Original Article

Public Personnel Management
1–26

DOI: 10.1177/0091026019845119
journals.sagepub.com/home/ppm
SAGE

How to Do a Salary Equity Study: With an Illustrative Example From Higher Education

Lori L. Taylor[1], Joanna N. Lahey[1], Molly I. Beck[2], and Jeffrey E. Froyd[3]

Abstract

Regular salary equity studies can be a best practice among employers committed to salary equity and fairly managed compensation. While a well-constructed salary study can identify inequities for amelioration, a poorly constructed study can create rather than solve problems. Organizations may be deterred from doing these studies because of their inherent analytical challenges. We provide a guide for human resource managers describing how to conduct their own salary studies, how to interpret the results, and how organizations can apply the results. We describe best practices across public sector organizations and illustrate them with an example from higher education. We also provide a link to an online appendix containing sample code that can be used to conduct such analyses using two popular software packages. The twin goals of the article are to increase the quality of salary analyses while reducing the barriers to conducting them.

Keywords

salary study, practitioner's guide to salary equity studies, detecting sex, race and loyalty pay gaps, hedonic wage analysis

In 2016, President Obama asked businesses to sign the White House Equal Pay Pledge. One of the essential planks of the pledge was a commitment to conducting "an annual company-wide gender pay analysis across occupations" ("White House Equal pay

[1]Texas A&M University, College Station, USA
[2]University of Arkansas, Fayetteville, USA
[3]The Ohio State University, Columbus, USA

Corresponding Author:
Lori L. Taylor, Department of Public Service and Administration, Texas A&M University, 4220 TAMU, College Station, TX 77843-4220, USA.
Email: lltaylor@tamu.edu

pledge," 2016). At roughly the same time, the U.K. government started implementing a policy that requires all employers—public and private—with at least 250 employees to provide annual reports on salary equity ("Equality Act, 2010 (Gender Pay Gap Information Regulations," 2017). The implication is clear: regular pay analysis can be a best practice among employers committed to salary equity.

Unfortunately, there is little evidence that the public sector in the United States—which could be expected to lead on this issue—has adopted this diagnostic best practice. A recent study by the U.S. Office of Personnel Management (OPM, 2014) found that few federal agencies conduct regular salary reviews. Reese and Warner (2012) found it "extremely difficult" (p. 326) to find anyone at the state level who could say whether a state was monitoring compliance with laws regarding pay equity. It is unlikely that local government is doing more.

Fear of litigation may contribute to the public sector's dearth of regular salary analyses. Indeed, when we began our analyses for a large public-sector organization, the first stop was with the lawyers. However, carefully conducted salary analyses are as likely to remove sources of grievance as to reveal them.

A belief that formal salary structures prevent inequity might also inhibit the use of public sector salary studies. Why invest in analyzing that which by rule cannot exist? However, the OPM (2014) report makes it clear that they found evidence of inequities arising from "discretionary authority" (p. 2) even in federal agencies with set salary schedules.

A more basic reason why the public sector hesitates to conduct regular salary analyses may be a lack of expertise. A well-executed salary analysis requires tools and techniques outside the skill set of most traditionally trained human resource managers, who may be tempted to rely instead on comparisons of average salary by position or unit when evaluating equity. After all, it seems intuitive that someone whose salary is close to average for their position is being compensated fairly. However, comparisons based on average salaries can be misleading. Average salaries can be skewed by the earnings of a small number of individuals, and within-group comparisons might not be appropriate if there are within-group differences in worker productivity. It can be equitable for more-skilled managers to earn more than less-skilled managers, for example.

A well-designed analysis of salary equity accommodates the multidimensional nature of the organization's labor force and overcomes many of the drawbacks associated with average salary comparisons.[1] It supports the pay for performance, broadbanding systems, and monetary bonuses and awards that are part of a manager's toolkit to enhance work motivation and reduce turnover among performers (e.g., Lawler, 1990; Rainey, 2014; Thompson, 2007). It can thereby provide a formal way to increase personnel flexibility in public organizations often seen to be mired in red tape (Feeney & Rainey, 2010). It also allows administrators to investigate alternative explanations for patterns that initially appear to indicate bias and to demonstrate, for example, that average salaries are higher for men than for women not because of bias but because the men have more professional experience.

This article draws on the existing literature on salary inequities to provide a detailed guide for human resource managers looking to conduct their own salary equity analyses. Our twin goals are to increase the frequency with which salary analyses are conducted in the public sector by lowering existing barriers to conducting such analyses, and to improve the quality of any such analyses conducted. We describe best practices across a variety of public sector organizations and illustrate the importance of adopting those best practices using a specific example from higher education. We then provide the specific code required to conduct such analyses using two widely used software packages (SAS and Stata), and detailed explanations regarding interpretation of regression results.

Why Do a Salary Study?

Public organizations may be concerned about two types of salary inequities. The first kind includes systematic biases in compensation schemes (Herzog, 2008). Previous salary studies have found systematic bias in wages based on gender (Alkadry & Tower, 2006; Binder, Krause, Chermak, Thacher, & Gilroy, 2010; Ginther, 2004; Ginther & Hayes, 2003; Rabovsky & Lee, 2017; Sneed, 2007), race and gender (Webber & González Canché, 2015), or even loyalty to an organization (Masakure, 2016). The second type of concern is for fairness at the individual rather than group level. Some individuals may have lower salaries for idiosyncratic reasons unrelated to productivity or the usual demographic suspects. For example, an individual who had an unusually productive year during a time when the organization was experiencing a salary freeze could earn less than another individual whose similarly productive year occurred when managers were less constrained. This type of inequity—which may have nothing to do with any systematic bias—may be particularly pernicious when raises are percentage-based because small differences early in the career of an employee will become magnified over time.

A well-constructed salary study can help an organization determine whether either type of inequity exists so that ameliorative actions can be taken. In contrast, a simple comparison of means or a poorly constructed study may not uncover the above inequities or may incorrectly imply inequities when none exist. This article draws on the literature from human resource management, economics, and public administration to provide a practitioner's guide to conducting a reliable salary review to provide the state of the art in salary analysis. Throughout, we illustrate our points using a recently conducted salary analysis for a large public university in the United States.

Salary Studies in the Literature

Typically, salary equity studies use regression techniques to model the relationship between a measure of compensation, a demographic of interest (such as sex or race), and an array of other possible explanatory factors, which are referred to as controls. Controls included in the salary analysis help rule out alternative explanations for the patterns in the data. (A supplemental table highlighting several examples from the literature and describing the controls used therein can be downloaded here: https://bush.tamu.edu/research/faculty/projects/.)

Credible salary studies include controls for differences in worker productivity. Some studies have used direct measures of productivity such as numbers of publications (Barbezat, 2004; Binder et al., 2010; Ginther, 2003; Renzulli, Reynolds, Kelly, & Grant, 2013), grant dollars received (Binder et al., 2010; Renzulli et al., 2013), or student performance on standardized tests (Meier & Wilkins, 2002). A few have used worker reports on their own productivity (e.g., Langbein & Lewis, 1998). Most have used indirect measures of productivity such as the worker's highest level of education (e.g., Bidwell, 2011), rank (e.g., Binder et al., 2010), or certification status (e.g., Alkadry & Tower, 2006). Years of experience (either overall or with the organization) has been used to control for productivity differences that arise from learning by doing or on-the-job training (e.g., OPM, 2014). Worker age or years since degree have been sometimes used as imperfect proxies for worker experience (e.g., Bidwell, 2011; Lewis, Boyd, & Pathak, 2018; Mincer, 1974; Webber & González Canché, 2015).

