Received: 27 January 2021 | Accepted: 15 December 2022

DOI: 10.1111/poms.13944

ORIGINAL ARTICLE

Bridging the gap: Applying analytics to address gender pay inequity

David Anderson[1] | **Margrét Vilborg Bjarnadóttir**[2] | **David Gaddis Ross**[3]

[1]Villanova School of Business, Villanova University, Villanova, Pennsylvania, USA

[2]Robert H. Smith School of Business, University of Maryland, College Park, Maryland, USA

[3]Warrington College of Business Administration, University of Florida, Gainesville, Florida, USA

Correspondence
Margrét Vilborg Bjarnadóttir, Robert H. Smith School of Business, University of Maryland, College Park, MD 20742, USA.
Email: mbjarnad@umd.edu

Handling Editor: Sean Willems

Abstract

Employers are under increasing pressure from regulators, investors, and their own employees to address gender pay inequity in their compensation structure. However, managers tasked with measuring and closing gender and other pay gaps have few helpful quantitative tools at their disposal. We design and implement an algorithm that measures whether an employer has a pay gap and suggests salary adjustments to close it, all while supporting the employer's goals with respect to equity, culture, and incentives in its pay structure. We present our approach through a case study of one of the first firms we worked with. We also discuss the application of the approach at other organizations ranging from 75 to 130,000 employees. Our work not only provides management personnel with an accessible solution to a problem of growing social and practical importance but also demonstrates how the operations management toolkit may be applied to improve human resources management and address an important social problem.

KEYWORDS

case study, compensation, demography, gender, human resources, pay equity, pay gap, race

1 | INTRODUCTION

One of the most widely researched, discussed, and litigated inequities in economic life is the gender pay gap: the empirical regularity that women are frequently paid less than men, even after accounting for differences in job responsibilities and qualifications (Blau & Kahn, 2017). In recent years, this attention has translated into increasing pressure on firms to achieve pay equity. U.S. states like Colorado, Massachusetts, New Hampshire, New Jersey, and New York have strengthened their equal pay legislation. Europe has seen similar activity: the UK requires any firm with more than 250 employees to report its gender pay gap (Gov.uk Collections, 2020); France has legislated stiff penalties for firms scoring low on an equality scale in which pay equity figures prominently (Society for Human Resource Management, 2019); and Iceland recently became the first country to require firms to obtain equal pay verification (Domonoske, 2018). Most recently, the European Parliament put forward a pay transparency directive that combines many transparency, reporting, and best practice measures into a single legislative framework (European Commission, 2021). Pay gaps with respect to other demographic characteristics, for example, race (Lang & Lehmann, 2012), have also received attention. Meanwhile, employers are increasingly embracing pay equity as part of their cultural values and as a way to enhance their reputation. For instance, Salesforce disclosed that it spent millions of dollars in 2015 and 2016 to reach gender pay parity (Horowitz, 2017; McGregor, 2015).

There is a profusion of research on the empirical drivers of the gender pay gap (e.g., Blau & Kahn, 2017, 2000; Weichselbaumer & Winter-Ebmer, 2005) and on how to establish more gender-equitable recruitment and promotion policies (Foley et al., 2019), but surprisingly little research guides firms in detecting a gender pay gap and making compensation adjustments to close it. To our knowledge, there are only three papers that directly consider the cost of correcting a gender pay gap (Anderson et al., 2019; Becker & Goodman, 1991; Becker & Toutkoushian, 1995). Both papers by Becker and co-authors use data from a class action lawsuit to compare the increase in a university's total wages from different equalization methods (e.g., giving women a percentage increase vs. paying them the salary they would be predicted to receive if they were male), but they do not offer a general solution.

Accepted by Sean Willems, after three revisions.

Anderson et al. (2019) adopt the theoretical conceit that firms will seek to close their gender pay gap at minimum cost and use this conceit to predict which employees are likely to get raises. However, their cost minimization approach (CMA) has been uniformly and emphatically rejected by the over 40 employers we have worked with because, as we discuss below, it severely distorts compensation structures in terms of both fairness and incentivization, and it may even expose an employer to legal liability given its apparent neglect of fairness considerations. In addition, none of the foregoing work discusses the many practical implementation issues that arise in allocating raises to close a pay gap (e.g., how to determine whether a gap is large enough to require closing or how to measure statistical significance in this context).

This lacuna in our theoretical and practical understanding has important consequences. In our fieldwork with firms ranging from 75 to 130,000 employees, managers have consistently told us that they had a keen desire to address pay inequity in their compensation structure but were frustrated by a lack of tools. They did not have a rigorous way of measuring a pay gap or of allocating raises to close it, nor did they understand how closing a pay gap would affect their other compensation goals, how much it would cost, and whether it would really address pay inequity. They were therefore left to choose between various ad hoc approaches, such as giving across-the-board raises to all female employees (Loriggio, 2016), which has been especially common at universities (e.g., University of Alberta, University of British Columbia, and McMaster), or asking supervisors to propose raises for women who appear to be underpaid. These approaches have significant shortcomings: they may not actually close the gender pay gap, they may allocate raises in a nonsystematic way that is difficult to justify to regulators or courts, they may fail to respect other compensation goals, and they may have a high cost. This paper aims to fill this theoretical and practical gap in our knowledge by presenting an optimization approach to address pay inequity.

To develop our approach, we followed the groundbreaking work of Holmstrom et al. (2009) on "ill-structured decision situations." In particular, managers generally understand the overall goal of (a) achieving pay equity while (b) respecting other compensation priorities, but they lack a precise definition of what this twofold goal means, to say nothing of knowing how to achieve it. Specific challenges include properly filtering out the effect of legitimate pay drivers when calculating the gender pay gap, rigorously operationalizing a firm's own sense of fairness and compensation goals, and developing a replicable method for optimally respecting these goals while closing the gender pay gap and ensuring pay equity problems do not reappear. As recommended by Holmstrom et al., we adopted a design sciences approach wherein we first applied the operations management toolkit to structure the problem before developing a solution in the form of a systematic, straightforward process that managers and their advisors may use to address pay equity. This process includes a software application that guides employers through every step of the pay gap elimination lifecycle, from data cleaning and visualization, to statistical modeling, to raise allocation to close the gap. Our approach thereby dovetails with the increasingly common strategic goal of maximizing the efficiency and equity of human resources (HR) policies through quantitative, data-driven techniques (Biro, n.d; Viser, 2018) and thereby establish scalable HR policies that do not overrely on the judgment of individual supervisors (Choi, 2019; Davies, 2019).

In line with Holmstrom et al. (2009), we emphasize that solving a particular field problem is only part of our contribution. The broader research question that we seek to address is how the statistical and optimization toolkit of operations management can contribute to HR management and, more generally, help managers close pay gaps in a systematic way. Thus, our paper contributes to an emergent stream of literature on people analytics. Boudreau et al. (2003) discuss the vital role of human factors in operations management and the impact of resources and constraints on traditional HR functions, calling attention to the potential value of research at the interface of operations and HR for both domains. Operations management scholars have looked at how to optimize performance through team composition (Huckman & Staats, 2011), employee training (Hopp & Oyen, 2004), and incentive schemes (Siemsen et al., 2007). There has also been significant research into workforce planning, employee engagement (Yee et al., 2008, 2010), and how to estimate future load and employee churn (see Cotton & Tuttle, 1986, and Hom et al., 2017, for reviews of the literature on employee turnover prediction and theory). However, many other areas, including optimal compensation, evaluation, and recruitment strategies, remain relatively understudied in the operations management literature. We contribute by studying how to close a gender (or other demographic) pay gap using quantitative, data-driven methods.

Our paper also contributes to the operations literature on the trade-off between fairness and efficiency, which has analyzed diverse operational settings (see, e.g., Bertsimas et al., 2013; Wagstaff, 1991; Woessmann, 2008). Across disciplines, researchers find that a narrow focus on either efficiency or fairness comes at significant cost to the other, but when decision-makers value both, there is a Pareto frontier where outcomes can be both equitable and efficient (Okun, 2015), for example, in organ sharing (Benvenuto et al., 2018; Mooney et al., 2019). In the context of closing pay gaps, incorporating efficiency (which in this case translates directly to compensation costs) can distort employee incentives and negatively impact the overall salary structure, exposing employers to legal and regulatory risk; our approach, which incorporates fairness and an employer's other compensation priorities into the raise allocation process, addresses these potential problems.

Finally, a long research tradition documents that women and men are frequently rewarded differently in the labor market for personal, performance, and job attributes, as demonstrated by running separate wage regressions for women and men in the same labor pool (Blau & Kahn, 2003; Harmon, Oosterbeek, & Walker, 2003; Sorenson & Dahl, 2016). This is the basis for the well-known Blinder–Oaxaca decomposition (Blinder, 1973; Oaxaca & Ransom, 1999) of a gender pay gap into (a) differences in how compensation factors (e.g., job role, human capital) are distributed between women and men and (b) differences in the rewards women and men receive for these factors (Becker & Toutkoushian, 1995; Toutkoushian & Hoffman, 2002). Blinder–Oaxaca has been widely applied in the social sciences to study the gender pay gap but is not used for regulatory purposes because it is not invariant to affine transformations and is thus easy to manipulate (Oaxaca & Ransom, 1999). We build on this insight but use it prescriptively rather than diagnostically: in the spirit of the Blinder–Oaxaca analysis, we use a regression that interacts gender with job, performance, and personal characteristics to determine the degree to which women and men are rewarded differently for these compensation factors in a firm's workforce. We then minimize this difference, subject to the firm's budget constraints and target pay gap (usually zero), all while rewarding employees for their performance, qualifications, and job in accordance with the firm's HR policies.

Most of this paper is organized around a case study of one of the first employers we worked with. To refine this approach, we supplement the narrative with discussion of issues that arose with other employers or in our analytical work. Thus, our discussion follows the steps we took to implement our approach and yields a process for others to follow.

2 | CASE STUDY

Our story begins in 2016 with our first visit to NordCo, a small European company with 471 employees, 70% of whom were male. Overall, women were paid 11.0% less than men. However, the unadjusted gender pay gap is often large partly because of gender imbalances across job roles and seniority levels. The company had been striving for gender equity writ large since 2013; for example, they balanced their executive team between men and women. NordCo wanted us to help them determine whether they had an adjusted pay gap, also called an "equal pay gap" (i.e., a pay gap after accounting for other pay drivers), which is the main focus of equal pay legislation. If they did have an adjusted pay gap, they wanted help eradicating it.

2.1 | Measurement

Measuring an equal pay gap is part structured and part unstructured. The wage equation itself is structured: academics, courts, and regulators around the world have established a standard methodology. As explained in detail in Becker and Toutkoushian (1995) and Anderson et al. (2019) and the references therein, this standard methodology for measuring pay discrimination in an organization, whether done internally or by expert witnesses in a lawsuit, is the log-linear regression, wherein one regresses the natural log of wages on an indicator variable for the demographic characteristic of interest, usually gender. Importantly, this regression controls for observable characteristics such as job role, performance, education, and experience and thus differs from the unadjusted pay gap obtained from comparing average pay for women and average pay for men. Although the unadjusted pay gap is often the one reported by the media and although there could in principle be other methods for calculating an equal pay gap, we base our methods on the log-linear regression because it underlies the standard to which employers are held.

