

NILG 2020 Virtual Conference Webinar Series

NILG Advisory Council Panel on Compensation Standards and Best Practices - Compensation Roundtable Follow-up with OFCCP – July 29, 2020

Experts answers to attendees' questions

1. How does David Cohen define "covered by statistical analysis"? This is very similar to the guidance the Agency already uses. The open question is how to define statistical sufficiency...

Cohen - Bob defines statistical coverage as using the 30 and 5 rule at least 70% of the workforce is covered under a regression and the remaining 20-30 get covered under small group tests. I would probably say 70% for both regression and small group but by all means if you can get a higher coverage without violating similarly situated that is great."]

2. How does OFCCP adapt the statistical analysis (especially the groupings) for smaller federal contractors?

The Roundtable experts refer this question to OFCCP's Branch of Expert Services for response.

3. What criteria does OFCCP use to group jobs?

The Roundtable experts refer this question to OFCCP's Branch of Expert Services for response.

4. For DK: on small sample tests, what about the common situation in which there are no comparators in small groups? And would contractors really want the Agency to use tests without controls for things like different experience levels between comparators when examining small groups?

DK - First, "**...what about the common situation in which there are no comparators in small groups?**" Generally, measuring pay equity, under a classic Title VII framework, requires comparators within a group of similarly situated individuals. For example, if you are interested in measuring gender pay equity for a particular job, but there are only women in the job, is it possible to statistically measure pay equity as a function of gender? The logical and common sense answer is obvious. However, for some analysts, they believe that you can measure pay equity in this situation if you aggregate and fold this job in with other jobs where there are comparators, and analyze them as one Pay Analysis Group (PAG). As a group of experts, we do not believe that it is appropriate to analyze a PAG that is comprised of disparate jobs where some have comparators and others do not have comparators. Those results and conclusions can be highly inaccurate and unreliable.

Second, "**...would contractors really want the Agency to use tests without controls for things like different experience levels...?**" Contractors desperately need the Agency to practice good science in a pay equity investigation. They want the Agency to measure pay equity among truly similarly situated groups of individuals, and not Pay Analysis Groups that merely "approximate" a similarly situated standard by grouping disparate jobs that are "expected" to be paid the same. Needless to say, accuracy and reliability are crucial. This question suggests that contractors must choose between two competing options in small sample situations: (1) measure pay equity with control factors but aggregate jobs into Pay Analysis Groups; or (2) measure pay equity with no control factors using small sample analytical methods among similarly situated individuals (e.g., jobs). While it is true that you cannot **statistically** control for explanatory factors when using small sample analytical methods, it is not true that explanatory factors are completely excluded from small sample pay equity investigations. Cohort analysis is a well established, recognized, and accepted method for evaluating pay differences while controlling for explanatory factors in small sample situations. Even the EEOC recommends using cohort analysis in small sample situations (EEOC Compensation Compliance Manual). Indeed, it is possible to accurately measure pay equity among truly similarly situated individuals in small sample situations while controlling for explanatory factors.

5. With all due respect, the Agency *never* continually combines groups until we find a problem.

Rob - That has not been our collective experience as we often see the Agency aggregating dissimilar groups of employees to achieve "critical mass." Whether they are combining groups or failing to separate distinct groups, we cannot say as the outcome is the same. If you are a contractor and did not have the same experience, we would like to discuss your experience with you so we have a better understanding of how the Agency conducted its analysis. If this is a commitment from the Agency to not combine dissimilar groups of employees for the sole purpose of achieving "critical mass" to be able to run a regression analysis and to instead rely on the appropriate application of widely accepted small sample tests, we welcome this change.

6. Could you clarify a comment from Dr Bob? Did he just say that an EEO Establishment equals similar jobs? That the way they are set up - the jobs at the location are usually similar?

Mike/Jora – Yes, the expert panel also heard the OFCCP Director of Enforcement make those comments, which we believe refers to using EEO-1 job categories as job groups (more properly, SSEGs). The panel respectfully disagrees with the belief that all or most jobs in an EEO-1 job category are similarly situated. A great example would be EEO Job Group 2: Professionals. For most contractors, this job group would contain such occupations as engineers, accountants, HR professionals, chemists, software programmers, lawyers, graphic designers, technical writers, economists, and health professionals. Not only do these jobs represent different job families, but they represent multiple grades and bands. Thus, employees in these jobs do not perform similar tasks, don't have similar skill levels, and are not expected to be paid the same. In short, EEO job groups don't meet the OFCP's own definition of similarly situated and EEO job groups do not represent the salary structure of any federal contractor.

7. Wouldn't I only interact two variables if I had a good reason to think the effect of the first factor varies importantly with the effect of the second one? And if I'm the employer as well as the author of my compensation manual, don't I almost necessarily have an identifiable good reason?

Rob - As discussed in the presentation, and as confirmed by Dr. LaJeunesse of the OFCCP in his comments, the application of interaction terms is complicated and can be confusing. It is difficult to do right. But, if you aggregate, either jobs or years of data or both, you need to be concerned with interaction terms – both how you implement them and how you test for their importance.