Most salary studies control for differences in the tasks workers are asked to perform. Such controls for job characteristics could include indicators for administrative or supervisory responsibilities (Barbezat, 2004; Binder et al., 2010; Carlin, Kidd, Rooney, & Denton, 2013; Ginther, 2003), law enforcement officer status (OPM, 2014) or occupational category, or departmental affiliation.

Finally, because many scholars have found that wages are higher in locations with a high cost of living or a lack of desirable amenities such as good climate or low taxes (Taylor, 2015), salary studies focusing on organizations with multiple locations need to control for the extent to which those differences in location drive differences in wages. For example, the salary study conducted by the OPM (2014) included controls for duty station.

The type of data used in these regression analyses varies, depending on the purpose of the study, the scope of the study (e.g., specific institution, state-wide study, or national study), and availability of data. Some studies rely on internal or nationally representative surveys, which are particularly effective in detecting industry-wide salary biases. For example, several recent studies have used data from the decennial U.S. Census or annual American Community Surveys to examine wage gaps based on both gender and race in the public sector (e.g., Lewis, 2018; Lewis, Pathak, & Galloway, 2018). Several studies used nationally representative surveys of specific populations such as the U.S. National Study of Postsecondary Faculty (Barbezat, 2004; Renzulli et al., 2013), the U.K. National Health Service Survey of Nursing (Pudney & Shields, 2000), or the U.S. Survey of Doctorate Recipients (Ginther, 2004; Ginther & Hayes, 2003; Webber & González Canché, 2015). Other researchers have analyzed survey data gathered either by professional organizations (Kenyon, 1997; Langbein & Lewis, 1998) or surveys of program graduates (Kenyon, 2003). Finally, some scholars have used study-specific survey instruments (Alkadry & Tower, 2006; Hobbs, Weeks, & Finch, 2005).

Administrative data such as payroll records are preferred for analyses of salaries within an organization. Examples include data covering the entire public sector in

Queensland, Australia (Bradley, Green, & Mangan, 2015), data from more than 1,000 Texas school districts (Meier & Wilkins, 2002), or university data (Binder et al., 2010; Carlin et al., 2013). Administrative data typically cover all personnel within an organization and, therefore, are not subject to sampling or reporting biases. Such records typically include current compensation, date of hire, percentage time, department, position or job description, and basic demographics like race and sex or gender.

Some researchers have access to multiple years of administrative data. When these data follow the same individuals over time, analysts can use panel data techniques to control for systematic, yet unobserved sources of salary variation. Controls for individuals using "fixed effects" are generally not used in this literature because when these controls are included, the effects of sex and race (which typically do not vary over time for individuals) cannot be estimated. However, scholars who fully exploit the panel nature of their data can use random effects for individuals. Random effects models partially control for unobserved individual characteristics but still allow for the estimation of the effects of group characteristics of interest. For example, see Binder et al. (2010).

Theory Into Practice: An Illustrative Example

To fix ideas and illustrate our recommendations, we used the techniques described in this article to analyze faculty salaries in a large, public, Research 1 (R1) university. We collected 10 years of administrative data on compensation, demographics of interest, and essential controls for all tenured and/or tenure track faculty members across the university. We present analysis from a naturally clustered but unnamed subset of departments, all of which reported to a single administrator. Table 1 describes the demographic variables included in our salary models while Table 2 presents descriptive statistics. While this example helps make our recommendations more concrete, the data, modeling, and interpretation issues we describe generalize to all types of public sector workers.

Dependent Variable: Earnings

To compare earnings between people with different levels of time obligation, we recommend inflating part-time salaries into full-time equivalents (FTEs), or what they would be paid if they were working full time instead of part-time. Either FTE monthly salaries or FTE annual salaries can be used. The FTE monthly salary for a person who works 60% of a full time job (i.e., 3 out of 5 days per week) is his or her monthly salary divided by 0.60. With this approach, additional adjustments to monthly salaries for employees who only work part of the year are not necessary, because they have FTE earnings while working and no time commitment for months they do not work. Thus, the annualized FTE salary for a seasonal worker—like a school teacher, lifeguard, or university professor—would be 12 times the FTE monthly salary. Our analysis used FTE monthly salaries for the month of October.

Table 1. Demographic Variables Included in the Salary Models.

Variable type	Specific variables
Sex Indicators	Male Faculty Indicator
	Male Associate Professor Indicator
	Male Full Professor or Above Indicator
	Male Faculty Indicator × Time Trend
	Male Faculty Indicator × Post 2011 Time Trend
Faculty Rank Indicators	Associate Professor Indicator
	Dean/VP Indicator
	Distinguished Professor Indicator
	Full Professor
Faculty Race/Ethnicity Indicators	Anglo Faculty Indicator
	Asian Faculty Indicator
	Foreign-born Faculty Indicator
Administrative Status Indicators	Current Administrator
	Ever Administrator; 2000-01 through 2015-16
Years Since Degree Indicators	Years Since Degree; and its square
	Years Since Degree Unknown Indicator
Years of Service at Current Employer	Years Since First Hired
	Break in Service Indicator
Annual Term Length	Months Under Contract per Year (9 To 12)
Rank at Hire Indicators	Hired as Advanced Assistant Professor
	Hired as Associate Professor
	Hired as Full Professor
	Rank at Hire Unknown
Highest Degree Held Indicators	MA
	Special Degree
	Dual Degree
Certification Indicators	Board Certified
	Dual Certified
Change In Status Indicators	Newly Tenured
	Other Promotion
	Other Title Change
Degree Field Indicators	Indicators for individuals with doctorates in fields that are salary outliers (accounting; business administration; economics and management)
Departmental Indicators	Indicator Variables for Each Department
Department Trends	Department Indicators × Time Trend
Year Indicators	Year Indicators
Rank-Specific Year Indicators	Associate Professor Indicator × Year Indicators
	Full Professor or Above Indicator × Year Indicators

Ideally, one would like to measure total compensation, not just salary. However, it can be difficult to put a dollar figure on the value of health insurance benefits or

Table 2. Descriptive Statistics, 2007-2008 Through 2016-2017.

Variables	*M*	*SD*	Minimum	Maximum
Monthly Salary (log)	9.275	0.351	8.658	10.385
Male	0.649	0.477	0.000	1.000
Male Associate	0.169	0.375	0.000	1.000
Male Full Plus	0.349	0.477	0.000	1.000
Associate Professor	0.294	0.456	0.000	1.000
Dean/VP	0.009	0.092	0.000	1.000
Distinguished Professor	0.009	0.095	0.000	1.000
Full Professor	0.397	0.489	0.000	1.000
White	0.702	0.458	0.000	1.000
Asian	0.135	0.342	0.000	1.000
Foreign Born	0.193	0.395	0.000	1.000
Current Admin	0.035	0.183	0.000	1.000
Ever Admin	0.122	0.328	0.000	1.000
Years Since Degree	19.368	12.970	1.000	53.000
Years Since Hire	15.574	12.036	1.000	51.000
Break in service	0.035	0.185	0.000	1.000
Term Length	2.214	0.059	2.197	2.485
Hired as Advanced Assistant	0.069	0.254	0.000	1.000
Hired as Associate	0.047	0.212	0.000	1.000
Hired as Full	0.040	0.196	0.000	1.000
Rank at Hire Unknown	0.003	0.051	0.000	1.000
MA or Less	0.007	0.081	0.000	1.000
Special Degree	0.009	0.095	0.000	1.000
Newly Tenured	0.031	0.173	0.000	1.000
Other Promotion	0.022	0.145	0.000	1.000
Title Change	0.003	0.051	0.000	1.000

pension accruals. Fortunately, within an organization, benefits packages tend not to vary across individuals within ranks and departments. All clerical staff are likely to have the same benefits package as one another, and all administrators are likely to have similar packages. Any variation in compensation that is common to all employees in a unit can be captured using departmental controls and will, therefore, have no influence on the ability to detect systematic biases in salary.