An unstructured and potentially contentious part of measuring an equal pay gap is determining what observable characteristics should be included in the log-linear regression and how they should be operationalized as variables. In principle, the wage regression should capture the value and nature of work as well as relevant qualifications. It is therefore important to determine what jobs are substantially similar. This can be done either by grouping similar jobs together, by rating each job on key factors, or by using a point schema that assigns a numerical value to each job. Since an employer may need to defend its choice of values to employees, regulators, or even courts, it is critical to maintain transparent, consistent standards that accurately reflect the value, importance, and demands of the work.

To do this, the firm first needs to collect employee data reflecting important determinants of pay. Although NordCo had this data to hand, this is an important first step for some organizations. Many of these data may also have multiple correlated elements, requiring experimentation and debate to determine which elements to discard and which to include.

After multiple iterations, which is common, NordCo chose to use job roles to capture the fundamentals of equal work. Using job roles usually begins with a broad overview, then jobs that contain similar tasks and responsibilities are combined into groups. Each job role is then studied to determine whether there are individual employees in that role whose tasks are significantly different from those of others in the same role. If so, the job role may need to be split in two (e.g., junior and senior project managers). NordCo settled on 18 job roles, created by an analysis of 37 factors applicable to every job (including responsibilities, skills, and work environment). NordCo also chose to include performance, growth potential (a measure of upward mobility in the firm), shift versus overtime work, and two unique and confidential elements: Important Factor A is set of job-related factors, and Important Factor B measures exposure to a stressor unique to NordCo.

Every employer needs its own set of pay drivers, as determined by its unique circumstances and pay philosophy. Moreover, circumstances change. NordCo evaluates its pay structure every year, and after it defined its job roles, it later combined many of them to simplify the system and better align it with the company's operations.

With a thoughtful set of job-related and personal characteristics in hand, we calculated an initial log-linear wage regression as follows. Formally, let X be an $n \times k$ matrix representing n employees and their k job-related and personal characteristics, including a constant and a gender dummy that we index by F. If employee i is female, then $x_{iF} = 1$, and 0 otherwise. Let $Y = \ln W$ be a column vector, where w_i is the wages of employee i. The standard definition of the equal, or adjusted, pay gap relates the gender regression coefficient from the following regression model to pay discrimination:

$$Y = \beta X + \epsilon, \tag{1}$$

where ϵ is a column vector of residuals. By standard calculations, we have

$$\begin{aligned} \beta &= (X'X)^{-1}X'Y = CY \\ \beta_F &= (X'X)^{-1}x_F'Y = c_F Y' \end{aligned} \tag{2}$$

where the constant C is the Moore–Penrose (pseudo) inverse of X and c_F is the row vector that yields β_F, the adjusted gender pay gap. The element of c_F associated with employee i is denoted c_{iF}. Although it is common to think of a coefficient in a log-linear regression such as β_F as referring to percentage changes, the correct formula for how β_F influences the conditional expectation of the untransformed wages is actually $e^{\beta_F}e^{\frac{\sigma^2}{2}}$, where σ^2 can be estimated by the mean square error of the regression (Duan, 1983). We report untransformed coefficients throughout.

Another unstructured and potentially contentious part of measuring the pay gap is deciding how to interpret the magnitude and statistical significance of β_F and the other coefficients. Although sophisticated audiences and regulators may focus on whether β_F is statistically significant at conventional levels (e.g., 5%), most managers we have worked with disregard statistical significance, instead focusing on "practical significance": whether the coefficient would be problematic in the eyes of internal and external stakeholders. This perspective can be justified from a theoretical point of view if an employer's workforce is taken as the entire population under study. Applying the finite population correction, standard errors converge to zero as the sample size approaches the entire population (Abadie et al., 2020; Ramachandran & Tsokos, 2009). The p-value of every coefficient in the wage regression is then by definition 0. (This view ignores "design-based uncertainty" if one regards, say, an employee having one job instead of another as a "treatment" (Abadie et al., 2020), but regulators or courts do not generally consider design-based uncertainty.) Regardless, to be broadly applicable, our methods must be able to accommodate different perspectives on whether β_F or other coefficients are large enough to matter and whether their statistical significance is relevant. We will return to this issue, especially in Sections 2.6 and 2.9.

We now present NordCo's wage regression, with some modifications to preserve confidentiality: we normalize salaries to 1, suppress the constant in reporting the results, and simplify the regression model to the following:

$$\begin{aligned} \ln(Salary) = {} & \beta_0 + \beta_1 ImportantFactorB + \beta_2 GrowthPotential \\ & + \beta_3 WorkingHours + \sum_j \beta_{4j} Job_j \\ & + \sum_k \beta_{5k} Performance_k \\ & + \sum_l \beta_{6l} ImportantFactorA_l + \beta_f Female + \epsilon. \end{aligned} \tag{3}$$

The model estimates are provided in the first columns of Table 1. The coefficient for gender, β_F, is -0.013, indicating that after controlling for job category, Important Factors A and B, performance, and growth potential scores, women had log-salaries 0.013 lower than men with similar job-related and personal characteristics. NordCo's pay gap was much lower than we typically see and reflected the company's earlier efforts to achieve gender equity. The gender coefficient is not statistically significant at conventional levels ($p = 0.315$), but, as is typical, the practicing managers at NordCo believed that it reflected a bias in pay structure and thus still wanted to close the gap.

2.2 | Cost minimization approach

Having measured the pay gap, the company's goal was to eliminate it, that is, to bring the gender coefficient to zero. (We discuss budget constraints and options for choosing a target pay gap below.) We started by proposing the algorithm developed by Anderson et al. (2019), that is, the CMA. The HR manager in charge of the project opposed this on principle, arguing that the "equity" in pay equity should be respected and that employees who received raises should deserve them. We thus ran the algorithm only as a benchmark. One immediate problem was that some of the suggested raises were astronomical, as high as 31.7%, levels difficult to explain to the firm's stakeholders, including employees.

We then tried modifying the CMA to cap raises at 10%. Formally, let raises given to employees be denoted by δ_i, then the modified CMA can be expressed as

$$\min_{\delta_i} \quad \sum_i \delta_i, \tag{4a}$$

$$\text{subject to} \quad \beta_F = \sum_i c_{iF} \ln(w_i + \delta_i) \geq 0, \tag{4b}$$

$$\delta_i / w_i \leq 0.1, \tag{4c}$$

$$\delta_i \geq 0 \quad \forall i. \tag{4d}$$

TABLE 1 Regression results for the model used to measure the equal pay gap. The constant term is suppressed for anonymity.

	Original		After CMA		After SA	
Coefficient	**Estimate**	**Pr(>|t|)**	**Estimate**	**Pr(>|t|)**	**Estimate**	**Pr(>|t|)**
Gender	−0.013	0.3153	0	0.9713	0	0.9999
Important Factor B	0.1079	0.0000	0.104	0.0000	0.1166	0.0000
Growth Potential	0.0018	0.8413	0.002	0.7971	0.0049	0.5908
Working Hours	0.0012	0.0000	0.001	0.0000	0.0012	0.0000
Job Cat 1	1.121	0.0000	1.128	0.0000	1.122	0.0000
Job Cat 2	0.1117	0.0197	0.114	0.0172	0.1148	0.0155
Job Cat 3	0.2702	0.0000	0.284	0.0000	0.2853	0.0000
Job Cat 4	1.015	0.0000	1.017	0.0000	1.009	0.0000
Job Cat 5	0.402	0.0000	0.413	0.0000	0.4106	0.0000
Job Cat 6	0.3821	0.0000	0.394	0.0000	0.3991	0.0000
Job Cat 7	0.5006	0.0000	0.511	0.0000	0.5084	0.0000
Job Cat 8	0.1454	0.0005	0.176	0.0000	0.1555	0.0002
Job Cat 10	0.4392	0.0000	0.444	0.0000	0.437	0.0000
Job Cat 11	0.2661	0.0000	0.267	0.0000	0.2658	0.0000
Job Cat 12	0.1951	0.0000	0.196	0.0000	0.1969	0.0000
Job Cat 13	0.9946	0.0000	1.002	0.0000	0.9936	0.0000
Job Cat 14	0.4058	0.0000	0.419	0.0000	0.4012	0.0000
Job Cat 15	0.1371	0.0015	0.137	0.0014	0.1531	0.0004
Job Cat 16	0.4134	0.0000	0.424	0.0000	0.4195	0.0000
Job Cat 17	0.9802	0.0000	0.979	0.0000	0.9708	0.0000
Performance Cat 1	0.0445	0.0992	0.032	0.2394	0.0321	0.2304
Performance Cat 2	0.0585	0.0002	0.045	0.0046	0.0479	0.0024
Performance Cat 3	0.1129	0.0000	0.1	0.0000	0.1069	0.0000
Important Factor A Cat 1	−0.295	0.0000	−0.295	0.0000	−0.2972	0.0000
Important Factor A Cat 2	−0.1497	0.0000	−0.151	0.0000	−0.1545	0.0000
Important Factor A Cat 3	−0.2488	0.0000	−0.248	0.0000	−0.2217	0.0000
Important Factor A Cat 4	−0.0389	0.3228	−0.037	0.3459	−0.0315	0.4188
Important Factor A Cat 5	0.0884	0.0240	0.093	0.0174	0.1032	0.0079
Important Factor A Cat 7	−0.2504	0.0019	−0.246	0.0021	−0.2458	0.0021
R-squared	0.905		0.904		0.9064	
Adjusted R-squared	0.8987		0.8977		0.9002	
MSE	0.00896		0.00889		0.00879	

Note: The base category for Job Categories is Category 9; for Performance Categories, Category 0; for Important Factor A, Category 8. Important Factor Category 6 is dropped for collinearity.

where (4d) reflects an assumption the firm will not lower salaries, (4a) is the sum of the wage increases (which the CMA minimizes), (4b) ensures we reach the target pay gap (zero in this case), and (4c) constrains individual raises to a maximum of 10%. We note that w_i, the starting salary, is constant for each employee, making (4c) a linear constraint and that the sign of β_F is negative when the pay gap is in favor of men. As the objective is to minimize cost, the $\beta_F \geq 0$ constraint will be binding at the optimal solution. In this case, the optimal solution resulted in 12 raises that altogether added 0.133% to the total payroll cost. The raises were to be concentrated in a heavily male, low-paid job role, with eight of the ten women in that role receiving raises. This makes sense because it is most cost effective to allocate raises to (a) low-wage employees (usually women) due to the concavity of the log-transformation of wages in the log-linear regression and (b) women who resemble men in terms of job-related and personal attributes because paying these women more increases the explanatory power of these attributes in the wage regression, thereby reducing the explanatory power of gender. We also note that seven raises were exactly at the 10% constraint, six in Job Category 8 and one in Job Category 5.

While these recommendations would close the pay gap at a relatively low cost, NordCo was not satisfied. Table 1, which shows the regression coefficients after the CMA raises, highlights some of their concerns. First, the relationship between performance and pay has been weakened by applying the CMA. The coefficients for performance above the baseline category 0 have all decreased, indicating that if the company were to follow the CMA and adhere to the resulting pay model in future salary decisions, they would reduce the incremental rewards for better performance. Second, NordCo did not like the distortion to the overall pay structure. For example, the coefficient on Job Category 8 increased by 17.4%, from 0.145 to 0.176.