There are two vague terms in your question – “good reason” and “importantly.” It is not clear what you mean by “good reason,” but the second question seems to imply that it should be found in a compensation manual. In our experience, compensation manuals have some general factors that impact employees’ pay, but they do not include every factor that impacts every employee’s pay, except in a collective bargaining agreement. Documenting every factor for every job would be an impossible and never-ending task. It is also unrealistic to assume that HR managers, with limited if any statistical training, would include a term like “interaction terms,” or even the concept, in their compensation manuals.

These manuals, however, will often include pay factors like time-in-job and service. The fact that they may not say that pay increases at different rates for those in different jobs does not preclude it from being true and reasonable. In our experience, pay may grow at different rates for those in different jobs. For example, pay grows faster for those with less experience than for those with more experience¹ and the difference in experience, among other things, may be captured by the job title. Moreover, the interaction of job and pay factors may extend beyond time-in-job and service.

Being trained in various disciplines that employ statistical analysis, we will respond to your question from that perspective. We would interpret “good reason” and “importantly” to mean are they statistically justified. The only way to know is to include the necessary interaction terms and to properly test them. If they are statistically significant, then there is “good reason” to include them and pay varies “importantly” with them.

One final comment, Dr. LaJeunesse discussed interacting gender with service, which is not the type of interaction we were discussing or even contemplating. Such an interaction greatly complicates the interpretation of the statistical results, in particular any observed gender pay difference.

8. To clarify: what slide in "Leen Prin. #7" is referred to as the "10:1 rule" is not a *rule*. It is never strictly followed. It is a very rough guideline for thinking about statistical power, at ranked fairly low in Agency considerations on an "optimal" PAG.

¹ See, for example, the well-known age-earnings profiles found in many labor economics textbooks and studies, which show pay increasing rapidly early in a career and then tapering off.

Jora - Thank you very much for your comment. We always appreciate feedback. OFCCP's FAQ explicitly describes following this rule ("OFCCP then additionally tries to ensure that there are at least 10 observations (or employees) per control variables to be able to conduct a sound statistical analysis.") and we have seen it followed in practice. If the Agency does have a ranking of considerations for "optimal" PAGs, it would certainly increase certainty and efficiency for contractors if it were explicitly codified and shared, as part of the Agency's commitment to improving transparency.

9. On the "5 rule", which is enforced, the relevant question for the experts is this: what is the minimum number of individuals in a protected class that you would accept for the Agency to say "we have a systemic pay disparity here." If there is a single female, and the female control applies only to her, would you accept this as a finding? What if there were two females, etc. What is the minimum you would accept?

Valentin/Jora - Thank you for your question about systemic discrimination, small groups, and the "5 rule."

There are two separate, but related "5 rules" that we think may be conflated in this question:

- First, directive 2018-05 states that control variables should "Control for race/ethnicity by creating a series of component dichotomous (0-1) variables for each race/ethnicity category using the category with more than five (5) employees and the highest average pay in each PAG as the reference category."
- Second, the OFCCP requires that "...each category [of a pay factor included in a regression analysis] contain at least five observations.". See, for example, the FAQ which states "To capture meaningful pay differentials across the categories, OFCCP requires that each category contain at least five observations. If a category has fewer than five observations, OFCCP will join those observations with their ordinal counterpart (e.g. nearest grade or level) or to the category with the nearest average pay."

The second rule, of combining levels of a particular variable to reach at least 5 employees per level, is referred to by the OFCCP as "variable fusion" and is what we were discussing in the webinar. It can lead to results that are inaccurate and

unreliable. We have seen small race categories being combined as part of this practice, which certainly leads to uninterpretable findings.

Regarding the issue of jobs with few incumbents, we recommend that the Agency recognizes that not all employees will have a comparator because of their duties and tasks. Aggregating dissimilar employees just to use multiple regression or small-group techniques is not a sound statistical practice.

Finally, the OFCCP recognizes that systemic discrimination refers to multiple and consistent (same direction) compensation differences in group averages. For example, during the presentation of the OFCCP's Branch of Expert Services (BES) at the NILG 2020 Virtual Conference the Agency recognized that isolated or inconsistent indicators are not proof of systemic discrimination. However, the Agency has not disclosed when it considers that a group of indicators is evidence of systemic discrimination. This lack of transparency contradicts the Agency's own directives and adds uncertainty to the audit process.

10. The presenters repeatedly state that the OFCCP should use the groups developed by contractors because contractors spend the time, energy, and money to develop them. But I'm not sure what we're talking about: Aren't these simply the AAP job groups? Well, the Agency always makes an effort to use AAP job groups so I'm not sure what this is referring to.

Val -The groups developed by contractors that we reference are not AAP job groups; they are "similarly situated employee groups" or SSEGs. SSEGs are groups developed under the principles articulated in Title VII caselaw in connection with identifying employment discrimination, including pay discrimination. AAP job groups are rarely SSEGs; SSEGs are comprised of employees who have similar work, level of responsibility, skills and/or qualifications. In contrast, AAP job groups generally are broader than a proper SSEG and do not limit themselves only to those employees with appropriately similar work, responsibility, skills or qualifications. Contractors should reject any OFCCP allegations of pay discrimination as not consistent with Title VII where overbroad AAP job groups or "pay analysis groups" (PAGs) are used in lieu of SSEGs.