EEO INSiGHT

How to Compute Liability
WHEN STATISTICAL EVIDENCE OF PAY DISPARITIES EXIST *Part I*

Daniel Kuang, Ph.D. | *Biddle Consulting Group, Inc.*
Dan A. Biddle, Ph.D. | *Biddle Consulting Group, Inc.*

Differences in average compensation between two groups (e.g., men and women) are not necessarily evidence of pay discrimination. However, if those differences are large enough to be statistically significant, and if they cannot be justified by job-related factors, they may be used as evidence of pay discrimination. Under these circumstances, non-defensible differences should be remedied through compensation adjustments.

This article will focus on how to address the gaps between the amounts that individual employees are paid and how much the regression model indicates they "should" be paid. However, it will not discuss the variety of options employers have regarding exactly how or when to make these adjustments, though that is just as important as identifying who needs an adjustment and how much.

The primary purpose in making compensation adjustments is to close the gap in compensation between the Focal (*e.g.*, women/minority) and Reference (*e.g.*, men/white) groups after controlling for differences in job-related factors. Despite how seemingly straightforward this should be, the massive volume of literature on compensation liability models provides sobering evidence of the complexity involved in properly modeling compensation adjustments.[1]

Fortunately, there is a very general and flexible method that may be applied in almost all circumstances. It is a methodological framework that is comprised of two steps:

Step 1: Compute Liability — the total amount of money to be paid to the negatively impacted group.

Step 2: Determine Distribution — the method of identifying how much of the total liability to distribute to each individual within a negatively impacted group.

This article will detail *Step 1—Compute Liability*. Step 2 will be detailed in a separate article.

COMPUTING LIABILITY—DESCRIPTION

The key to a general and flexible compensation liability model is a commonality in the underlying statistical model. The typical compensation analysis falls into one of two categories when it tests for differences between groups: 1) without controlling for explanatory variables (*e.g.*, *t*-Test, ANOVA – analysis of variance) or 2) with controlling for explanatory variables (*e.g.*, multiple regression). Despite differences in the analytical strategies, their underlying statistical model is the same; they are all variants of multiple regression.

Evidence of a significant between-group difference in compensation exists when the regression coefficient

(b) for the group variable (*e.g.*, men/women, white/minority, etc.) remains significant after controlling for differences in job-related variables. The following statistics are needed from the regression output to compute the amount needed to "eliminate" group differences (to varying degrees):

1. b_{Group}: Regression Coefficient for the group variable (*e.g.*, gender/race).
2. SE_b: Standard Error for the Regression Coefficient (b_{Group}).
3. N: Total sample size.
4. k: Total number of independent variables in the model.

If the group variable is coded properly (0=Focal/Women/Minority and 1=Reference/Men/White), the b is the mean difference in pay between Focal and Reference group members, shown in Equation 1 as:

$$b = \text{Mean}_{Focal} - \text{Mean}_{Reference} \qquad \text{(Eq. 1)}$$

- If $b < 0$, then the Focal group is negatively impacted.
- If $b > 0$, then the Reference group is negatively impacted.

In a Title VII context, a significant b, irrespective of its directionality (positive or negative) is an indication of potential compensation discrimination.

When the regression model does not contain explanatory variables, then the b obtained in Eq. 1 can be interpreted literally: raw mean difference in compensation between Focal and Reference groups.

However, the typical regression model will include one or more explanatory variables. In such instances, the b in Eq. 1 is the mean difference between the Focal and Reference group *after* controlling for differences in the explanatory variables. This is often referred to as the "adjusted mean."

The statistical test to determine whether b is significant is:

$$t = b \,/\, SEb \qquad \text{(Eq. 2)}$$

With

$$df = N - k - 1 \qquad \text{(Eq. 3)}$$

METHOD

Once these statistics are obtained, computing compensation adjustments becomes a fairly straightforward mathematical exercise. The steps for computing the amount needed to "eliminate" significance are as follows:

1. Determine the desired legal defensibility: In standard deviation units, what is the tolerable pay disparity (*i.e.*, 2, 1, or 0)? Once the desired standard deviation difference in pay disparity is determined, compute the p-value. Common thresholds are computed and presented in Table 1.

   For the advanced analysts who desire to apply specific standard deviation units, they may convert standard deviation units into 2-tail p-values with a statistics table or apply the following formula in Excel:

   =2*(1-NORMSDIST(Standard Deviation)).

TABLE 1: Establishing Acceptable Levels of Legal Defensibility

Standard Deviation	p-value[a]
2	40
1.95	25
1	4
0	2

Note: [a] 2-tail p-value.

It is important for employers to understand that reducing salary differences to two (2) standard deviations will cost less than reducing the disparities to one (1) or zero (0) standard deviations. However, this gives the employer very little "cushion." Meaning, that even small changes in salaries or workforce composition can/may cause the statistically significant disparity to reappear.

2. Determine the non-significant *t*-value: Once the desired and tolerable *p*-value is determined (Step 1), the next step is to compute the *t*-value for the available degrees of freedom (*df*). The non-significant *t* may be obtained from a statistics table or Excel with the following formula:

$$t_{non\text{-}significant} = \text{TINV}(p\text{-value}, df) \qquad \text{(Eq. 4)}$$

3. Compute compensation adjustment: Once the non-significant *t*-value is determined (Step 2), the next step is to insert this value and component from the original *t*-test formula into the following formula:

$$\text{Liability}_{\text{individual}} = (t_{non\text{-}significant} \times SE_b) - |b_{\text{Group}}| \qquad \text{(Eq. 5)}$$

This computed compensation liability is at the individual level. Specifically, it is the amount that needs to be adjusted for each individual in the impacted group to reduce the pay disparity to the desired level. The total liability for the impacted group is:

$$\text{Total Liability} = N \times \text{Liability}_{\text{individual}} \qquad \text{(Eq. 6)}$$

To confirm the validity of these adjustments, a "what-if" simulation analysis can be performed. In such an analysis, calculated adjustments are added hypothetically to the appropriate employees in the database and the pay disparity between Focal and Reference members is re-evaluated. If the results of the statistical test matches the desired pay disparity (*e.g.*, 0, 1, 2 standard deviations), then the computed liability is valid.

CONCLUSION

This paper detailed a general method of computing liability within a multiple linear regression framework. Although the mechanics of computing liability is fairly straightforward, it is important that analysts understand the concepts of this method prior to making any pay adjustments.

Please note that the method detailed in this paper is only one of two steps in a comprehensive pay adjustment study. This first step details how to compute the total amount necessary to diminish the pay gap between focal and reference members in a group. A future article will detail methods of distributing the computed liability.

ENDNOTES

1. Colquitt, J., Conlon, D.E., Wesson, M.J., Porter, C., & Ng K.Y. (2001). Justice at the millennium: A meta-analytic review of 25 years of organizational justice research. *Journal of Applied Psychology, 86*(3), 425-445.