The results only strengthened NordCo's resolve to address any inequities in their compensation structure rather than "make the problem go away" at minimum cost. In particular, giving raises predominately to low-wage women as determined by the CMA would have had three consequences: (a) the firm's wage structure would be compressed, reducing the ratio between high- and low-paid employees and compromising incentives; (b) the raises would not necessarily have gone to women who appeared underpaid in the wage regression, that is, the CMA sometimes overpays some women to compensate for underpaying others; and (c) as one U.S. lawyer later told us: "I'd love to get my hands on a document in [legal] discovery showing that my client's employer decided how to allocate raises so as to achieve a zero pay gap at the lowest cost; that violates the intent of equal pay legislation." In short, the theoretical conceit underlying Anderson et al. (2019) that employers would seek to close their pay gaps strictly at minimum cost is simply not correct in practice.

2.3 | Existing alternatives to the CMA

We then discussed alternatives to the CMA. The most obvious was an across-the-board log-raise of 0.013 to every woman, equal in magnitude to the gender coefficient. This would have added 0.356% to the payroll cost, considerably more than the CMA. Moreover, NordCo's managers objected that it would leave some women underpaid and might overpay others. In other words, it would represent a missed opportunity in correcting underlying biases.

Another option we investigated was to run the wage regression without the gender indicator and then give a raise to every woman who had a negative residual, that is, who appeared underpaid. This method is similar to what many organizations have historically done when conducting internal pay reviews. But this method has problems, too. It compresses the pay structure of women relative to that of men (whose wages are not adjusted), does not necessarily close the pay gap to a specific target (zero, in this case), and can result in raises being given to the lowest paid women (who thus have negative residuals) in job roles where, on average, women are paid more than men, which can magnify existing inequalities. The implication was clear: existing methods for systematically closing pay gaps in a workforce simply were not practical or effective. We needed a new approach.

2.4 | Structured approach

We, therefore, proposed the structured approach (SA). The SA starts from the following premise: not only should women and men be paid the same across the salary structure (as measured by the gender indicator in the wage regression), but gender equity in pay should also hold across different parts of the organization and with respect to different pay drivers. In other words, in wage regressions for women and for men (as in the Blinder–Oaxaca decomposition), the intercepts and the slopes of the regression lines with respect to job-related and personal qualifications should be roughly the same. This is not only fair, but it also respects the firm's compensation policies, for example, that pay increases as desired in response to job-related and personal qualifications.

This is the case because rather than running separate wage regressions, the SA uses the statistically equivalent formulation of augmenting the log-linear wage regression with a full set of interaction terms between each job-related and personal qualification and the female gender dummy. In effect, the SA measures how much each characteristic contributes to men's pay (estimated by the coefficient on the variable for that characteristic) and how much more or less women are rewarded for each characteristic (estimated by the coefficient on the interaction of that characteristic with gender). We can then see where the pay gap comes from. In our work with employers, we have often found specific job roles or departments with very large pay gaps and others with very small or no pay gaps. We also commonly see that all else equal, women do not receive the same rewards in terms of pay for additional education, experience, or performance ratings as men do.

Importantly, the SA still measures and closes the gender pay gap using the regression specified in Equation (3). Thus, it effectively uses a combination of two wage regressions: the base log-linear regression to measure the size of the gap, plus a regression with interaction terms (denoted as BO given its link to the Blinder–Oaxaca decomposition) to support the raise allocation, as described below.

Formally, let $X_{-1,-F}$ be the original X matrix excluding the constant and gender column (x_F). Then, let X_{BO} be an $n \times (k-2)$ matrix defined as $X_{BO} = x_F' X_{-1,-F}$. The resulting elements of the matrix of interactions X_{BO} are zero for all male employees and have the same values as X for all female employees. We will reference a row in X_{BO} corresponding to employee i as x_{BOi}. Let $Y = \ln W$ be a column vector, where w_i are the wages of employee i, as previously defined. The regression formula that measures factor-specific bias (i.e., whether women are rewarded differently than men for job-related and personal characteristics) can then be expressed as

$$
\begin{aligned}
Y &= \beta X + \beta_{BO} x_F' X_{-1,-F} + \epsilon \\
Y &= \beta X + \beta_{BO} X_{BO} + \epsilon,
\end{aligned}
\tag{5}
$$

where ϵ is a column vector of residuals. The elements of β_{BO} represent the difference between the contribution of each characteristic to female employees' wages and its contribution to male employees' wages. In other words, β_{BO} represents the magnitude by which women get rewarded more or less than men for job-related and personal characteristics included in the regression model. Therefore, for a female employee i, $\beta_{BO}x_{BOi}$, represents the difference in pay due to specific inequities (e.g., the degree to which she is underpaid because she is rewarded less than a man for possessing a certain job or human capital attribute). For a female employee i who is underpaid, $\beta_{BO}x_{BOi}$ will be negative, so multiplying by -1 gives the estimate of the overall factor-specific inequity she faces. Thus, we define $\Gamma_i = -1 \cdot \beta_{BO}x_{BOi}$.

For NordCo, the regression with interactions is

$$\begin{aligned}\ln(Salary) = {} & \beta_0 + \beta_1 Female + \beta_2 ImportantFactorB \\ & + \beta_3 GrowthPotential + \beta_4 WorkingHours \\ & + \sum_j \beta_{5j} JobCategory_j + \sum_k \beta_{6k} Performance_k \\ & + \sum_l \beta_{7l} ImportantFactorA_l \\ & + \beta_8 ImportantFactorB \cdot Female \\ & + \beta_9 GrowthPotential \cdot Female \\ & + \beta_{10} WorkingHours \cdot Female \\ & + \sum_j \beta_{11j} JobCategory_j \cdot Female \\ & + \sum_k \beta_{12k} Performance_k \cdot Female \\ & + \sum_l \beta_{13l} ImportantFactorA_l \cdot Female + \epsilon. \quad (6)\end{aligned}$$

The regression results, in Table 2, show significant heterogeneity in how job-related and personal characteristics (factors) influence pay for women and men. Women with Important Factor B receive a much smaller increase in salary for that additional stressor than men do, as shown by the coefficient on the associated interaction of −0.165. Similarly, each additional work hour per month increases a woman's salary by less than it does for men (coefficient of −0.00118 on the interaction between gender and work hours). The difference in total work hours between the 25th and 75th percentiles of NordCo's workforce was 39.1 h. All else equal, a man who works an extra 39.1 h per month would see his log-wage increase by an average of $39.1 \cdot 0.001216 = 0.0475$. A similar woman would see her log-wage increase by an average of only $39.1 \cdot (0.001216\text{-}0.00118) = 0.0014$ for the additional hours worked. In addition, the pay gap is relatively large in certain job roles, particularly Job Categories 2, 12, and 15, as indicated by large negative interaction terms between the job dummy and the gender variable. In other job categories, the interaction is small or even positive, indicating that the pay gap favored women, as in Categories 4, 5, 10, 11, and 16. (We address the statistical significance of these interactions below.)

NordCo's goal, in line with the SA, was to close their pay gap while reducing the factor-specific differences in pay represented by the interaction terms. To do this, the SA prioritizes raises for employees who, based on their pay-related characteristics, are expected to experience the greatest inequity. Specifically, we raise the salaries of employees whose Γ_i is the largest until the target value T_F for β_F (from the base log-linear wage regression) is reached. In this formulation, we limit the raises to female employees (or, more generally, to employees of the underpaid gender), denoting the set of female employees by F. This approach to reducing the pay gap can be expressed as a minimax optimization problem, minimizing Γ, which we define as the maximum factor-specific inequity that any employee faces (i.e., $\Gamma = \max_i(\Gamma_i) = \max_i(-\beta_{BO}x_{BOi})$) as follows:

$$\min \quad \Gamma, \tag{7a}$$

$$\text{subject to} \quad \beta_F = \sum_i c_{iF}(\ln(w_i) + \delta_i') = T_F, \tag{7b}$$

$$\Gamma + \beta_{BO}x_{BOi} - \theta_i = 0 \quad \forall i \in F, \tag{7c}$$

$$y_i^+ \geq \theta_i / M_1 \quad \forall i \in F, \tag{7d}$$

$$y_i^- \geq -\theta_i / M_1 \quad \forall i \in F, \tag{7e}$$

$$y_i^- + y_i^+ \leq 1 \quad \forall i \in F, \tag{7f}$$

$$\delta_i' \leq y_i^- M_2 \quad \forall i \in F, \tag{7g}$$

$$\delta_i' + (\Gamma + \beta_{BO}x_{BOi}) \geq 0 \quad \forall i \in F, \tag{7h}$$

$$\delta_i' \leq -(\Gamma + \beta_{BO}x_{BOi}) + y_i^+ M_3 \quad \forall i \in F, \tag{7i}$$

$$\delta_i' = 0 \quad \forall i \notin F, \tag{7j}$$

$$\delta_i' \geq 0 \quad \forall i, \tag{7k}$$

$$y_i^+, y_i^- \in \{0, 1\} \quad \forall i, \tag{7l}$$

where Γ, δ_i', y_i^+, y_i^-, and θ_i are all decision variables. Γ is the objective, the maximum factor-specific inequity faced by any female employee, and δ_i' is the *log-scale* raise given to employee i (we note this differs from the CMA, where δ indicates a dollar-scale raise). Note that β_F is calculated using

TABLE 2 Regression results from the SA

Coefficient	Estimate	Std. Error	t value	Pr(>|t|)
Gender	0.1501	0.1314	1.143	0.2538
Important Factor B	0.1239	0.020	6.1930	0.0000
Growth Potential	0.0087	0.0105	0.8350	0.4042
Working Hours	0.0012	0.0001	12.4420	0.0000
Job Cat 1	1.2230	0.1202	10.179	0.0000
Job Cat 2	0.2279	0.1331	1.713	0.0875
Job Cat 3	0.3882	0.1208	3.214	0.0014
Job Cat 4	1.0890	0.1253	8.689	0.0000
Job Cat 5	0.5071	0.1127	4.498	0.0000
Job Cat 6	0.4997	0.1180	4.235	0.0000
Job Cat 7	0.6021	0.1203	5.004	0.0000
Job Cat 8	0.2510	0.1146	2.190	0.0291
Job Cat 10	0.4944	0.1230	4.020	0.0001
Job Cat 11	0.3107	0.1174	2.646	0.0085
Job Cat 12	0.2829	0.1329	2.129	0.0338
Job Cat 13	1.0830	0.1555	6.966	0.0000
Job Cat 14	0.4768	0.1193	3.997	0.0001
Job Cat 15	0.2775	0.1263	2.198	0.0285
Job Cat 16	0.5140	0.1151	4.464	0.0000
Job Cat 17	1.0610	0.1000	10.613	0.0000
Performance Cat 1	0.0123	0.0302	0.407	0.6845
Performance Cat 2	0.0330	0.0186	1.777	0.0763
Performance Cat 3	0.0961	0.0218	4.401	0.0000
Important Factor A Cat 1	−0.2985	0.0269	−11.098	0.0000
Important Factor A Cat 2	−0.1580	0.0288	−5.486	0.0000
Important Factor A Cat 3	−0.1624	0.0670	−2.425	0.0158
Important Factor A Cat 4	0.0079	0.0572	0.138	0.8905
Important Factor A Cat 5	0.1470	0.0562	2.617	0.0092
Gender · Important Factor B	−0.1650	0.071	−2.325	0.0206
Gender · Working Hours	−0.0012	0.001	−1.960	0.0507
Gender · Growth Potential	−0.0353	0.022	−1.597	0.1109
Gender · Perf Cat 1	0.139	0.074	1.882	0.0605
Gender · Perf Cat 2	0.104	0.038	2.744	0.0063
Gender · Perf Cat 3	0.072	0.047	1.533	0.1261
Gender · Job Cat 1	−0.019	0.120	−0.161	0.8720
Gender · Job Cat 2	−0.073	0.134	−0.547	0.5846
Gender · Job Cat 4	0.018	0.096	0.191	0.8490
Gender · Job Cat 5	0.065	0.129	0.507	0.6123
Gender · Job Cat 6	−0.035	0.135	−0.260	0.7952
Gender · Job Cat 7	−0.007	0.148	−0.044	0.9648
Gender · Job Cat 8	−0.054	0.112	−0.482	0.6301
Gender · Job Cat 10	0.040	0.113	0.358	0.7203
Gender · Job Cat 11	0.037	0.109	0.335	0.7377
Gender · Job Cat 12	−0.084	0.128	−0.653	0.5141
Gender · Job Cat 15	−0.139	0.121	−1.149	0.2513
Gender · Job Cat 16	0.111	0.158	0.699	0.4849