Independent Variables of Interest

The first set of independent variables looks at specific demographic differences of interest to the organization. Demographics of interest can include sex, race, ethnicity, age, or combinations of these. Tenure at the organization or rank at hire may be other variables of interest, if there is concern about a “loyalty tax” (Barbezat, 2004).

Our illustrative model focused on sex differences, so we included indicators for sex, sex interacted with faculty rank, and (in some models) sex interacted with a linear time trend.[2] The interaction between sex and faculty rank allowed for differences in salary between male and female faculty at the different ranks: assistant professor, associate professor, and professor. An interaction between sex and a time trend allowed differences between male and female salaries to widen or narrow over time (which is a modeling option that is only available with longer data panels). An organization interested in looking at race differences would instead substitute race for sex.

It is common to include worker demographics as controls, even if they are not specific variables of interest. In our illustrative model, we included indicators for White/non-Hispanic, Asian, foreign-born status, and rank at hire as additional demographic controls. (Greater racial detail was not possible given the small number of faculty of color in many departments.)

Other Essential Controls

The second set of demographic variables are used as controls for things that may legitimately explain salary differences—differences in worker productivity, differences in job characteristics, and differences in location characteristics. Such factors capture plausible alternative explanations for why two people have different earnings.

In our example model, we controlled for faculty rank, current and previous administrative status, a quadratic for years since degree, years of service and whether or not there was a break in service, whether the annual contract is 9, 10, 11, or 12 months, rank at hire, highest degree held, certification, whether there has been a promotion or title change, indicators for doctorates in fields that are high salary outliers, departmental indicators, year indicators, Rank × Year interactions, and, in some cases, department time trends. The Rank × Year interactions allowed salary growth to be systematically different for the three ranks: assistant professors, associate professors, and professors. Departmental time trends allowed salary growth to be higher in some departments than in others. The illustrative analysis focused on a single campus, so there was no need to control for location characteristics.

In an ideal world, we would have included direct controls for productivity. Unfortunately, research productivity and teaching excellence (primary responsibilities of university faculty) are difficult to measure consistently across academic units. (See Fairweather, 2005, for further discussion.) There is no consensus on the validity of available indices of research productivity, such as citation counts or Academic Analytics scores. Teaching loads that would be considered equitable in one department could be inequitable in another. Similarly, a large lecture or small section may mean something different in a humanities class than it does in an engineering class. Rather than introduce noisy and potentially unreliable measures of research productivity and teaching excellence, we have excluded these variables from our example analysis.

As a general rule, public sector managers will not have access to a full set of direct measures of worker productivity, and the measures they have may be of sufficiently poor quality that they cannot be used. Fortunately, all is not lost if a complete set of productivity measures is unobtainable, so long as the people using the results of the salary model are informed of their importance, that they are missing from the model, and how to interpret the results to account for the lack of productivity measures. Indeed, it is better to exclude poorly measured productivity and allow managers or managerial committees to apply their own filters than it is to measure productivity in ways that can introduce bias to the model. We discuss how to deal with this concern in more detail in the section on how to use the results of the salary study.

Some measures of productivity may be problematic to use in a salary study even if they can be measured. In our university example, measurable attributes of teaching productivity include student evaluations of teaching, which have been shown to be biased in ways that disproportionately and negatively affect underrepresented minorities and women (e.g., Ho, Thomsen, & Sidanius, 2009). Other scholars have found evidence of racial differences in performance appraisals in the public sector (e.g., Lewis, 1997). We recommend caution when using subjective performance evaluations as a productivity measure, and recommend excluding the metric whenever there is a suspicion that raters may exhibit conscious or unconscious bias regarding the population of interest.

Note also that the number of controls that can be included in the model is a function of the size of the data set. If the department is too small to support the full model or if the department does not include any variation in one of the controls, then by necessity, these departments will use an abbreviated model with fewer controls. For example, we were unable to run regressions using race controls or time trends for either the policy school or the business school at the university in our example. When sample size limits the number of possible controls, it is more important to include well-regarded proxies for productivity (such as years of experience) than to include other demographics (such as age). While it is impossible to have a perfect model with every possible explanatory variable, if the manager is aware of what has been included in the model and what is still missing, the model is still useful because he or she can use managerial discretion and knowledge to interpret the results.

Data Collection Concerns

When working with administrative personnel data, it is vital to be careful about anonymity. As a first pass, follow your institution's guidelines for dealing with confidential data. At many institutions, additional precautions may need to be taken on top of these guidelines. For example, if the dataset uses Social Security numbers as a unique identifier, it is important to strip those from the dataset and instead develop a unique employee ID (which could be a scrambled Social Security number) to follow individuals over time. Data storage is also important—keep data on a secure server or drive that is kept in a locked drawer in a locked room. Do not send personnel data over email or store it on a laptop. An additional anonymity concern is important for publicly reported data—do not report the coefficient of an indicator variable for any group that has a small population. For example, if a department in question has only one or two non-White employees, make sure that you do not publicly report the coefficient of the indicators for race because people will be able to back out information for those specific employees. Obviously, these results can be reported to those making salary decisions or who are in a need-to-know situation, but individual data should remain publicly obscured and anonymous.

Once the data are collected, it is important to examine the data to make sure that it represents what you think it does. Administrative data are often not kept with research goals in mind, so historical data may be altered to bring it up to date. For example, when we first got personnel data, what we believed to be rank at hire (assistant vs. associate vs. full professor) turned out to be current rank because the administrative variable was changed when the employee was promoted. There may also be reporting errors that need to be addressed. For example, an employee's age should never be less than his or her years of experience.

If important data pieces are missing, as in our case with historical rank-at-hire, then you will need to back-fill the data. This is an example of initial infrastructure development that will not be required in subsequent salary studies. In general, we have found it easy to contact current employees to ask them about their rank-at-hire. Contacting employees who have separated from the agency is much more difficult. Fortunately, the Internet has made it easier to find people's employment information online either through the current webpages and online resumes or CVs or, more generally, through LinkedIn pages of professionals. If this missing information is not online, it may be important to offer token compensation (e.g., a coffee gift-card) to employees who have separated from the organization in exchange for filling out a survey.

Determining the appropriate set of controls for any salary analysis is an iterative process. Factors that resonate in one context might be uninteresting or unimportant in another. We found it useful to solicit feedback from institutional units to ensure that we were capturing all the important alternative explanations for salary differences. We began with our initial best guess of what should be included in our salary models and presented it across campus to various colleges. These colleges then informed us of omitted variables important to salary determinations in their units. For example, these discussions led us to include board certification for some departments. While we feel that our salary model is the best model possible to determine salary differences at an R1 institution given the constraints of data collection, there will be important omitted variables outside of this institutional structure. In practice, it is important to talk with the departments you are studying to determine any additional necessary controls, such time spent on travel, supervisory status, job-specific knowledge, or languages spoken.

The initial period of data collection may be time consuming and intensive. However, once data collection systems have been established and data have been back-filled, the process becomes less onerous in each succeeding year. In our example, we were able to work with the personnel office so that they kept the original rank at hire variable and created a new variable for each promotion and time at promotion for data collection going forward. In addition, in our example, departments began to see more value in the process in each succeeding year and by the third year were requesting to know when results from the annual salary study would be available.

How Much Data?

Just as salaries can differ from one person to the next, they can also differ over time. In any given year, some individuals will undoubtedly earn more (or less) than can be explained by the variables in the data set. A salary difference observed in a single year could be reflection of an unusual event or unusually productive period. Although it is possible to conduct salary studies using a single year of data, such analyses are likely to both misidentify one-time anomalies as red flags, and miss subtle but persistent evidence of bias. We recommend taking a panel analysis approach and analyzing multiple years of data.

