Expert Rebuttal Report

Response to Dr. Ali Saad's Expert Report
on Gender and Racial Differences in Compensation at Oracle USA

Janice Fanning Madden, PhD
Econsult Corporation

August 16, 2019

D-449
449.1

Table of Contents

INTRODUCTION

In this report, I respond to the comments and analyses of Dr. Ali Saad that are included in his expert report, submitted on July 19, 2019. Dr. Saad's report responds to data and approaches that the Office of Federal Contract Compliance Programs (OFCCP) included in their Second Amended Complaint (SAC). The data and approaches in the SAC differ in numerous ways from those used in my July 19, 2019 report, "Analysis of Gender and Racial Differences in Compensation at Oracle, 2013-2018." Some of the data and approaches, however, are similar to those that I used in my report. I discuss here only those parts of Dr. Saad's report that are relevant to the data and approaches used in my report. To the extent that the data and approaches Dr. Saad reviews are not relevant to my analyses, I do not respond.

I focus on the principal issues raised in Dr. Saad's report: how to measure and analyze whether there are patterns of gender and racial differences in compensation-related outcomes for Oracle employees, and which are the appropriate employee characteristics or controls to include in analyses of gender and racial differences in such outcomes.[1] The question I analyze is whether there is evidence that women, Asian, or African American employees who come to Oracle with equivalent credentials to men or white employees have systematically different compensation outcomes, including

[1] Dr. Saad objects to the use of Medicare earnings in his report. He implies, incorrectly, that Medicare earnings do not reflect current year earnings accurately because they are decreased by varying choices of contributions to retirement plans. Dr. Saad is incorrect. Contributions to pension plans are included in Medicare earnings He also objects to Medicare earnings because they sometimes reflect earnings based on decisions made in other years, such as exercising bonus stock options. While his point is empirically accurate, and potentially an issue for measuring individual earnings, it is not an issue for measuring group differences. If we want to get average earnings for a group, the exercising of stock options would "average out" for the group by combining employees who are awarded such options but do not yet cash them with those who are cashing in from past compensation. On average, they should "even out" for the group to reflect their overall experience at getting such compensation.

whether there are gender and racial differences in Oracle's initial job assignments and promotions or subsequent assignments that contribute to the differences in current compensation. Any differences in outcomes by gender or race evident in the statistical analyses must come from one or a combination of the following reasons: (1) gender or racial differences in treatment when setting compensation; (2) systematic differences by gender or race in job assignments, or (3) systematic differences by gender or race in unmeasured characteristics *after controlling for any gender or race differences in measured characteristics*. The first two are forms of gender or racial discrimination, while the last reason ascribes gender and racial differences to a systematic inferiority in qualifications that are observed by Oracle management (but that are not recorded in the database) for women, Asian, and African American employees with the same database-recorded qualifications as men and white employees.

The principal opinions discussed in more detail below are:

- Statistical analyses of whether there is gender and racial discrimination in compensation by an employer are required to use only exogenous characteristics of employees. Exogenous characteristics are those that the employee, and not the employer, control.
- Dr. Saad's report presents no statistical analyses of gender and racial compensation discrimination using only exogenous characteristics. He does not control for the obvious exogenous characteristic of education. Rather, most of his control variables, such as job title and organizational name, reflect Oracle's decisions. These are the very decisions whose gender and racial neutrality are to be determined by the statistical analysis, and not assumed.

- Exogenous characteristics include the skills that employees have as they come to the employer, such as educational attainment and prior experience, and time at the current employer. The analyses in my prior report control for these characteristics in the estimation of gender and racial compensation disparities. Dr. Saad's report identified another potentially exogenous control variable, patent production. When I add this characteristic (which is likely to be endogenous, that is affected by Oracle assignments) to the exogenous characteristics I previously included, the estimated gender and racial compensation disparities decrease by about fifteen percent and remain large and highly statistically significant.
- Dr. Saad's use of cumulative years of leave of absence as a control effectively justifies compensation discrimination against mothers, biasing the measurement of gender compensation differences. Leaves of absence decrease experience. Adjustment of experience for the leave of absence is the correct approach to the consideration of the compensation effects of taking a leave of absence.
- While endogenous characteristics, such as Oracle's job assignments, may be used to assess the sources of gender and racial compensation disparities, they cannot be used as measures of discrimination. Dr. Saad's use of controls for time in job and organizational name of job are clearly endogenous variables. Organizational name of job is a problematic control because Oracle indicated it does not measure the product produced, labor economics theory indicates it should not affect compensation, almost all employees move between these