(Continues)

TABLE 2 (Continued)

Coefficient	Estimate	Std. Error	t value	Pr(>|t|)
Gender · Important Factor A Cat 1	0.041	0.059	0.693	0.4886
Gender · Important Factor A Cat 2	0.028	0.065	0.434	0.6642
Gender · Important Factor A Cat 3	−0.151	0.090	−1.671	0.0955
Gender · Important Factor A Cat 4	−0.076	0.080	−0.943	0.3462
Gender · Important Factor A Cat 5	−0.113	0.083	−1.371	0.1711

The base category for Job Categories is Category 9; for Performance Categories, Category 0; for Important Factor A, Category 8. Important Factor A Categories 6 and 7 and the interactions between Gender and Job Category 17 are dropped for collinearity.
R-squared: 0.911, Adjusted R-squared: 0.8999, MSE: 0.00839.

Equation (3), while β_{BO} is calculated using Equation (6). M_1, M_2, and M_3 are constants bounded by $\max_i |\beta_{BO} x_{BOi}|$, which is the maximum possible factor-specific inequity (and therefore relatively small). Most firms simply want to set T_F to zero, aiming to completely eliminate the pay gap. Since NordCo had begun pay equity audits years before, they had a relatively small pay gap at the beginning of our collaboration and were able to completely close it within a year. However, other companies may have an initial gap that is too large to close at once given budget constraints, or they may be satisfied with a small pay gap. In these cases, T_F can be set to an intermediate value or to a certain tolerance level, as we will discuss later.

In the formulation, constraint (7b) calculates β_F as previously described (using the original log-linear regression without interaction terms) and restricts the value to be equal to the target value, T_F. The term $\beta_{BO} x_{BOi}$ calculates the difference in pay for employee i due to gender differences in how characteristics are rewarded. For each employee, the constraint (7c) captures the difference between the maximum bias level, Γ, and that employee's factor-specific inequity; θ_i reflects this difference. In other words, constraint (7c) ensures that θ_i measures the initial amount of factor-specific inequity beyond Γ faced by employee i. Constraints (7d)–(7f) translate each θ_i into binary indicators y_i^+ and y_i^-. If employee i experiences more factor-specific inequity than Γ, θ_i is negative, and constraint (7e) ensures that $y_i^- = 1$. Conversely, when employee i faces less factor-specific inequity than Γ, θ_i is positive and constraint (7d) ensures that $y_i^+ = 1$. Constraint (7f) ensures that either y_i^+ is positive or y_i^- is positive, but not both. Constraint (7g) sets the pay raises, δ_i', to zero for everyone whose θ_i is positive (i.e., for everyone who experiences less factor-specific inequity than Γ). Constraint (7h) ensures that the raises are large enough to reduce factor-specific inequity to at least the threshold Γ, and constraint (7i) restricts the raises so as not to exceed this threshold. Effectively, constraint (7g) ensures that the raise for each employee is 0 if $y_i^+ = 1$, and constraints (7h) and (7i) set the raise δ_i' equal to θ_i if $y_i^+ = 0$, thus bringing the effective factor-specific inequity faced by employee i down to exactly Γ. This constraint is needed to ensure that the minimum cost solution that achieves Γ is selected.

While the mathematical expression of the optimization formulation is complex, the optimal solution can be characterized and found with a binary search for the smallest Γ that satisfies the constraints, denoted as Γ^*. An arbitrarily large Γ' is selected, such that when raises are awarded, β_F will overshoot the target T_F. All employees (of the underpaid gender) with $\beta_{BO} x_{BOi} > \Gamma'$ are given raises of $\beta_{BO} x_{BOi} - \Gamma'$, and β_F is recalculated. If $\beta_F > T_F$ (which will happen in the first iteration), then Γ' is lowered, and the process is repeated. If $\beta_F < T_F$, then Γ' is increased, and the process is repeated. With each iteration, the size of the change in Γ' is decreased by 50%. The process terminates when $\beta_F = T_F$ within some tolerance limit. The Γ' from the final iteration is optimal, and it becomes Γ^*. This Γ^* is the smallest Γ, which satisfies the constraints given in the optimization formulation.

In effect, given a value for Γ, the constraints of the formulation ensure that every woman's raise is set to $\max(0, \Gamma_i - \Gamma)$. The objective is to find the minimum Γ which will satisfy the constraint that β_F equals the target pay gap. This ensures we choose the smallest set of raises necessary to close the pay gap, thus minimizing cost while prioritizing raises based on each employee's factor-specific inequity.

Firms typically desire additional constraints to reflect other compensation goals and limitations. For instance, a limit on any individual raise is often incorporated, $\delta_i' \le \delta_i'^{max}$; $\delta_i'^{max}$ could be based on job roles. Or, in the case of a partially unionized workforce, $\delta_i'^{max}$ can be set to zero for employees for whom salary raises are not feasible or desirable. Absent such considerations, $\delta_i'^{max}$ can be set as the same fixed max log-scale raise δ'^{max} for everyone, which translates to the same maximum percent raise for each employee. In addition, organizations typically determine salary ranges for specific ranks or job roles. An employee's placement within the salary range may be a function of deterministic attributes such as educational level or of more subjective factors such as performance. Let l_j and u_j represent the lower and upper bounds of the salary range for job role j. To limit each employee's salary to a specified range, we restrict δ_i' such that $\exp(\ln(w_i) + \delta_i')) \in [l_j, u_j]$ if employee i has role j. Wage ranges may be set internally (e.g., accounting managers' salaries should fall within some upper and lower bounds $l_{\text{accounting manager}}, u_{\text{accounting manager}}$), by union contracts or by regulation. The lower bound constraints $\exp(\ln(w_i) + \delta_i') \geq l_j$ are redundant if the organization has been following their own internal salary ranges.

To implement these additional constraints, the constraints that require $\Gamma_i \geq \Gamma^*$ need to be relaxed (please refer to the Supporting Information for the expanded formulation). However, the optimal solution can still be characterized and found through binary search for Γ^*. In this expanded formulation, each employee at the optimal solution, $\delta_i' = \delta_i'^{max}$ if $\Gamma_i + \delta_i'^{max} \leq \Gamma^*$, $\delta_i' = 0$ if $\Gamma_i \geq \Gamma^*$, and $\delta_i' = \Gamma^* - \Gamma_i$ for all other employees.

We finally note that if there exists an employee i in group j such that $w_i \geq u_j$, then the optimization problem incorporating these additional salary range constraints is infeasible. To account for such a situation, the formulation can be relaxed by introducing additional variables, ensuring $\delta_i' = 0$ for employees currently exceeding the upper limit.

It is also possible for the SA to incorporate constraints that ensure employees are rewarded as much as desired for a certain job-related or personal characteristic (e.g., performance). This is done by incorporating a constraint such that the coefficient in the wage regression associated with that factor cannot drop below a certain threshold. This is structurally similar to ensuring that the coefficient on β_F is close enough to zero, as in the SA as formulated above.

Using the SA to close the pay gap to zero, with a 10% max raise constraint, results in 40 raises, with an average of 3.7% per raise, at a total cost of 0.279% of the pay base. Although that is more than double the 0.127% cost of the CMA (without max raises), it is still 22% cheaper than giving equal raises to each woman, which would cost 0.356% of the pay base. Using the SA, the largest raises go to women with long working hours, those who have Important Factor B, and those in job roles with large estimated pay differences between men and women. Compared to the CMA, the SA awards more smaller raises, which are more dispersed throughout the organization. Table 1 shows the regression results after closing the gap. The gender coefficient is now exactly 0.00, while changes to other coefficients are minimal vis-a-vis their values in the original regression, except for variables where the factor-specific inequity was concentrated. The coefficient for Important Factor B has gone up by 8% (from 0.1079 to 0.1166). Similarly, the coefficients for Job Categories 2 and 15 have increased. Conversely, among employees with performance scores of 1 or 2, women were relatively highly paid, and these women have received at most only small raises; in consequence, the coefficients for those variables decreased, from 0.045 to 0.032 for Performance Category 1 and from 0.059 to 0.048 for Performance Category 2.

2.5 | Model transparency

Because the adjustments are based on regression coefficients and each employee's objective characteristics, raise suggestions can be explained, providing transparency. In particular, we can calculate the factor-specific bias for each employee, and the final raise for employee i is the difference between the final Γ and Γ_i. An example is presented in Figure 1, which shows the contribution of each pay driver for two employees. The employee with the large raise has Important Factor B, plus large positive contributions from Important Factor A and Growth Potential. The employee with the small raise has a large negative contribution from her performance score (category 2, interaction coefficient = 0.104), which is counterbalanced by her growth score and by Job Category 15 (interaction coefficient = −0.139).