Since data for many years may be available in administrative personnel records, managers may want to use as many years as are available in a salary study. However, things change over time. Analysis of salary patterns over the last 30 years may obscure patterns that occurred only in the last 5 years. We recommend using more than 3 years of data whenever possible, and caution against using more than 20 or so years. Our example relies on 10 years of salary data.

Specification

Hedonic wage models have a long history in economics and are particularly well suited to detecting systematic bias in an organization's salaries.[3] These models use regression analysis to describe the relationship between the dependent variable—FTE earnings—and the independent variables of interest, given the other control variables. Acknowledging that there are worker characteristics that cannot be observed in even the most extensive database, a typical hedonic wage model would be structured as:

$$\ln\left(W_{idt}\right) = T_{idt}\beta + X_{idt}\delta + \gamma_t + \tau_i + e_{idt} \tag{1}$$

where W_{idt} is the FTE wage for person *i* in department *d*, in year *t*, T_{idt} is a vector of the independent variables of interest, X_{idt} is a vector of the other controls, γ_t are year dummies, τ_i are individual-specific random effects, and e_{idt} are random errors. Any common statistical package, including SAS, Stata, SPSS, R, or Microsoft Excel can be used to estimate the hedonic wage model, although SPSS and Excel do not accommodate random effects easily.

The dependent variable in the hedonic salary model is the natural log of each worker's FTE salary. This "log-levels" specification presumes that a small change in an independent variable has a common percentage change in the dependent variable. We recommend using the natural log of salary rather than the salary itself because a log-levels specification is more consistent with how salaries work. It, for example, allows the salaries in an administrative unit to all increase by a constant percentage rather than by a constant dollar amount. Furthermore, the interpretation of coefficients is straightforward. For example, a coefficient of −0.05 on an indicator variable for foreign born would indicate that foreign-born employees earn 5% less than other workers, all else equal; a coefficient of 0.12 on a continuous variable for years of experience would indicate that each additional year of experience adds 12% to the person's predicted salary, all else equal. The interpretation of coefficients for a set of indicator variables will depend on which indicators are included and which ones are dropped, so, for example, if the variable White has a coefficient of 0.04 and the variable Asian has a coefficient of 0.06 and the remaining race variable not included in the regression is "other," then White employees earned 4% more than those of other, non-Asian races, Asian employees earned 6% more than those of other non-White races, and White employees earned 2% less than Asian employees.

The individual-specific random effects (τ_i) capture the influence of persistent worker characteristics that are not observable in the data but could affect earnings, such as verbal ability or public service motivation. They cannot be estimated in a model that uses a single year of data. If these unobservables are correlated with the demographics of interest (i.e., sex, race, or years of service) then the estimated effects of the demographics of interest could be biased. As a result, multi-year models that are estimated with random effects for individuals tend to be more reliable than single-year specifications.

Illustration

This section contrasts effects of different analysis choices using our illustrative example. We began with a simple comparison of means from a single year of data (2016-2017) using our university dataset. As Table 3 illustrates, a *t*-test found no statistically significant difference between male and female salaries, by rank, but did find that the average FTE monthly salary of female faculty members ($11,496) was significantly lower than that of male faculty members ($14,459) if rank was not taken into consideration. On the other hand, an *F*-test (needed where there are more than two possible categories) indicates that the average salary for Asian faculty members ($21,735) was significantly higher than the average salary for non-Hispanic White or other faculty members ($15,587) at the full professor rank, but not at the other ranks.

Next, consider a model based on the same year of data but estimated with the full array of controls, as in the first column in Table 4. We could not estimate random effects for individuals in this model because each person was observed only once, so the random effects were dropped from Equation (1). This omission means that nothing in the model captured person-specific yet unobserved characteristics. The coefficient of *Male* was 0.037, meaning that male assistant professors earned 3.7% more than all female faculty, all else equal. However, the coefficient was not statistically significant, so we cannot confidently assert that any difference was found. The coefficient of Male × Associate was 0.040, which would indicate that male associate professors earned 3.7 + 4.0 = 7.7% more than female associate professors, on average. Again, we cannot be confident that the effect was significantly different from zero. Similarly, male full professors and higher earned 6.7% more than female full professors, but the sum of the coefficients was not statistically distinguishable from zero.[4] The hypothesis that the male indicator and all its interaction terms were jointly zero cannot be rejected because the *F*-statistic probability was 0.19, $F(3, 128) = 1.60$. With this single cross-section, we cannot say with confidence that there were salary differences by gender.

Table 3. A Comparison of Means for FTE Monthly Salaries, by Sex and Race.

	Number	*M* FTE Salary	*SD*
All Tenured and Tenure Track Faculty	158	13,371	4,956
Males	100	14,459***	5,253
Females	58	11,496	3,753
Whites	112	13,299	4,905
Asian	16	14,942	6,745
All Other	31	12,820	3,993
Assistant Professors			
Males	16	11,525*	3,063
Females	19	9,656	2,872
Whites	24	10,328	2,773
Asian	5	12,665	4,406
All Other	6	9,447	2,620
Associate Professors			
Males	28	12,650	4,294
Females	24	11,052	2,644
Whites	31	11,855	3,812
Asian	6	11,179	3,828
All Other	15	12,324	3,552
Full Professors			
Males	56	16,201	5,549
Females	15	14,536	4,518
Whites	56	15,371	5,251
Asian	5	21,735**	6,891
All Other	10	15,587	3,633

*$p < .1$. **$p < .05$. ***$p < .01$.

We also cannot say that there were significant salary differences by race. Although the coefficient on the Asian indicator was significantly different from zero (at the 10% level), the coefficient on the White indicator was not statistically significant, and we cannot reject the hypothesis that the two racial indicators (White and Asian) were jointly zero, $F(2, 128) = 1.49$, $p = .23$. On the other hand, the single cross-section did reveal evidence of a significant loyalty penalty. Each one-year increase in the number of years since a faculty member was hired was associated with a 1.4% lower salary than would have otherwise been expected. Thus, for example, the model predicted that a full professor who was hired 5 years earlier would have earned 7% more than an otherwise equal full professor who was hired 10 years earlier.

The second column in Table 4 illustrates the same hedonic wage model estimated using 10 years of data. The additional years of data have increased the sample size from 158 observations in the simple cross-section to 1,513 observations. This specification, which is typically referred to as an analysis of repeated cross-sections, did not exploit the panel nature of the data set or control for unobservable characteristics. It indicated that male salaries were significantly different from female salaries at the

Table 4. The Estimated Relationship Between Salaries and Faculty Demographics in Selected Departments, 2007-2008 Through 2016-2017.

Variables	One-year cross-section 2016-2017	10-year panel without random effects	10-year panel with random effects	10-year panel with random effects and male trend
Male	.037	.018	.057***	.067**
	(.056)	(.014)	(.021)	(.028)
Male Associate	.040	.060***	.018	.023
	(.071)	(.020)	(.019)	(.019)
Male Full Plus	.030	.048**	.011	.020
	(.071)	(.022)	(.029)	(.028)
Male * Time trend				−.001
				(.002)
Associate Professor	.165***	.154***	.193***	.192***
	(.061)	(.036)	(.017)	(.017)
Full Professor	−.856***	−.603***	.341***	.337***
	(.171)	(.044)	(.027)	(.028)
White	.038	.020*	.011	.011
	(.034)	(.011)	(.026)	(.026)
Asian	.087*	.048***	.057*	.057*
	(.051)	(.015)	(.032)	(.032)
Foreign Born	−.065*	−.035***	−.039	−.040
	(.037)	(.012)	(.025)	(.025)
Current Admin	.024	.009	.085***	.085***
	(.090)	(.025)	(.021)	(.021)
Ever Admin	.082	.118***	.120***	.121***
	(.054)	(.015)	(.045)	(.045)
Years Since Degree	.007	.009***	.023***	.023***
	(.007)	(.002)	(.004)	(.004)
Years Since Degree Squared	.000	−.000	−.000***	−.000***
	(.000)	(.000)	(.000)	(.000)
Years Since First Hired	−.014***	−.015***	−.013***	−.013***
	(.003)	(.001)	(.003)	(.003)
Break In Service	.147**	.113***	.114	.113
	(.072)	(.021)	(.070)	(.070)
Term Length	−.363	.014	−.067	−.068
	(1.483)	(.077)	(.054)	(.054)
Hired as Advanced Assistant	−.102*	−.064***	−.054	−.054
	(.059)	(.016)	(.056)	(.056)
Hired as Associate	.018	−.016	.037	.037
	(.058)	(.021)	(.042)	(.042)
Hired as Full	.069	.050*	.137**	.137**
	(.077)	(.026)	(.069)	(.069)
Rank at Hire Unknown	.347**	.230***	.153	.158
	(.157)	(.077)	(.120)	(.116)
MA or Less	.056	.096**	.062	.061
	(.159)	(.043)	(.105)	(.105)

