ascertain reliability. However, there are portions of other surveys that can provide some context for our estimates.

One of those surveys is the BLS's National Compensation Survey, or NCS. Some comparisons made in the micro data review stage of ORS can also be used in the estimate validation. Using just the microdata, we have looked at the strength measurement in ORS compared to the Physical Environment measures in the NCS. The physical environment measure includes an aspect of risk in the job in addition to the physical strain of the job, so it will necessarily be different from the strength measure, but there still appeared to be a general trend of agreement between the two surveys.

Another comparison was between the Job Controls and Complexity element in NCS and the complexity element in ORS. We found that a general trend of agreement exists between the two, with some interesting outliers that can be explained in part by the difference in the two measurements. In testing, only a small amount of the data was a clear non-match between the two sources.

NCS also contains leveling information that has already been used in ORS collection and micro-data review to prompt extra scrutiny when a collected response differs from the NCS data. But this leveling data, if calculated as an average NCS level by detailed SOC, could also be used for estimate validation.

Another useful data source is the Dictionary of Occupational Titles (DOT) that the ORS will be replacing. It is relevant to our survey because it is widely acknowledged as the predecessor to this data. It covered similar topics, such as physical requirements and educational preparation. But it was last substantially updated in 1991 and fully updated in 1977.

The DOT used a more detailed classification of occupations by Standard Occupational Classification (SOC) than ORS uses. The idea of an occupation is obviously a major building block of occupational data. An occupation is a construction, however, not a naturally occurring phenomena that we can definitively draw a line around. The representatives from the major governmental statistics agencies convene every eight years to update the SOC system in order to have a relatively modern classification of occupations in the United States. Obsolete jobs are removed from this list and new jobs are added. But in practice many jobs combine the tasks of several distinct on-paper occupations, leaving some natural variability within a SOC classification. The data on occupations divided by SOC will differ from the data in the DOT because the occupational classifications are aggregated differently.

As a limited exercise, the ORS estimates for the amount of job preparation, SVP, were compared side by side to a closely related occupation offered in the DOT, which also has an SVP measurement. Less than a third of these estimates, which were a small subset of the total occupations collected in ORS, differed by more than one level.

Another useful data source is the Occupational Information Network, or O*NET, which includes updated lists of the tasks that the incumbent performs, something that is also collected in ORS. The O*NET is a database of occupational characteristics that was intended to move beyond the DOT. The ORS estimates for SVP were lined up side by side to the O*NET, which has a measurement called the "job zone" that is roughly comparable. About a third of the mode estimates for ORS SVP differed by more than a single level from the O*NET job zone. A single level could be a difference in a single day of preparation up to a thousand or more days at the higher SVP categories.

We can use these other surveys and data sources to inform our expectations for ORS estimates but the drawbacks described above also indicate the need for a survey that covers all of this information reliably and timely. The validation process that we execute on our estimates is one part of helping to determine that the estimates are FFU. Determining whether an estimate is FFU isn't about liking the data, it is ensuring that quality thresholds are met. It is not the purpose of validation 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 the estimates follow expectations. By validating the reasonableness of the estimates, investigating anomalies, and documenting results, validation is able to support the publication of accurate estimates.

4. What are the Validation Processes?

The assumption in validating the estimates is that the underlying data have been reviewed in the ORS data review processes. This review is designed to ensure the accuracy, consistency, integrity and quality of the microdata. Once that data are ready for analysis, typically estimation is run along with additional processes such as weighting, imputation, suppression and applying publication and confidentiality criteria [10]. In ORS, imputation and suppression are still in development and are not yet fully applied. As detailed in the previous section, one part of validation is understanding the purpose of our study and using other surveys and data sources as a general comparison to get a feel for our data. This helps to evaluate the reasonableness of the estimates. Estimates that do not meet expectations are investigated, explained, and documented. Next is an investigation into the additional processes that are run in estimation.

Weighting accounts for how much influence a particular response have on the estimate based on the sample design. Weights are applied and run in estimation, but if there is an unusually large weight that sways the estimate it may be picked up in validation.

Imputation will aim to correct for item non-response in the data. Research on imputation is ongoing. In general, a usable quote must have a valid SOC code and at least one non-missing data entry for an ORS element. Quotes without the minimum amount of data will be treated as non-respondents during unit non-response adjustment and may not be donors or recipients for imputation. Responses coded as unknown or that are blank are considered to be a non-response.

The validation process will need to include checks on imputation when the procedure is implemented. The recipients could be checked to see if they all matched to a donor, and the maximum times a donor is used could be checked to make sure usage is low. The change in the estimate can also be compared from before imputation and after.