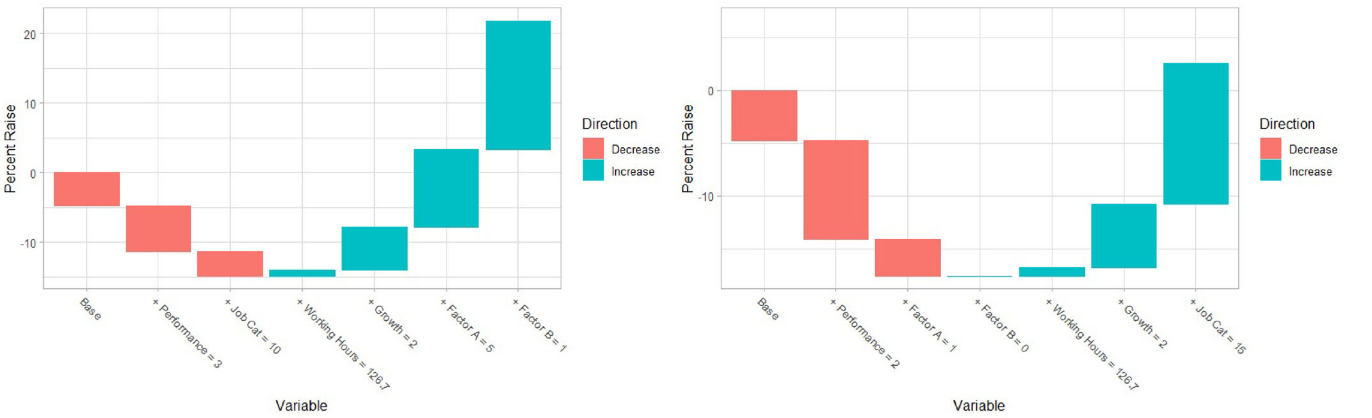
2.6 | Model robustness

The vast majority of managers with whom we have worked did not focus on the statistical significance of individual regression parameters, yet it is social sciences convention to regard an employer's wages as random samples from an (unspecified) data-generating process. For example, at NordCo we found out early on that a variable capturing employees' education was insignificant for all education levels. This was because job roles were largely homogeneous with respect to education, so this variable offered very little "additional information" about compensation. We used this absence of statistical significance to build a better wage model. Most importantly, regulators and courts frequently consider statistical significance on the basis that coefficients with large standard errors and p-values do not reflect real relationships. In particular, we want to ensure that the SA does not award raises based on estimates supported by very few employees, who may represent high-leverage outliers. Thus, we make it possible to augment the SA with robustness procedures to eliminate statistically insignificant *interaction* terms in the regression model. These procedures do not change how the gender pay gap itself is calculated.

One option is a lasso model that penalizes large regression coefficients, shrinking them towards zero (Tibshirani, 1996). This method completely zeroes out some coefficients while reducing others. We use elastic net regression, selecting the penalty value on coefficient size that minimizes the cross-validation prediction error. We only penalize the interaction terms because they determine the raises, and the main effects are all intentionally chosen by the HR team, regardless of statistical significance. Formally, we have

$$\min_{\beta,\beta_{BO}} \frac{1}{n} ||Y - (\beta X + \beta_{BO} X_{BO})||_2^2 + \lambda ||\beta_{BO}||_1. \tag{8}$$

An alternative approach is a backwards elimination model, where we iteratively eliminate the interaction term variable with the highest p-value until no interactions with a p-value above a given threshold remain. This approach offers ease of comprehension and transparency, but it introduces a discontinuity at the significance threshold, which the lasso technique avoids.



(a) **Large raise** (b) **Small raise**

FIGURE 1 Raise breakdown by pay driver for two representative employees [Color figure can be viewed at wileyonlinelibrary.com]

TABLE 3 Comparison of regression methods.

Variable	Base	Lasso	Backwards
Gender	0.150	0.0212	0.1319
Gender · Working Hours	−0.001	−0.00028	−0.0012
Gender · Perf Cat1	0.139	0.0909	0.1373
Gender · Perf Cat2	0.104	0.0676	0.1015
Gender · Perf Cat3	0.072	0.0415	0.0855
Gender · Important Factor B	−0.165	−0.1077	−0.1596
Gender · Growth Potential	−0.035	−0.0218	−0.0352
Gender · Job Cat1	−0.019	-	-
Gender · Job Cat2	−0.073	−0.0127	-
Gender · Job Cat4	0.018	0.0292	-
Gender · Job Cat5	0.065	0.0338	0.1187
Gender · Job Cat6	−0.035	-	-
Gender · Job Cat7	−0.007	0.0336	-
Gender · Job Cat8	−0.054	−0.00018	-
Gender · Job Cat10	0.040	-	-
Gender · Job Cat11	0.037	0.0725	0.0831
Gender · Job Cat12	−0.084	-	-
Gender · Job Cat15	−0.139	−0.0559	−0.0827
Gender · Job Cat16	0.111	0.02094	0.1664
Gender · Important Factor A Cat1	0.041	-	-
Gender · Important Factor A Cat2	0.028	-	-
Gender · Important Factor A Cat3	−0.151	−0.0793	−0.0934
Gender · Important Factor A Cat4	−0.076	−0.0137	-
Gender · Important Factor A Cat5	−0.113	−0.0579	−0.0649
R-squared	0.911	.910	.911
RMSE	.0916	.0922	.0919

Table 3 compares the NordCo regression coefficients in the base model, the lasso model, and after backwards elimination (here, we use a threshold of $p = 0.2$). The interaction for Important Factor B is large and negative across all three models, indicating that effect is large and significant. The working hours variable also remains negative across all three models. On the other hand, the −0.084 coefficient for Job Category 12 in the base model is eliminated by the lasso and backwards elimination models, indicating these differences may have been due to chance or regression leverage associated with a small number of employees; indeed, this category has only 20 employees.

Allocating raises based on the lasso model (the backwards elimination model) with a 10% maximum raise results in slightly lower costs, reducing the cost from 0.279% of the pay base in the base case to 0.267% (0.261%). The lasso model awards more smaller raises, compared to fewer larger raises in the base case. This is a direct result of penalizing large regression coefficients.

Another issue is that every employer measures its HR variables differently in terms of currency, units, and categories. This not only makes comparison across employers difficult but also means that coefficient magnitudes in a given employer's wage regression cannot be meaningfully compared. An easy fix for this is to standardize the variables (mean center and divide by standard deviation). However, regulators and courts do not calculate gender pay gaps using standardized variables, we have found that managers strongly prefer to work with the untransformed variables with which they are familiar, and standardization has no material effect on the allocations of raises suggested by any method considered here. We thus do not propose standardizing variables unless a client requests it.

2.7 | Model verification

It is important to carefully inspect the suggested raises to see whether the wage model needs refinement. For example, an employer may find that the raises go to many part-time employees, perhaps reflecting the fact that the organization does not pay part-time employees at the same level as full-time employees, even after normalizing their salaries to full-time equivalents. (Such disparities may be illegal in some jurisdictions.) If this differentiation is part of the compensation strategy, a new indicator variable for part-time employees can be incorporated into the regression formula and the analysis rerun. As another example, one firm found that within one job role, customer-facing employees were all paid more than those with no client interaction. The SA suggested raises for employees with no customer contact, which was not in line with the company's compensation strategy. Based on this observation, the company split the job role into two roles in its internal classification system. At another firm, the largest suggested raise was given to an employee who was on a performance improvement plan, and the company was hoping the employee would find a new employer. The SA correctly identified that the employee was underpaid relative to expectations, but it was a conscious choice reflecting poor performance (which the employer wanted to leave outside the scope of the model). Such cases can be accounted for by using additional constraints that set the appropriate $\delta_i^{\prime}$s to zero in the formulation.

2.8 | Budget constraints

NordCo was fortunate that eliminating all factor-specific differences in compensation was financially feasible, but some employers must more carefully balance fairness with cost. To show the trade-offs faced by these employers and give them a menu of options, we proceed as follows.

First, define the amount of factor-specific inequity in a company's wage model, θ, calculated as $\theta = \sum_i \max(0, \Gamma_i)$ for all employees of the underpaid gender, depending on the context. Completely eliminating the factor-specific inequity would result in $\theta = 0$. We can normalize the reduction in inequity by calculating the percentage change in θ before and after raises are awarded. Second, we augment the SA optimization formulation by adding a budget constraint.

We illustrate the procedure with the NordCo data by imposing budget constraints going from 0.1% of the pay base (less than the cost of the CMA) up to 0.3% of the pay base (more than the cost of the SA) and varying the target β_F. Figure 2 shows the performance of the SA across a range of budgets and target thresholds. For any budget, there is a trade-off between reducing the pay gap itself and reducing factor-specific inequity. With a budget of 0.1% of the total pay base, we can either reduce factor-specific inequity by 60% while bringing β_F to -0.008, or we can almost close the pay gap entirely, bringing β_F to -0.002 but only eliminating 20% of the factor-specific inequity. With a budget of 0.2% of the pay base, we can eliminate over 83% of factor-specific inequity while bringing β_F to -0.002, or we can set the target threshold to 0 and reduce θ by 72%.

Presenting managers with a figure like Figure 2 gives them an easy-to-understand menu of options (budget, pay gap, factor-specific inequity) from which to choose. For the reasons previously discussed, we encourage employers to focus on fairness, but clients who are facing tight budget constraints and have strong pressures to reduce their equal pay gap need to balance these considerations.

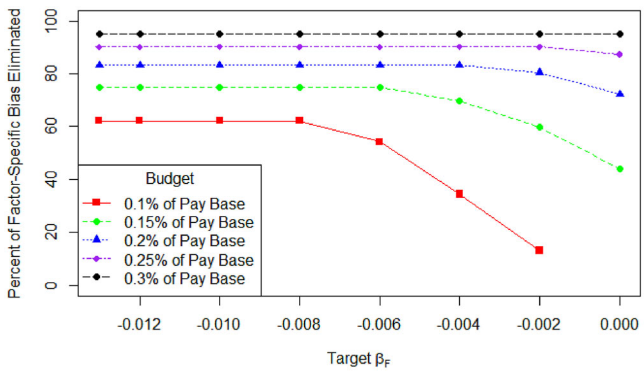


FIGURE 2 Pay gap and factor-specific inequity reduction as a function of budget [Color figure can be viewed at wileyonlinelibrary.com]

2.9 | The target pay gap

A natural question that arises as a company measures its equal pay gap is "what is the right target?" Even in the absence of true bias or discrimination, employee turnover, new hires, and salary changes all impact the pay gap. Therefore, an organization's pay gap is an evolving statistic, but in the absence of bias it should oscillate around zero.

For many firms, then, an equal pay gap of 0.0% is the goal. This target is easy to communicate to employees, regulators, and other stakeholders as a clear standard for pay equity. Depending on the magnitude of β_F and budget concerns, some organizations choose intermediate values for T_F, perhaps reaching zero over a few compensation cycles.

In these cases, internal and external stakeholders may have very different expectations. External stakeholders (especially regulators) might focus on the statistical significance of β_F, whereas internal stakeholders might be more interested in the economic significance (size) of the pay gap. Thus, a 3% pay gap that is not statistically significant at conventional levels may satisfy regulators but dissatisfy employees. Accordingly, we use three approaches to determine when action needs to be taken: statistical significance, practical significance, and reversing the burden of proof.

2.9.1 | Statistical significance

Firms that are exclusively focused on external stakeholders, such as regulatory bodies or certification agencies, or on minimizing legal risk will want to focus on the statistical significance of β_F according to the following hypothesis:

$$H_0 : \beta_F = 0,$$
$$H_a : \beta_F \neq 0. \tag{9}$$

If the regression coefficient is significantly lower than 0 at, say, the $\alpha = 0.05$ level, then a gap exists and needs to be closed. If it is not, then external stakeholders are unlikely to impose sanctions, and the firm need take no action. Firms with this focus may accordingly choose a target pay gap a little closer to zero (to be safe) than $t_{\alpha/2,n-2} \cdot SE(\beta_F)$, corresponding to the upper bound of the $(1-\alpha)$% confidence interval around zero.