(continued)

Table 4. (continued)

Variables	One-year cross-section 2016-2017	10-year panel without random effects	10-year panel with random effects	10-year panel with random effects and male trend
Special Degree	−.033	.032	.130*	.130*
	(.153)	(.040)	(.076)	(.076)
Newly Tenured	.012	−.060**	−.029***	−.029***
	(.087)	(.024)	(.011)	(.011)
Other Promotion	−.081	−.131***	.002	.002
	(.119)	(.027)	(.009)	(.009)
Title Change		.018	−.026*	−.026*
		(.076)	(.013)	(.014)
Year Indicators	No	Yes	Yes	Yes
Year × Rank Indicators	No	Yes	Yes	Yes
Department Indicators	Yes	Yes	Yes	Yes
Degree Field Indicators	Yes	Yes	Yes	Yes
Random Effects for Individuals	No	No	Yes	Yes
Observations	158	1,513	1,513	1,513
Number of Individuals	158	238	238	238

Note. Standard errors in parentheses. The omitted control category for race is "Non-Asian Minority." The omitted control category for rank is "Assistant Professor."
$*p < .1$. $**p < .05$. $***p < .01$.

associate (coefficient of 0.060) and full professor (coefficient of 0.048) levels, but not at the assistant professor level (coefficient of 0.018). The model also suggested that there were significant salary penalties associated with both race ("Other" made 2% less than "White" and 4.8% less than "Asian") and loyalty (each additional year since first hired decreased salary by 1.5%, controlling for everything else in the model).

It is important to note that while moving from a single cross-section to repeated cross sections often increases significance because of the greater sample size, this is not always the case. In addition, the effect on coefficient magnitudes of moving from a single cross section to a panel can go either direction. For example, while the magnitude of the coefficient on Male × Associate increased from an insignificant 0.040 to a significant 0.060, the effect on *Foreign Born* diminished from a marginally significant −0.065 to a statistically significant −0.035. The single cross-section can be affected by outliers, while the repeated cross-section can smooth out the effects of outliers over time.

The third model in Table 4 introduced random effects to control for person-level unobserved characteristics. As such, it more reliably isolated the impact of sex, race, or longevity on compensation. Thus, we can confidently say with 99% certainty that male assistant professors in our sample earned 5.7% more than did female assistant professors. Note that the coefficient on the *Male × Associate* and *Male Full Plus* terms were no longer significant. Because the *Male* variable was significant, we know that male associate professors made at least 5.7% more than did female associate professors, but we cannot say with certainty that they made 5.7 + 1.8 = 7.5% more than did female associate professors because the coefficient of 0.018 was not significantly different than zero. Similarly, male full professors made at least 5.7% more than did female full professors with 99% confidence.

Incorporating random effects into the analysis based on 10 years of data also changed our perspective with respect to race—but not with respect to loyalty. We cannot reject the hypothesis that the White (0.011) and Asian (0.057) indicators were jointly zero, but increases in the number of years since hire were still significantly associated with a 1.3% decrease in salary (controlling for everything else in the model). Furthermore, the 10-year panel with random effects indicated that full professors hired at the rank of full professor systematically earned 13.7% more than full professors hired at the rank of assistant professor and promoted from within, even if they had otherwise equal personnel files. This increased salary may reflect the need to pay more to overcome an endowment effect to get someone to move (e.g., Kahneman, Knetsch, & Thaler, 1991) or may reflect that more productive moves are more likely to occur than equally or less productive moves.

The fourth column of Table 4 explores the possibility that sex differentials have been narrowing or widening over time. This specification includes *MaleTrend*, a variable that is the interaction between the male faculty indicator and time.[5] A positive coefficient on this indicator would mean that the salary gap between male and female faculty was widening over time; a negative coefficient would mean that the salary gap was narrowing. As the model illustrates, there was no evidence that the size of the salary gap was changing over time because the coefficient on the *MaleTrend* variable (−0.001) was not significantly different from zero.

The final model in Table 4, column (4) is the most general and, therefore, preferred specification for analyses with 3 or more years of data. The data requirements for Models 2, 3, and 4 are identical. We provide Stata and SAS code for our preferred specification in a supplemental appendix available here: https://bush.tamu.edu/research/faculty/projects/.

While we believe that the final specification in column (4) using panel data and random effects is the best choice for doing a salary study, it should be noted that the results in our example are similar across the last two columns. Although the move from cross-section to panel data with random effects and the inclusion of a gender- or race-specific trend may be important in your organization, even information from a 1-year cross-section can be useful so long as you are aware of the limitations of this type of analysis.

Applications

Salary studies can be used to accomplish two important objectives. The first is to detect group-level differences in salary such as gender. The second is to detect individual differences in salary. The two goals may be valued differently at different levels of the organization. Lower levels may be more interested in the individual differences, since these managers are charged with individual salary decisions. Higher levels may be more interested in group-level differences, since these managers are charged with decisions across multiple units.

Regardless of the study objectives, it is important to recognize that unexplained differences in salary do not necessarily mean that lower paid groups or individuals are unfairly suffering wage discrimination. No empirical model is complete, and models based on administrative data may omit key variables associated with differences in productivity or costs.

Detecting Group-Level Salary Differentials

Returning to our illustrative example, recall that in our preferred specification we found that male professors were paid 6.7% more than female professors at all ranks. Although we included many controls in our example, we were unable to completely control for research and teaching productivity. Thus, the observed pattern we found (i.e., that male salaries were higher than female salaries) could be interpreted in at least three ways:

1. Female faculty members were systematically paid less than comparable male faculty members because of their sex, or
2. Female faculty members were systematically less productive than male faculty members, or
3. Female faculty members were systematically less qualified than male faculty members for some qualification that we failed to measure in the model (e.g., if we are missing the effects of a relevant type of board certification).

Any of these interpretations would be troublesome, though for different reasons. The first might indicate unlawful salary discrimination, which should be addressed directly. The second could indicate climate problems that contribute to lower productivity by female faculty members, indicating the need to study how department practices and actions are causing female faculty members to be less productive and address those core problems.[6] The third indicates the need to improve the model and collect more data, and (if there are systematic gender differences in obtaining this certification) could indicate a pipeline problem. In this case, it is important for the scholars to go back to departments to determine what belongs in the model, as well as to explore ways to provide opportunities for women to earn those certifications. The appropriate managerial response differs depending on the interpretation.