Suppression, which would occur after the publication criteria has been applied, involves eliminating an estimate if it is determined that there was an error in the underlying data or the construction of the estimate and the impact of the error moves the estimate outside of its confidence interval. This is performed when it is not possible to go back and correct the data or the construction, which is the preferred action. In ORS it is not yet determined the extent to which suppression will be necessary or whether it will be more possible than in other programs to go back and correct the data. If the final estimates lend themselves to revealing suppressed estimates through deduction, such as by figuring out a suppressed total as a share of 100, there will also need to be secondary suppression where additional estimates are removed to preserve the table's integrity.

The publication criteria, which involves reliability and confidentiality criteria, are the standards that the estimates need to meet in order to be published. Currently in ORS the plan is to have a publication criteria that would deem an estimate publishable if the estimate cell has a minimum number of unique establishments and usable quotes, with a relative standard error that does not exceed our given threshold. It also includes a dominance rule, which means no single establishment can dominate the estimate. As the survey is ongoing, it is unknown whether these specific criteria will yield the quantity and quality of estimates that we want to have going into publication, so this will be something to keep an eye on moving forward. Increasing the unique establishment and quotes numbers may yield slightly more accurate estimates with a lower relative standard error, but it would come at the cost of losing occupations for publication. Likewise, having these thresholds too low could result in an increase in the quantity of estimates available at a cost of data quality.

The publication criteria also establish a ceiling for the relative standard error (RSE) of an estimate. If the RSE is above that ceiling, it will not be eligible for publication. The RSE helps to validate the reliability of these estimates. As there will be many more units in the next phase of production, we expect the number of estimates failing the RSE criteria to drop. The pattern currently in the estimates is that they tend to pass if they are one of the more frequent responses, like 'not present.' Moving forward the analysis on RSE will help to determine exactly where the cutoff for publication should be and where there is the most variation in the data.

Estimates that do not meet certain criteria will violate confidentiality standards and thus not be published. Confidentiality criteria tests are used to identify these unpublishable estimates. These tests are be performed after estimates have been calculated, and a set of confidentiality flags indicate whether the estimate can be published. These tests look at the number of firms and quotes in a cell as well as whether a single establishment dominates the estimate.

5. What tools do we use for Validation?

The validation tools at our disposal in these investigations into the estimates and the estimation processes include reports, visualizations, and automation. The end-goal for ORS validation is to have in place a system that allows a reviewer of estimates to quickly

examine all the notes and reviews that have already taken place on a given data point or collected establishment. This will involve merging systems of collection, data validation, and estimate validation.

Validation Reports – These are automatically generated reports that we can run at any given moment to give an up-to-date look at how the estimates are performing. They cover different topics that we may want to look at in the estimates. For example, one report looks at any cases where one employer has more than a certain percent contribution to the total weight. This report is intended to find any instances where an estimate is driven primarily by one respondent. It reports any estimate where a single employer has a large contribution to the total weighted employment and what that contribution is. This dominance criteria is also in place for the publication criteria.

For any of these reports, once an item is flagged for attention it is assessed as to whether that flag is the result of either an imputation or the collected data. If the item was imputed, the donor is examined in a separate imputation database to see at what level the match occurred as it may have been a less than ideal match. If it is collected data, the item is reviewed using logged data from the review process to see if that item was a part of a review for that schedule. If it wasn't, perhaps whether it should be looked at.

If an error is found, there are options including suppressing the estimate No decisions have yet been made on how exactly this will be run.

The other reports that are currently available include:

Unusual Specific Vocational Preparation (SVP) Results for the 2-digit Standard Occupational Classification (SOC): This dashboard shows the estimate for no, low, and high skilled jobs for easy review.

As an example, one might see there are slightly large high skills estimates for traditionally lower skilled occupations, which would merit further review. The sub cell estimates could be quickly checked to see that none show a large deviation from the others. However, some highly weighted observations might be coded as requiring college degrees. As the sample gets larger these few observations will become less important and the estimate should even out to a lower and more expected SVP.

Unusually high prevalence of exposure to environmental conditions for the SOC: This shows any meaningful estimate for a present environmental condition. The assumption is that these environmental conditions, like extreme temperature or exposure to chemicals, wetness, or heights, do not exist often and can quickly be reviewed to ensure they make sense.

Unusually high standard errors within a two-digit SOC for selected physical demands: This report shows any estimate with a high standard error in a physical demand characteristic.

Unusually high standard error within a two-digit SOC group and SVP category: This report shows any estimate with a high standard error in these categories.