However, it is unclear whether the assumptions underlying standard hypothesis testing apply. If not, bootstrapping is a reasonable alternative (Efron & Hastie, 2021). To illustrate, we take 10,000 samples of 471 NordCo employees, with replacement, and run the regression in Equation (3). The average coefficient is −0.01333, and the standard error of the coefficients is 0.01285. These numbers are very similar to the −0.0130 coefficient and −0.01295 standard error from the original analysis above. Indeed, with the log-linear wage regression, bootstrapping tends to produce similar standard errors to those generated by conventional regression statistics. Another reason for caution is that many practitioners are not familiar with bootstrapping.

2.9.2 | Rules of thumb

For firms that are focused on pay equity not as a regulatory obligation but as an integral part of their larger HR strategy, statistical insignificance of the gender pay gap is insufficient. This is especially true given that each organization is analyzing not a random sample but their entire data; as noted previously, this could be construed as a population, meaning that p-values are zero by definition. We, therefore, suggest that these organizations set an internal threshold below which any pay gap, regardless of statistical significance, is not considered to be of practical importance. To illustrate these threshold guidelines, we simulate companies of different sizes with no true underlying bias in pay (true $\beta_F = 0$) and measure the estimated gender pay gap in a sample population. Figure 3 shows the 80th percentile of the absolute value of the equal pay gaps for these simulated companies. The magnitude of the expected gap is a function of three underlying factors: the gender split of the employee population, the fit of the regression model, and, most importantly, the size of the organization. Small firms and firms whose salary models explain a relatively low proportion of pay variation (i.e., a low R^2) are likely to have larger measured pay gaps simply due to sample variation, which in the real world would reflect some randomness in employee recruitment, retention, and turnover. For firms with well under 100 employees, this randomness, even absent discrimination, can generate a large pay gap so easily that we omit them from the figure.

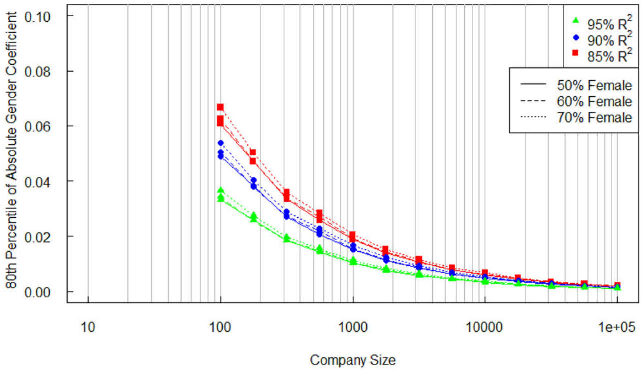


FIGURE 3 The 80th percentile of the absolute value of the gender coefficient in the absence of underlying gender inequity in pay, as a function of company size and different levels of variability in the salary structure and gender split [Color figure can be viewed at wileyonlinelibrary.com]

TABLE 4 The approximate 80th percentile of the absolute measured pay gap by company size and strength of the compensation model, assuming an equal employee gender split

	Model fit (R^2)		
Company size	**85%**	**90%**	**95%**
Small (200 employees)	5%	4%	2.5%
Medium (1000 employees)	2.25%	2%	1.75%
Large (5000 employees)	1%	0.5%	0.3%
Very large (more than 10,000 employees)	0.2%	0.1%	0.1%

Typically compensation models have an R^2 of at least 85% or above. Government agencies or other organizations with very rigid pay structures typically have an R^2 around 95%. Further, we have found that as firms get bigger, they have more sophisticated employee data, more homogeneous job roles, and thus a higher model R^2 than smaller companies. From Figure 3, we note that companies of 500 (1000) employees can expect an absolute gender coefficient of 0.022 (0.02) or less 80% of the time when their salary model explains 90% of the variability in pay. For organizations with 10,000 employees, the measured coefficient is below 0.002 more than 80% of the time as long as the compensation model's R^2 is at least 85%.

Table 4 extracts key information from Figure 3 and presents rules of thumb. These are rounded estimates based on a firm's size; firms should generally expect their measured pay gap to be smaller than these guidelines in the absence of discrimination. The following points summarize the conclusions from the different approaches to determining the significance of a pay gap:

1. If the pay gap is statistically significant according to conventional regression statistics, then it should be reduced irrespective of its size, as discussed above.
2. Even if statistically insignificant, a pay gap that exceeds the magnitude that can be expected in 80% of scenarios

for an organization's size, gender composition, and model fit, as per Figure 3, should be reduced.

3. Organizations for which the threshold guidelines represent a pay difference that significantly affects employees should set their own internal thresholds that are *lower* than those proposed by the simulation, such as 0.5%, 1%, or 2%. Likewise, small organizations should set their own internal thresholds.

2.9.3 | Reversing the burden of proof

The rules of thumb proposed above are defensive: they only allow an employer to claim that there is no clear evidence of a pay gap. Likewise, an insignificant coefficient on gender only implies that one cannot claim that a firm has a pay gap. But for firms who want to be pay equity leaders, we propose a way to assert that the evidence strongly suggests they do *not* have a pay gap, even though at any given time, a small pay gap may appear (e.g., due to random fluctuations in hiring and retention).

In these cases, we shift the burden of proof onto the firm to test the hypothesis that the absolute pay gap is less than some practical significance threshold, δ: $H_0 = \beta_F \notin [-\Delta, \Delta]$. Abadie (2020) discusses the information contained in a non-significant regression coefficient (with respect to a null of zero) and how it shifts the posterior distribution of possible coefficient values in a Bayesian framework. We take a similar approach, setting a null hypothesis that there is a pay gap bigger than some practical significance level and measuring the statistical significance of the evidence against it. Since our goal is to demonstrate that β_F is equivalent to zero, this is where the burden of proof rests. We conclude that the pay gap is effectively zero if and only if we have sufficient evidence that the pay gap is smaller than a given magnitude—that is, when we determine with a certain level of confidence that the true difference in pay between otherwise similar men and women is within a certain range of zero. To construct our test, we adapt methods from the statistical literature on equivalence testing (Wellek, 2010).

To begin, we set Δ as the level of practical significance, that is, a value that is close enough to zero to be considered essentially zero. We construct the following $(\Delta, \varkappa)$ hypothesis test:

$$\begin{aligned} H_0 &: |\beta_F| \geq \Delta, \\ H_a &: |\beta_F| < \Delta, \end{aligned} \tag{10}$$

where $\varkappa$ is the required significance level and the absolute value makes the test symmetrical between women and men. (This is the reverse of the traditional hypothesis test associated with regression coefficients, where H_0 corresponds to $\beta_F = 0$.) In this case, if the resulting entire $(1-\varkappa)\cdot 100\%$ confidence interval around the measured pay gap is within $[-\Delta, \Delta]$, we can claim with $\varkappa\%$ confidence that the company does not have a pay gap larger than Δ. Here, Δ can be thought of as the maximum acceptable β_F. Values below this are accepted as being practically equivalent to zero, indicating no pay gap.

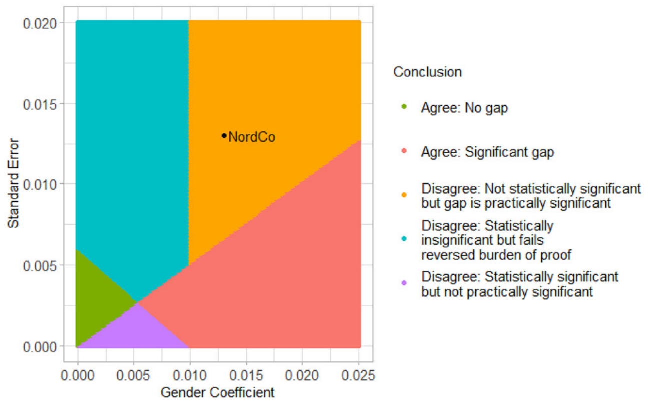


FIGURE 4 A comparison of methods for assessing the significance of a gender pay gap coefficient [Color figure can be viewed at wileyonlinelibrary.com]

The power (defined as the probability of rejecting the null hypothesis when the alternative hypothesis is true) can be related to key dimensions of a firm's workforce. We will assume that ϵ in Equation (1) is normally distributed with a mean of zero and a standard deviation of σ, and that the statistical assumptions of the log linear regression model are met. Then b_F (the sample estimate of β_F) follows a t-distribution. It is important to note that the variance of X_F, $Var(X_F)$, and as a result the variance of b_F, is dependent on the gender distribution of the company and are minimized when there is an even gender split. By applying the normal approximation to the t distribution (which is reasonable for all but the smallest firms), the power of the equivalence test is the probability that the absolute value of a $\mathcal{N}(0, \frac{\sigma^2}{nVar(X_F)})$ random variable is smaller than $\Delta - w_\varkappa/2$, where w_k is the length of a $(1-\varkappa\%)$ confidence interval around the random variable. The power of the statistical model depends on the width of the confidence interval and is thus a function of the population size (through n, the larger the population, the higher the power), the gender split (the more even the gender split, the higher the power), and the goodness of fit of the salary model (the better the model fit, the higher the power). Further, by definition $\Delta \geq \frac{w_\varkappa}{2}$, so that when the variability in our estimate of the pay gap is large, Δ needs to be large in order for the test to have any power. If Δ is less than the half-width of the confidence interval, the power of the test is zero.

2.9.4 | Method comparison

Figure 4 compares the methods by plotting the measured gender coefficient on the x-axis and the standard error of this coefficient on the y-axis. In this graph, we use a Δ practical significance level of 0.01 for the gender coefficient, a 90% confidence interval for the reversed burden of proof test, and $\alpha = 0.05$ for the traditional significance test. There are five regions in the graph. In the red region in the lower right, the gender coefficient is large and the standard error is small. Here, a statistically significant large gap is precisely measured, and all methods agree that there is a gap that needs to be closed. In green, in the lower left, we have a small coefficient and a small or modest standard error, and the methods agree there is no significant pay gap. In light blue, at the top left of the graph, the methods disagree. In this zone the standard error is large, and thus the traditional hypothesis test cannot reject the null hypothesis of no pay gap. The gap is not practically significant either, but due to the large standard error (typically due to small firm size) the firm also cannot affirmatively claim that they have no pay gap. A firm trying to reverse the burden of proof might seek to lower its pay gap and move into the green region, but it may also be satisfied. In the orange region, the methods disagree; the pay gap is statistically insignificant, but the regression coefficient is above the practical significance boundary and clearly fails the reversed burden of proof test. In this region, we recommend firms attempt to close the gap to move into the blue region. This was NordCo's position at the time of our initial engagement. Because the pay gap was above their practical significance threshold, they allocated raises to close the gap, even though the coefficient was not statistically significant.

Lastly, in the center bottom of the graph, the purple region indicates a precisely measured, very small pay gap. Because the standard error is so small, even a very small coefficient may have a p-value of less than 0.05. However, the reversed burden of proof method says with 90% confidence that the pay gap is not practically significant, so a firm using this method need not take action (but would have the option of doing so). Firms in this region are likely to be very large firms with correspondingly low standard errors.

By providing this menu of options to firms and discussing their implications given the firm's culture, regulatory environment, and stakeholders, we allow the firm's HR professionals and top managers to choose the right approach for their firm's unique situation.