Furthermore, if employees are segregated by position, occupation, or administrative role—and the model includes controls for those segregating dimensions—then underlying differences by sex, race, or loyalty will not be detected, despite clear cause for concern (Alkadry, Bishu, & Ali, 2017; Arulampalam, Booth, & Bryan, 2007; Blau & Kahn, 2017; Smart, 1991; Smith, 2012). For example, if all the women are in one occupation and all the men in another, then a model with indicator variables for occupation would be unable to detect differences in compensation by sex. Managers must always interpret the results of a salary study within the context of an institution's distribution of employees and should confirm that the organization does not suffer from segregation along racial or gender lines before interpreting null results as evidence of equity.

Note that the earlier examples used male and female as the two groups with systematic salary differences. However, the similar scenarios could be constructed for any two groups with systematic salary differences, for example, minority and nonminority faculty members.

Further complicating the matter, none of these findings necessarily indicate that female faculty members are more likely than their male colleagues to have salaries or productivity that are unusually low. It could just as easily be the case that female faculty members are less likely than their male colleagues to have salaries or productivity that are unusually high. Differences in the findings matter because it is easier for a manager to raise salaries of underpaid individuals than to decrease those of overpaid individuals (particularly in a low inflation environment). Nor do managers want to punish productivity or reward underachievement. This consideration leads us to the goal of examining individual differences in salary. Aggregate differences may be resolved by focusing on individual differences.

Detecting Individual Salary Differentials

A salary study can be used to detect individual cases where salaries may be out of alignment. To do so, the practitioner uses the salary model to predict the earnings of each individual and then compares those predicted earnings with the actual earnings the individual received. The goal is to identify outliers—individuals whose salary is unusually high or unusually low given their characteristics. Those outliers are then brought to the attention of managers so that they can determine whether these gaps can be explained by factors not included in the salary model.

When using this approach, it is crucial not to perpetuate group-level differences. Therefore, salary predictions should be constructed assuming no estimated differences by race and sex. The easiest way to do that is to predict what each individual would have been paid had she or he been native-born, White, and male. In the illustrative example earlier, we also adjusted the salary predictions for systematic differences by rank in the average random effects.

The first step in this process is to perform a regression analysis that includes all divisions or departments within the unit or organization being studied. The second step is to use the results of that model to predict what each individual's salary would have been had the individual been native born, White, and male. Essentially, the regression results in Step 1 provide the coefficients for a linear model into which we substitute the value 1 for the indicator variables of native born, White, and male, and then plug in the actual characteristics of each person for all other variables. In models with random effects, we also add in the value of the mean random effect for each position or rank to ensure that mean predicted wages by rank track mean observed wages. We can then plot the predicted values against the actual values as in Figure 1.

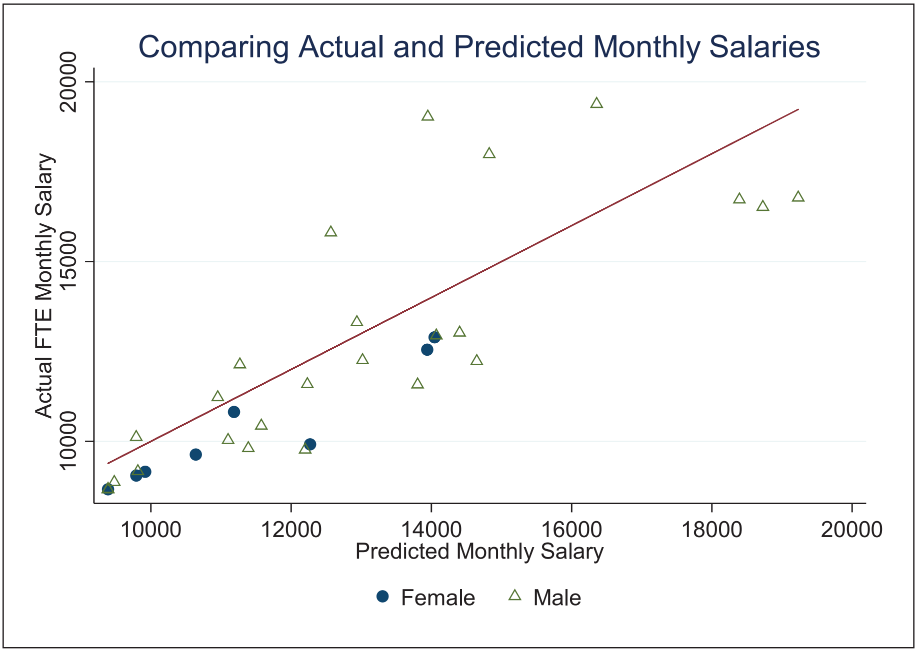

Figure 1. Selected actual and predicted salaries in a sample department.

The next step is to mark records that appear to be outliers within each division for current year salaries. To do this, we created a "normalized" or "z" score for each person's salary by taking the salary that would have been predicted for the current year were the person a White male, subtracting from that the person's observed salary, and then dividing by the standard error of the predicted salary. The equation is thus,

$$z_i = \frac{Salary_{predicted} - Salary_{actual}}{SE\left(Salary_{predicted}\right)}. \tag{2}$$

All individual salaries will deviate from the predictions to some extent, and in large organizations, it could be costly to review salaries of all the individuals. Therefore, it is necessary to choose a threshold or cut point to label outliers and flag for follow-up. Although the cut point is purely arbitrary, we chose to label as outliers the 10% of faculty members in each division with the highest absolute values of *z* score. Therefore, in each division, the 10% of records with the largest difference between actual and predicted were flagged for follow-up, as were the 10% of records university-wide with the largest difference for each faculty rank (assistant, associate, and full). We did not include people with specialized job titles whose salaries vary dramatically because of their specialized title. In practice, that meant that we included faculty records for individuals holding the rank of assistant, associate, or full professor to flag for follow-up; higher ranks, such as distinguished professors, deans, and vice presidents, were not included in this prediction exercise, though they were included in the estimation of the underlying salary model.

We do not recommend using the results from the individual salary difference calculations to mechanically increase wages because the regression results do not include a full set of productivity controls.[7] Instead, managerial discretion is necessary to interpret the outlier results. In general, the flagged results of salary regressions for individuals should be sent to someone with knowledge about individual employee productivity factors not included in the salary regressions. This person could be the employee's manager, or it could be a committee tasked with looking at compensation differences that has access to employee evaluations and other difficult-to-quantify information that affects salary differences.

In our example, each year, the dean of faculties and associate provost sends academic department heads lists of faculty members whose differentials between actual and predict salaries are in the top 10% of differences using the process described previously. This list shows the faculty member, the salary differential, and identifies whether the faculty member was identified in the previous year as having one of the largest differentials. Heads are sent only the list of those with largest salary differentials to provide the most relevant information, but if a department head requests it, salary differentials for every faculty member in the department are provided.

We found it useful to provide managers with a spreadsheet describing the actual and predicted salaries for all personnel who were flagged for follow-up as well as some graphical presentations of the data. For example, it can be helpful to share a scatter plot (by rank or occupation) of actual salaries on the vertical axis and years of service on the horizontal, with different markers for male and female employees. If there are large differences by sex or longevity, they can pop out of such a chart. This unmodeled information also reminds administrators about the range of salaries by group within their organization, and can head off criticisms that the predicted salaries are not credible.

Another powerful graphic is a scatter plot with the predicted salaries on the vertical axis and the actual salaries on the horizontal, again with different markers for individuals with different demographics of interest. For example, Figure 1 illustrates the predicted and actual salaries for male and female faculty in one of the departments included in our illustrative example.

How Our Analysis Has Affected Gender Differences

Although large, bureaucratic organizations are often resistant to change, we believe our salary analyses have made a difference. We have been conducting such analysis annually for the last 5 years. When our work began, there were statistically significant differences by sex in two of five science, technology, engineering, and mathematics (STEM) divisions. In the most recent year, we were able to detect such differences in only one.