Worker Characteristics: This report compares estimates across worker characteristics and highlights unexpected differences. It can also be adapted to show differences in other categories where estimates are run, so that if estimates by industry or establishment size are run, those numbers could also be compared.

Identifying which estimates are eligible for publication, i.e. by passing the publication criteria, the confidentiality criteria, and having enough estimates to be useful, helps to narrow down which estimates need to be examined. The expectation is to keep refining these procedures and add or remove reports to get looks at the data that ensure they meet the quality expectation.

For pre-production the estimates were run when 75% percent of the data were collected and again at 100%. While the dashboard creation can begin with the earliest round of estimates, the possibility of data or estimation changes prevents some report activity until the database closes after 100% collection.

Dashboards have become a standard practice around the tech world for slicing data in ways that users can easily digest and take away meaning. With estimate review the goal is to have dashboards that any reviewer or interested internal party can view in order to get a sense of if the data is reliable and valid. The visualization part of this tool is beneficial over the old style tables of data because it takes that information, which may be spread across tables thousands of rows long, and puts it into a graphical format where the user can easily see the outcomes. The information will be uniformly available to staff because of the easy adoptability of software that reads the visualizations.

A key component of the data visualization effort is to have the ability to drill down or click through to the underlying data. This is of use during the further review of estimates that occurs after they are flagged in the reviews. Older survey systems utilize different programs for different tasks which results in a lot of duplicate work in terms of review. By tracking all of the review comments from the earliest stages of a schedule through estimation, it is easy in the final stage to isolate what part of the process needs to be re-examined. Moving forward, this is envisioned to include the ability to move from a high level view of the graphical validation reports to any notes on the schedule, information on imputation, and information on suppression.

6. What We Learned in Pre-Production Validation

As mentioned earlier, the first step in running the reports for pre-production was to narrow down which estimates were eligible for publication and focusing the validation effort on those estimates. This involved using the publication criteria and narrowing the list of occupations to the ones where BLS had collected the most data.

The environmental dashboard was helpful in reviewing the estimates because there are only a few environmental elements that code for present and all of those could be reviewed. In one instance there was an estimate for how often an occupation was exposed to wetness on the job, and it was quick to drill down to see which particular data point was driving that estimate and make sure there was no weighting or imputation issue that would render the estimate invalid.

The SVP dashboard functioned much in the same way, and estimates that were unexpected were documented. Then the micro-data was examined for an explanation. Any estimate that is reviewed is documented.

The strength element is a unique case where several pieces of information are combined to come up with a measurement. Strength depends on the amount and frequency of weight lifted as well as how often the worker stands and pushes with their extremities. In validation for pre-production the possible minimum and maximum strength measure for data points were hand coded in order to give us an idea of how the Strength measure changes based on what is coded as unknown. This dataset may form the basis for how Strength is coded in the future.

SVP is another case where other elements are put together to calculate the SVP measure. Analysis helped to identify which of these underlying pieces- such as amount of post-employment training or degree requirements, most contribute to the measure and the frequency of when it is unknown. This in turn helped to refine collection training for production in order to maximize the information gained from respondents.

The relative standard errors are still high in this pre-production test but we expect them to perform better with more data. For most occupations, the most popular estimates such as whether a condition was present had acceptable relative standard errors.

How will we know that the validation effort has been a success? We aim to produce reliable and timely data that has little variation within an occupation and characteristic, thereby demonstrating reliability, and that matches up well with any pre-existing context we might have for the information, thereby demonstrating validity.

7. Conclusion and Next Steps

The pre-production validation effort evaluates the estimates for reasonableness, explains any anomalies, and provides better tools to make our efforts more efficient. In the Pre-production test we tried to eliminate duplicate review work that exists in older surveys. By using new technology, the information gathered at one stage in the survey process is readily available to the rest of the review staff down the line, and this has helped to make the process incredibly efficient. As more decisions are made regarding imputation procedures and benchmarking procedures in the future, the visualization system of validation will be able to adapt to the new procedures and provide information that is needed to ensure the estimates are fit for use.

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]   National Compensation Survey, www.bls.gov/ncs

[6]   See North American Industry Classification System http://www.census.gov/eos/www/naics/

[7]   See Standard Occupational Classification http://www.bls.gov/soc/

[8]   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

[9]   http://www.bls.gov/ors/#pretesting

[10]   Rhein, B. and C. Ponikowski. 2015. Estimation Processes Used in the Occupational Requirements Survey", In *JSM Proceedings*, Survey Methods Research Section. Alexandria, VA: American Statistical Association.

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.