3 | IMPLEMENTATION

3.1 | Pay gap project lifecycle

Managing pay equity is an ongoing process and not a single fixed-term effort. Historically, pay equity has been seen as a regulatory issue or a legal risk, but rarely as a strategic priority. However, this is changing as new external pressures (e.g., from activist investors, changing social attitudes, or government regulations) and internal pressures (e.g., employee morale or retention) are leading more firms to focus on pay equity.

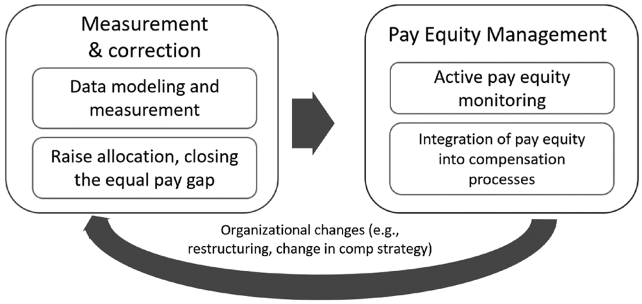


FIGURE 5 An overview of the pay gap project lifecycle, from addressing the legacy pay gap, through initial modeling and correction, through maintenance and integration

Every personnel decision a firm makes, from hiring and initial salaries to promotion and raises, influences the evolution of its pay gap. A pay gap is thus not only a reflection of current compensation decisions but also many past decisions. Once an organization has addressed its "legacy pay gap," it becomes crucial to maintain a fair pay structure so that the gap does not reemerge. In order to avoid getting stuck in an annual cycle of running an equal pay audit and then fixing any identified issues, firms need to proactively manage their pay gaps throughout the year by maintaining constant vigilance and awareness of how individual decisions impact the pay gap, particularly when setting salaries for new employees and awarding raises during merit review cycles. In other words, the integration of pay equity into compensation decision processes is necessary in order to maintain pay equity. Figure 5 illustrates this by depicting the lifecycle of a pay gap project.

3.1.1 | NordCo's pay equity journey

In the decade before we started working with NordCo, the firm had made some progress in reducing its pay gap, but they could never achieve their goal of zero. By 2017, within 12 months of the start of our collaboration, NordCo had eliminated their pay gap and has kept it closed since then. Figure 6 shows NordCo's equal pay gap measured annually from 2016 to 2020, as well as the pay gap progression for two other firms we have worked with; these other firms took a few years to close their pay gaps. Panel (b) shows NordCo's pay gap measured monthly in 2018. We see that at the end of each year after our project with NordCo, the pay gap has been completely closed, or very nearly so. However, over the course of the year, the pay gap fluctuates around 0% in either direction. If a gap begins to emerge and exceeds their internal practical significance threshold of 1% in either direction, managers are empowered to take action immediately rather than waiting for the results of an annual process. NordCo has achieved these results by using their salary model to help inform decision-making throughout the pay cycle, from starting salaries for new employees and newly promoted

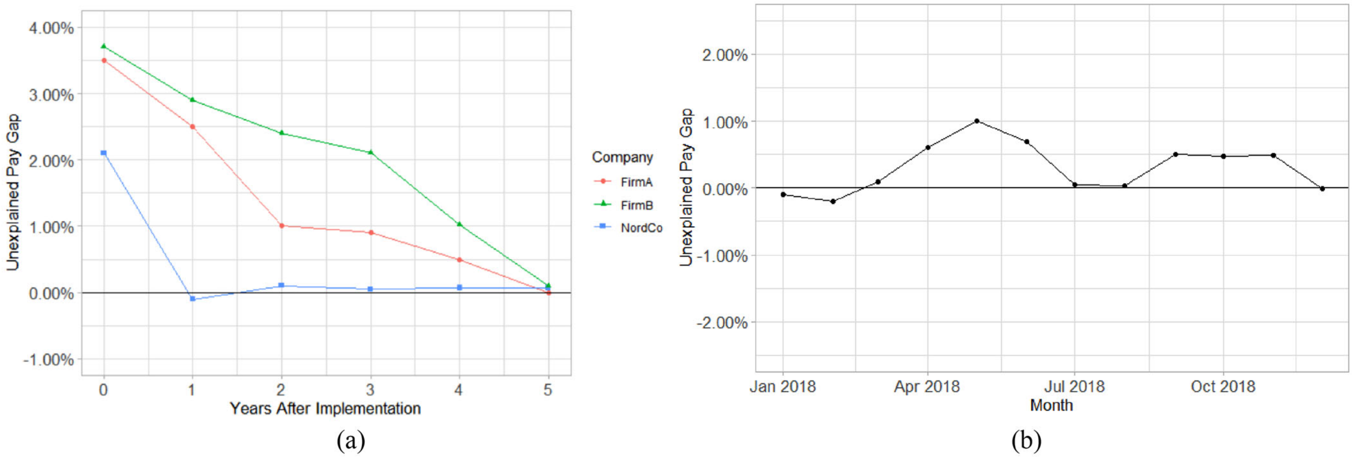


FIGURE 6 The (a) annual pay gap from 2016 to 2020 for NordCo and from the first 5 years of our projects for two other firms and (b) NordCo's pay gap over the course of 2018 [Color figure can be viewed at wileyonlinelibrary.com]

employees to understanding at the time of any compensation decision its exact impact on the pay gap. As one HR manager said to us, "We were always looking in the rear-view mirror by analyzing old data. When the results for external analysis came in, a lot had changed within the company. It was therefore unclear how we should fix the pay gap. As a result, we searched for a solution that could show the situation and enable us to evaluate salary decisions in real time."

3.2 | Software application

To help NordCo and other organizations achieve and maintain pay equity throughout every step of the pay gap lifecycle, we have developed an intuitive yet powerful software solution. The software is linked to a firm's HR systems to support data cleaning and visualization, statistical modeling, and raise allocation to close any pay gap identified. It also helps firms measure the pay gap impact of HR decisions, like allocating merit raises or setting the compensation of new hires. Figure 7 shows a screenshot from the overview page after a raise allocation analysis has been run. Given that some of our clients have tens of thousands of employees across multiple continents, manually calculating individual raise suggestions for pay equity would be implausible. Our experience has been that once the initial salary model has been built satisfactorily, which usually requires a few iterations of defining new variables and standardizing definitions, customers have been happy with the raises suggested and are able to work independently with the software.

4 | CONCLUSION

The methodological approach described in this paper has been used as part of the pay equity process in over 40 organizations ranging from 75 to 130,000 employees, across a range of industries, including finance and insurance, utilities, logistics, retail, hospitality, telecommunications, healthcare, and local and state governments. Our tools are also used by several international consulting firms. Altogether, our tools have been used to measure and help address pay equity in organizations collectively employing over 500,000 people. One of our largest partners, a Fortune Global 500 company, has successfully eliminated their equal pay gap in each of their national subsidiaries.

In our work with companies, we have found that the hardest, most time consuming, and most important step is defining the variables for the salary model: deciding which variables to include and ensuring that they sufficiently capture the nuances of compensation within the firm. The next step of standardizing variable definitions across departments, locations, and managers also requires careful implementation. Once the HR managers are confident in the statistical salary model, they have typically been receptive to the raise suggestions it offers as a basis for closing pay gaps, but depending on the organization those suggestions may not only be reviewed by managers but by legal counsel as well. We have also found that companies are appreciative of tools that help them proactively monitor the effects of their HR decisions on the pay gap. When managers can anticipate these impacts, they have more control and can prevent a pay gap from re-emerging.

The SA developed herein explicitly corrects biases throughout a firm's pay structure. The cost of the approach is heavily dependent on where the largest inequities manifest themselves and the operational constraints included in the optimization routines. Eliminating pay inequity throughout a salary structure may cost significantly more than simply eliminating the pay gap as typically measured in the cheapest way, but the SA is the only approach to date that addresses the spirit of equal pay for equal work legislation and thus minimizes regulatory risk. The approach can also incorporate budget constraints and a firm's desired trade-off between closing the pay gap and addressing factor-specific inequity.

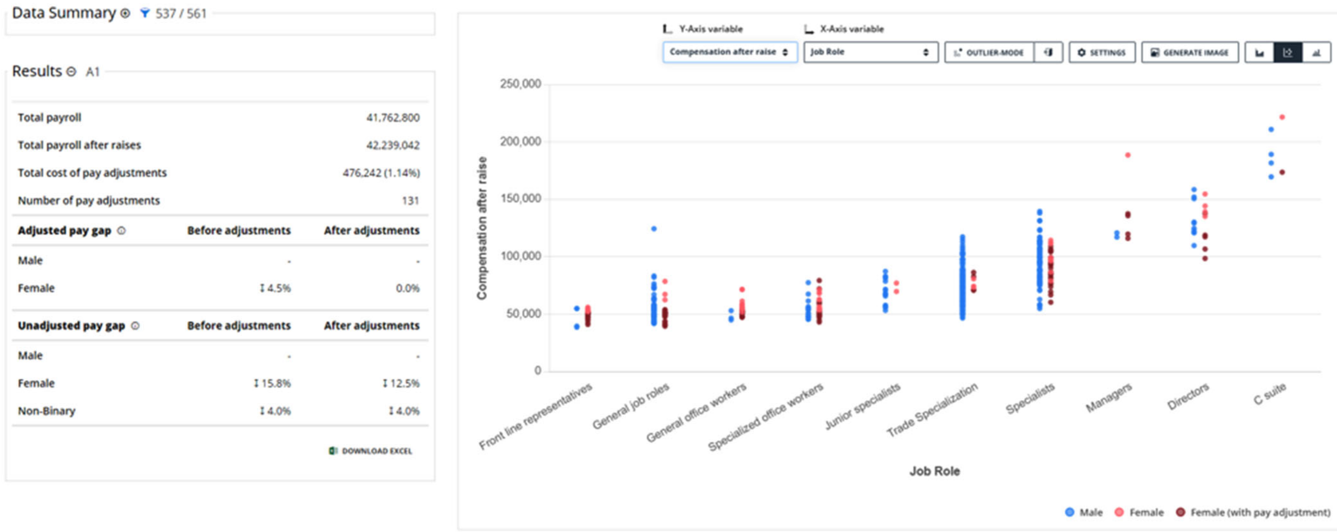


FIGURE 7 Screenshot from the software system [Color figure can be viewed at wileyonlinelibrary.com]

Subjective factors always play a role in salary decisions. However, the approach presented in this paper offers HR managers an optimal starting point, which can be applied at scale, from which to make adjustments. Any factors that lead to consistent changes to the suggested raises can easily be added as constraints to the optimization formulation or as factors in the salary model.

It is worth noting the limitations of our approach. While the SA can close the "unexplained" gap efficiently and fairly, the overall difference in earnings between female and male employees may not be significantly reduced if men are disproportionately in higher paying job categories or endowed with characteristics like education that tend to increase pay. To wit, we find that for NordCo, the raw difference in average salaries between men and women was only reduced from 11.0% to 10.2%. Applying the operations toolkit to representation across job categories and endowments, which relate to the processes of hiring and promotion, is beyond the scope of this paper, as is ensuring equal access to promotions and overtime. These are exciting avenues for future research.