A major contributing factor to the progress we observe was the process by which we engaged with each division. Colleges and departments receive a list of salary outliers from the office of the dean of faculties and associate provost with a request either to make sure salaries of the outliers can be justified or to adjust the salary. As an example, the Dean of Faculties and one of the coauthors met the dean of a college that was small enough that the dean would know the faculty members. The meeting was to discuss the salary survey, how it was done, and the results. The dean looked at the list of outliers who were assistant professors. For the first and second, the salary was justified based on performance. For the third, the dean said, "Wait a minute, that seems low; that needs to be investigated." Although only the dean of faculties would know whether a salary adjustment was submitted in this case, it seemed likely that an adjustment was requested for this female faculty member.

Another contributing factor may have been the transparency surrounding the analysis. Group-level results were published to the web and presented internally and externally in a variety of forums. Team members also met with individuals concerned about implications of the analysis at their request. Administrators were presented with group-level results for other units as well as their own, which facilitated internal benchmarking regarding group-level equity and (we suspect) applied additional motivation for change.

A final contributing favor may be the university's long-term and expanding commitment to salary equity studies. University administrators have recently expanded the scope of the annual analyses to include tenured and tenure-track faculty from university divisions not included in the original work. Following requests from multiple units for similar salary studies for non-tenure-track faculty members, the university has also commissioned a salary equity study for non-tenure-track faculty members. Team members are not aware of efforts by the university to conduct salary studies for employees who are not faculty members but would not be surprised if such analyses were to become part of the administrative toolkit in the future.

Conclusion

The goal of fair compensation in an organization is not perfect homogeneity of earnings. Earnings vary across individuals in most organizations. Some of that variation is explained by differences in the jobs workers perform and how well they perform them. It is expected that a high skilled worker will earn more than a low skilled worker, or that a worker who takes on unusual risks or unusual burdens is compensated accordingly. An organization that operates in multiple locations may also need to pay higher salaries where the cost of living is higher. Instead, an organization with a fair compensation schedule has no systematic variation in earnings that cannot be explained by differences in worker productivity or working conditions.

We discuss procedures for and complications inherent in conducting a salary study in a large organization. We provide suggestions for what variables to use and how to put in place long-term data collection procedures. We provide code in a supplemental, online appendix (https://bush.tamu.edu/research/faculty/projects/) to help individuals run their own salary studies, and explain with an illustrative example how to interpret the results. Finally, we discuss how to apply these results to implement changes in salaries, focusing on making it clear to decision-makers that even the best salary study will be unable to completely capture productivity differences between employees so that their judgment is still necessary even after accounting for more easily measurable factors.

Two things helped us to solicit buy-in from units likely to be affected by the salary study. First, consulting with units about what measurements are important to include in your salary study engages their participation in construction of the models and helps alleviate initial skepticism. A first attempt at a salary study may omit important factors that contribute to salary differences that is obvious to people in the unit but less so to the researchers. Once these factors are controlled for, managers are more likely to accept the results. Second, managers are more likely to appreciate the salary study over time, so plans for conducting the salary study should include multiple iterations rather than base acceptance on the initial study. In our experience, while managers were initially skeptical of our salary study, they became convinced of its usefulness both after working with the research team to include necessary variables, and after they had seen the results and had been able to implement changes.

It is generally easier to bring up artificially low salaries in an organization than it is to decrease artificially high salaries, particularly in years of low inflation. Workers do not enjoy salary cuts and, if word gets out, such cuts can decrease morale of the entire unit rather than just the affected member. Indeed, in our illustrative example, over the years, we saw many unusually low salaries being brought up closer to their predicted salary but did not see unusually high salaries lowered. While it may simply be the case that unusually high salaries capture a large amount of unmeasured productivity, managers are reluctant, with good reason, to lower salaries.

We hope that this article and its accompanying code (available in a supplemental, online appendix) are useful to organizations conducting their own salary studies. While our illustration used an academic example, the methods and procedures discussed apply to any organization interested in equity, fairness, and rewarding excellence.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Portions of this work were funded by the U.S. National Science Foundation under grants HRD-1008385. Any opinions, findings, and the conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect those of the National Science Foundation.

Notes

1. The term *salary study* here denotes the type of salary review that is concerned with equitable pay within an organization. We are not using it in the context of an assessment of how salaries compare across organizations or sectors (but see Lewis, Pathak, & Galloway, 2018, for a recent review).
2. A linear time trend is a variable that takes the value of 1 for the first year (or other unit of time) in your sample, a 2 for the second year, and so on. Whether or not to include a linear time trend in the model depends on the number of years in the data. If there are only two years, then it will be mathematically impossible to include both year fixed effects and a linear time trend. Time trends are better measured and more interesting in longer panels.
3. For example, see Bacolod and Blum (2010), Binder et al. (2010), or Ginther (2003).
4. The hypothesis that the sum of the male and Male × Associate terms are jointly zero cannot be rejected at the 10% level, $F(1, 128) = 2.62$, $p = .11$. The hypothesis that the sum of the male and Male × Full terms are jointly zero cannot be rejected at the 10% level, $F(1, 128) = 2.09$, $p = 0.15$.
5. The main effect of the time trend is captured by the indicator variables for year. Including indicator variables allows the wage level to grow more rapidly in some years than in others.
6. For example, Glover, Pallais, and Pariente (2017) find that racist managers in French grocery stores decrease the productivity of minority workers compared with nonminority workers while nonracist managers do not.
7. For a discussion of the pitfalls associated with mechanical applications of regression results from salary studies, see Haignere (2002).

ORCID iD

Lori L. Taylor https://orcid.org/0000-0002-2865-8598

Supplemental Material

Supplemental material for this article is available online.

References

Alkadry, M. G., Bishu, S. G., & Ali, S. B. (2017). Beyond representation: Gender, authority, and city managers. *Review of Public Personnel Administration*. doi:10.1177/0734371X17718030

Alkadry, M. G., & Tower, L. E. (2006). Unequal pay: The role of gender. *Public Administration Review*, *66*, 888-898. doi:10.1111/j.1540-6210.2006.00656.x

Arulampalam, W., Booth, A. L., & Bryan, M. L. (2007). Is there a glass ceiling over Europe? Exploring the gender pay gap across the wage distribution. *ILR Review*, *60*, 163-186. doi:10.1177/001979390706000201

Bacolod, M. P., & Blum, B. S. (2010). Two sides of the same coin: US "residual" inequality and the gender gap. *Journal of Human Resources*, *45*, 197-242. doi:10.3368/jhr.45.1.197

Barbezat, D. A. (2004). A loyalty tax? National measures of academic salary compression. *Research in Higher Education*, *45*, 761-776. doi:10.1023/B:RIHE.0000044230.25883.46

Bidwell, M. (2011). Paying more to get less. *Administrative Science Quarterly*, *56*, 369-407. doi:10.1177/0001839211433562

Binder, M., Krause, K., Chermak, J., Thacher, J., & Gilroy, J. (2010). Same work, different pay? Evidence from a US public university. *Feminist Economics*, *16*, 105-135. doi:10.1080/13545701.2010.530605

Blau, F. D., & Kahn, L. M. (2017). The gender wage gap: Extent, trends, and explanations. *Journal of Economic Literature*, *55*, 789-865. doi:10.1257/jel.20160995

Bradley, S., Green, C., & Mangan, J. (2015). Gender wage gaps within a public sector: Evidence from personnel data. *The Manchester School*, *83*, 379-397. doi:10.1111/manc.12061

Carlin, P. S., Kidd, M. P., Rooney, P. M., & Denton, B. (2013). Academic wage structure by gender: The roles of peer review, performance, and market forces. *Southern Economic Journal*, *80*, 127-146. doi:10.4284/0038-4038-2010.267

Equality Act 2010 (Gender Pay Gap Information) Regulations. (2017). United Kingdom. Retrieved from http://www.legislation.gov.uk/uksi/2017/172/made.

Fairweather, J. S. (2005). Beyond the rhetoric: Trends in the relative value of teaching and research in faculty salaries. *The Journal of Higher Education*, *76*, 401-422. doi:10.1080/00221546.2005.11772290

Feeney, M. K., & Rainey, H. G. (2010). Personnel flexibility and red tape in public and non-profit organizations: Distinctions due to institutional and political accountability. *Journal of Public Administration Research and Theory*, *20*, 801-826. doi:10.1093/jopart/mup027

Ginther, D. K. (2003). Is MIT an exception? Gender pay differences in academic science. *Bulletin of Science Technology & Society*, *23*, 21-26. doi:10.1177/0270467602239767

Ginther, D. K. (2004). Why women earn less: Economic explanations for the gender salary gap in science. *AWIS Magazine*, *33*, 1-5.

Ginther, D. K., & Hayes, K. J. (2003). Gender differences in salary and promotion for faculty in the humanities 1977-1995. *The Journal of Human Resources*, *38*, 34-73. doi:10.3368/jhr.XXXVIII.1.34

Glover, D., Pallais, A., & Pariente, W. (2017). Discrimination as a self-fulfilling prophecy: Evidence from French grocery stores. *The Quarterly Journal of Economics*, *132*, 1219-1260. doi:10.1093/qje/qjx006

Haignere, L. (2002). *Paychecks: A guide to conducting salary-equity studies for higher education faculty*. Washington, DC: American Association of University Professors.

Herzog, S. (2008). A four-step faculty compensation model: From equity analysis to adjustment. *New Directions for Institutional Research*, *140*, 49-64. doi:10.1002/ir.269

Ho, A. K., Thomsen, L., & Sidanius, J. (2009). Perceived academic competence and overall job evaluations: Students' evaluations of African American and European American professors. *Journal of Applied Social Psychology*, *39*, 389-406. doi:10.1111/j.1559-1816.2008.00443.x

Hobbs, B. K., Weeks, S. H., & Finch, H. J. (2005). Estimating the mark-to-market premium required to fill vacant business school faculty lines: The case of finance. *Journal of Education for Business*, *80*, 253-258. doi:10.3200/JOEB.80.5.253-258

Kahneman, D., Knetsch, J. L., & Thaler, R. H. (1991). Anomalies: The endowment effect, loss aversion, and status quo bias. *The Journal of Economic Perspectives*, *5*, 193-206. doi:10.1257/jep.5.1.193

Kenyon, G. L. (1997). Gender and income among Ontario social workers: The source of disparity. *Canadian Social Work Review*, *14*, 155-166.

Kenyon, G. L. (2003). Why does he always get to be the boy? Examining the gender income gap in social work. *Canadian Social Work Review*, *20*, 185-202.

Langbein, L. I., & Lewis, G. B. (1998). Pay, productivity, and the public sector: The case of electrical engineers. *Journal of Public Administration Research and Theory*, *8*, 391-412. doi:10.1093/oxfordjournals.jpart.a024389

Lawler, I. I. I. (1990). *Strategic pay: Aligning organizational strategies and pay systems*. San Francisco, CA: Jossey-Bass.

Lewis, G. B. (1997). Race, sex, and performance ratings in the federal service. *Public Administration Review*, *57*, 479-489. doi:10.2307/976959

Lewis, G. B. (2018). Diversity, pay equity, and pay in social work and other professions. *Affilia: Journal of Women and Social Work*, *33*, 286-299. doi:10.1177/0886109917747615

Lewis, G. B., Boyd, J., & Pathak, R. (2018). Progress toward pay equity in state governments? *Public Administration Review*, *78*, 386-397. doi:10.1111/puar.12897

Lewis, G. B., Pathak, R., & Galloway, C. S. (2018). Trends in public–private pay parity in state and local governments. *Review of Public Personnel Administration*, *38*, 303-331.

Masakure, O. (2016). The effect of employee loyalty on wages. *Journal of Economic Psychology*, *56*, 274-298. doi:10.1016/j.joep.2016.08.003

Meier, K. J., & Wilkins, V. M. (2002). Gender differences in agency head salaries: The case of public education. *Public Administration Review*, *62*, 405-411. doi:10.1111/0033-3352.00194

Mincer, J. (1974). *Schooling, experience, and earnings* (Human Behavior & Social Institutions No. 2). New York, NY: National Bureau of Economic Research.

Pudney, S., & Shields, M. (2000). Gender, race, pay and promotion in the British nursing profession: Estimation of a generalized ordered probit model. *Journal of Applied Econometrics*, *15*, 367-399.

Rabovsky, T., & Lee, H. (2017). Exploring the antecedents of the gender pay gap in U.S. higher education. *Public Administration Review*, *78*, 375-385. doi:10.1111/puar.12827

Rainey, H. G. (2014). *Understanding and managing public organizations* (5th ed.). New York, NY: John Wiley.

Reese, C. C., & Warner, B. (2012). Pay equity in the states. *Review of Public Personnel Administration*, *32*, 312-331. doi:10.1177/0734371X11428616

Renzulli, L. A., Reynolds, J., Kelly, K., & Grant, L. (2013). Pathways to gender inequality in faculty pay: The impact of institution, academic division, and rank. *Research in Social Stratification and Mobility*, *34*, 58-72. doi:/10.1016/j.rssm.2013.08.004

Smart, J. C. (1991). Gender equity in academic rank and salary. *The Review of Higher Education*, *14*, 511-525. doi:10.1353/rhe.1991.0011

Smith, R. A. (2012). Money, benefits, and power: A test of the glass ceiling and glass escalator hypotheses. *The Annals of the American Academy of Political and Social Science*, *639*, 149-172. doi:10.1177/0002716211422038

Sneed, B. G. (2007). Glass walls in state bureaucracies: Examining the difference departmental function can make. *Public Administration Review*, *67*, 880-891. doi:10.1111/j.1540-6210.2007.00775.x

Taylor, L. L. (2015). When equality is not equity: Regional cost differences and the real allocation of educational resources. In A. H. Normore, P. A. L. Ehrensal, P. F. First, & M. S. Torres (Eds.), *Legal frontiers in education: Complex law issues for leaders, policymakers and policy implementers* (pp. 247-266). Bingley, UK: Emerald Group Publishing.

Thompson, J. R. (2007). *Designing and implementing performance-oriented payband systems*. Washington, DC: IBM Center for the Business of Government.

United States Office of Personnel Management. (2014). *Governmentwide strategy on advancing pay equality in the federal government* (ES/PL/PA-2014-05). Washington, DC: United States Office of Personnel Management.

Webber, K. L., & González Canché, M. (2015). Not equal for all: Gender and race differences in salary for doctoral degree recipients. *Research in Higher Education*, *56*, 645-672. doi:10.1007/s11162-015-9369-8

White House equal pay pledge. (2016). Retrieved from https://obamawhitehouse.archives.gov/webform/white-house-equal-pay-pledge.

Author Biographies

Lori L. Taylor is Head of the Department of Public Service and Administration at Texas A&M University. She earned her doctorate in economics from the University of Rochester. Her primary research interests are educator labor markets and educational efficiency.

Joanna N. Lahey is an associate professor in the Bush School of Government and Public Service at Texas A&M University. She earned her doctorate in economics from the Massachusetts Institute for Technology. Her main research interests are employment and the economics of discrimination.

Molly I. Beck is a doctoral student in the Department of Education Reform at the University of Arkansas. She earned a master's in public service and administration from Texas A&M University. She studies teacher recruitment, quality, and retention.

Jeffrey E. Froyd is a professor in engineering education at the Ohio State University. He earned his doctorate in electrical engineering from the University of Minnesota. His main research interests are faculty development and assessment in engineering education.