Additionally, one major challenge has been internal communication: upwards to upper management, horizontally to other HR managers, and outside to other workers involved in the compensation process, like recruiters and department managers. These workers need to understand the difference between the unadjusted pay gap and the adjusted equal pay gap and how their actions influence each.

Even so, the SA can play an important role in illuminating the importance of various factors in determining pay inside a firm, and it can prompt reevaluation of what factors should influence pay. For instance, a firm may run the SA and be puzzled to discover that some employees are paid a premium. Investigation may then reveal that these employees have better communication and negotiation skills. The firm can then decide whether these skills should be explicitly included in the firm's pay model. We note, however, that a pay driver may become "tainted" if it is used to justify spurious differences in compensation, for example, giving members of one demographic group higher titles to justify their higher pay. Another complicated issue is the influence of external offers. On the one hand, some employees may reasonably be paid a premium because they have received an outside offer. On the other hand, the propensity to seek or receive external offers may vary along demographic lines. In practice, however, many employers want to avoid allowing the external market to give rise to a pay gap, both for regulatory and legal reasons and because it would conflict with the firm's values.

We have also observed that although the SA focuses on pay disparities among groups and is not designed to correct disparities at the individual level, clients have consistently reported that the salary models they develop as part of the SA help them identify and correct individual pay disparities. Similarly, the SA, narrowly construed, is prospective and therefore does not address accumulated pay inequity from having underpaid a demographic group over a period of years. However, it is straightforward to apply the SA to historical pay data and apply compensatory payments for past inequity.

Quantitative approaches to HR management and hiring decisions are becoming mainstream, and there is a call for greater transparency and equity in pay. Our work lies at the intersection of these two cultural and research currents, and we hope it will provide a basis for a better understanding of the complex factors inherent in salary structures and for the development of practical salary decision support tools. There is also potential for our work to be applied to investigate and, if necessary, address other types of demographic pay gaps. In summary, by using optimization and statistics, we support organizations in delivering on the promise of equal pay for equal work.

CONFLICT OF INTEREST

The authors have declared no conflict of interest.

ORCID

David Anderson https://orcid.org/0000-0002-1128-6206
Margrét Vilborg Bjarnadóttir https://orcid.org/0000-0003-2955-1992
David Gaddis Ross https://orcid.org/0000-0001-6561-6234

REFERENCES

Abadie, A. (2020). Statistical nonsignificance in empirical economics. *American Economic Review: Insights*, *2*(2), 193–208.

Abadie, A., Athey, S., Imbens, G. W., & Wooldridge, J. M. (2020). Sampling-based versus design-based uncertainty in regression analysis. *Econometrica*, *88*(1), 265–296.

Anderson, D., Bjarnadóttir, M. V., Dezső, C. L., & Ross, D. G. (2019). On a firm's optimal response to pressure for gender pay equity. *Organization Science*, *30*(1), 214–231.

Becker, W. E., & Goodman, R. (1991). The semilogarithmic earnings equation and its use in academe. *Economics of Education Review*, *10*(4), 323–332.

Becker, W. E., & Toutkoushian, R. K. (1995). The measurement and cost of removing unexplained gender differences in faculty salaries. *Economics of Education Review*, *14*(3), 209–220.

Benvenuto, L. J., Anderson, D. R., Kim, H. P., Hook, J. L., Shah, L., Robbins, H. Y., D'Ovidio, F., Bacchetta, M., Sonett, J. R., Arcasoy, S. M., & the Columbia University Lung Transplant Program. (2018). Geographic disparities in donor lung supply and lung transplant waitlist outcomes: A cohort study. *American Journal of Transplantation*, *18*(6), 1471–1480.

Bertsimas, D., Farias, V. F., & Trichakis, N. (2013). Fairness, efficiency, and flexibility in organ allocation for kidney transplantation. *Operations Research*, *61*(1), 73–87.

Biro, M. M. (n.d.). Why HR data is the key to strategic business decisions. *Viser Insights*, https://www.visier.com/clarity/why-data-is-the-key-to-strategic-hr-decisions/

Blau, F. D., & Kahn, L. M. (2000). Gender differences in pay. *Journal of Economic Perspectives*, *14*(4), 75–99. http://www.aeaweb.org/articles?id=10.1257/jep.14.4.75

Blau, F. D., & Kahn, L. M. (2003). Understanding international differences in the gender pay gap. *Journal of Labor Economics*, *21*(1), 106–144.

Blau, F. D., & Kahn, L. M. (2017). The gender wage gap: Extent, trends and explanations. *Journal of Economic Literature*, *55*(3), 789–865.

Blinder, A. S. (1973). Wage discrimination: Reduced form and structural estimates. *Journal of Human Resources*, *4*, 436–455.

Boudreau, J., Hopp, W., McClain, J. O., & Thomas, L. J. (2003). On the interface between operations and human resources management. *Manufacturing & Service Operations Management*, *5*(3), 179–202.

Choi, C. S. (2019). How fast-growing companies can scale their hr processes. *Glassdoor Homepage*. https://www.glassdoor.com/employers/blog/scale-hr-processes/

Cotton, J. L., & Tuttle, J. M. (1986). Employee turnover: A meta-analysis and review with implications for research. *Academy of Management Review*, *11*(1), 55–70.

Davies, R. (2019). How to structure your HR department as your company grows. *Software Advice*. https://www.softwareadvice.com/resources/hr-department-structure/

Domonoske, C. (2018). Companies in Iceland now required to demonstrate they pay men, women fairly. *NPR*. https://www.npr.org/sections/thetwo-way/2018/01/03/575403863/companies-in-iceland-now-required-to-demonstrate-they-pay-men-women-fairly

Duan, N. (1983). Smearing estimate: A nonparametric retransformation method. *Journal of the American Statistical Association*, *78*(383), 605–610.

Efron, B., & Hastie, T. (2021). *Computer age statistical inference, student edition: Algorithms, evidence, and data science*. Institute of Mathematical Statistics Monographs, Series Number 6. Cambridge University Press.

European Commission (2021). *Pay transparency: Commission proposes measures to ensure equal pay for equal work* [Online]. https://ec.europa.eu/commission/presscorner/detail/en/ip_21_881

Foley, M., Cooper, R., & Mosseri, S. (2019). *Gender equitable recruitment and promotion: Leading practice guide* (WGEA Commissioned Research Paper). The Australian Women's Working Futures (AWWF) Project, University of Sydney, Sydney, Australia. https://ses.library.usyd.edu.au/bitstream/handle/2123/21261/Recruitment_and_Promotion_0.pdf?sequence=2&isAllowed=y

Gov.uk Collections (2020). *Gender pay gap reporting* [Online]. https://www.gov.uk/government/collections/gender-pay-gap-reporting

Harmon, C., Oosterbeek, H., & Walker, I. (2003). The returns to education: Microeconomics. *Journal of Economic Surveys*, *17*(2), 115–156.

Holmström, J., Ketokivi, M., & Hameri, A.-P. (2009). Bridging practice and theory: A design science approach. *Decision Sciences*, *40*(1), 65–87.

Hom, P. W., Lee, T. W., Shaw, J. D., & Hausknecht, J. P. (2017). One hundred years of employee turnover theory and research. *Journal of Applied Psychology*, *102*(3), 530.

Hopp, W. J., & Oyen, M. P. (2004). Agile workforce evaluation: A framework for cross-training and coordination. *IIE Transactions*, *36*(10), 919–940.

Horowitz, J. (2017). Salesforce just spent another $3 million to close its pay gap. *CNN*. https://money.cnn.com/2017/04/04/news/companies/salesforce-equal-pay-women/index.html

Huckman, R. S., & Staats, B. R. (2011). Fluid tasks and fluid teams: The impact of diversity in experience and team familiarity on team performance. *MSOM*, *13*(3), 310–328.

Lang, K., & Lehmann, J.-Y. K. (2012). Racial discrimination in the labor market: Theory and empirics. *Journal of Economic Literature*, *50*(4), 959–1006.

Loriggio, P. (2016). University of Waterloo boosts female faculty pay after wage gap uncovered. *The Globe & Mail*. https://www.theglobeandmail.com/news/national/university-of-waterloo-boosts-female-faculty-salaries-after-gender-pay-gap-found/article31272503/

McGregor, J. (2015). Salesforce CEO combs through 16,000 workers' salaries to make sure women are paid fairly. *The Washington Post*. https://www.washingtonpost.com/news/on-leadership/wp/2015/04/27/salesforce-ceo-combs-through-16000-workers-salaries-to-make-sure-women-are-paid-fairly/?noredirect=on&utm_term=.fa8f742e7f9c

Mooney, J. J., Bhattacharya, J., & Dhillon, G. S. (2019). Effect of broader geographic sharing of donor lungs on lung transplant waitlist outcomes. *The Journal of Heart and Lung Transplantation*, *2*, 136–144.

Oaxaca, R. L., & Ransom, M. R. (1999). Identification in detailed wage decompositions. *Review of Economics and Statistics*, *81*(1), 154–157.

Okun, A. M. (2015). *Equality and efficiency: The big tradeoff*. Brookings Institution Press.

Ramachandran, K. M., & Tsokos, C. P. (2009). *Mathematical statistics with applications*. Academic Press.

Siemsen, E., Balasubramanian, S., & Roth, A. V. (2007). Incentives that induce task-related effort, helping, and knowledge sharing in workgroups. *Management Science*, *53*(10), 1533–1550.

Society for Human Resource Management (2019). *France: New gender equality obligations established* [Online]. https://www.shrm.org/resourcesandtools/legal-and-compliance/employment-law/pages/global-france-gender-equality-obligations.aspx

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, *58*(1), 267–288.

Toutkoushian, R. K., & Hoffman, E. P. (2002). Alternatives for measuring the unexplained wage gap. *New Directions for Institutional Research*, *115*, 71–89.

Viser (2018). *The datafication of HR: How you can make better people decisions* (White Paper). https://www.visier.com/wp-content/uploads/2018/02/The-Datafication-of-HR-Graduating-From-HR-Metrics-to-People-Analytics.pdf

Wagstaff, A. (1991). QALYs and the equity-efficiency trade-off. *Journal of Health Economics*, *10*(1), 21–41.

Weichselbaumer, D., & Winter-Ebmer, R. (2005). A meta-analysis of the international gender wage gap. *Journal of Economic Surveys*, *19*(3), 479–511. https://EconPapers.repec.org/RePEc:bla:jecsur:v:19:y:2005:i:3:p:479-511

Wellek, S. (2010). *Testing statistical hypotheses of equivalence and noninferiority*. Chapman and Hall/CRC.

Woessmann, L. (2008). Efficiency and equity of European education and training policies. *International Tax and Public Finance*, *15*(2), 199–230.

Yee, R. W., Yeung, A. C., & Cheng, T. E. (2008). The impact of employee satisfaction on quality and profitability in high-contact service industries. *Journal of Operations Management*, *26*(5), 651–668.

Yee, R. W., Yeung, A. C., & Cheng, T. E. (2010). An empirical study of employee loyalty, service quality and firm performance in the service industry. *International Journal of Production Economics*, *124*(1), 109–120.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Anderson, D., Bjarnadóttir, M. V., & Ross, D. G. (2023). Bridging the gap: Applying analytics to address gender pay inequity. *Production and Operations Management*, 1–19. https://doi.org/10.1111/poms.13944