organizational names, and the control adds hundreds of variables to the analyses resulting in insufficient data to precisely estimate the effects of any control variables, including gender and racial disparities.

- Dr. Saad's clustering of word descriptions of jobs also creates endogenous variables, as Oracle creates the job descriptions and assigns employees to them. The sorting of one job for 500 employees into 24 word clusters is descriptive, but does not appear to meet scientific standards for explanation. The example he gives for one job title is limited in scope. The example shows that clusters are not connected to the racial disparity in compensation and have a fractional effect on the gender disparity in compensation for this one job.
- Dr. Saad's study of global career level at hire for experienced hires does not refute my findings of gender and racial differences in initial assignments.
    - His analyses include less than a fifth of employees in the compensation analyses.
    - His analyses do not control for the global career levels of the job requisitions. When global career levels of the job applied for are controlled, there are significant gender and racial disparities in assignments.
- Dr. Saad's study of starting salary for experienced hires does not refute my findings of gender and racial differences in initial assignments.
    - His analyses include less than a fifth of employees in the compensation analyses.

    - His analyses include detailed controls for job assignment at hire, which is the outcome of interest. He gets his results by including endogenous variables and not including any control for education, an exogenous variable.

- Dr. Saad's study of pay growth effectively removes the two most important sources of pay growth, job title and global career level changes, and does not control for prior pay. When all sources of pay growth are included and initial pay is controlled, women experience significantly less pay growth.

I present the bases for these conclusions in the next three sections. The first section examines the criterion for including specific characteristics or control variables in order to determine whether gender and racial compensation disparities exist. The subsequent section compares the approaches taken by Dr. Saad and me, both conceptually and empirically. The next two sections examine the role of initial job assignments and pay, and of job assignments and pay decisions subsequent to hire, on gender and racial compensation disparities.

EMPLOYEE CHARACTERISTICS INCLUDED IN THE ANALYSIS

In my July 19, 2019 report, I present a series of statistical analyses estimating compensation differences by gender and race, for each year between 2013 and 2018. After discussing the differences between a compensation analysis that explains individual differences and one that explains group differences, I report my analyses showing the effects on measured gender and racial differences of adding characteristics or controls. Specifically, I compare the gender or racial coefficients across the compensation regression analyses. Tables 1 through 3 in my report, presented in several panels, show the effects on the measured gender or racial disparity of adding controls for various characteristics. The columns of each panel for each table show the gender or racial coefficients as I add various controls to the regression analyses. The individual panels differ in the dependent variable used (Medicare compensation versus base pay versus restricted stock units), and in the employee observations included (all employees versus those with education data versus employees with records of job at hire).

My analyses establish that the compensation differences by gender and by race are not the result of differences in exogenous characteristics. Exogenous characteristics are the characteristics of Oracle employees when they arrive at Oracle (education and prior experience) and the tenure or quantity of company experience they accrue after arrival. Exogenous characteristics are characteristics that Oracle does not control, but that employees themselves control.

Table 1 of my July 19, 2019 report presents clear evidence that the measure of compensation disparities for women is not affected by experience or education controls, showing that women are comparable to men with respect to these characteristics, at least in terms of the effects of the characteristics on compensation. When I add job descriptors (for example, column 6 of Table 1(a)), the measured gender disparity falls by about a quarter, implying that women are in different job areas or fields. To the extent that this variable accurately (and only) reflects gender differences in areas of prior experience and education, it is an appropriate control. The disparities in compensation by gender, after adding the job descriptors control, remain highly statistically significant and are generally over ten percent. If the job descriptors -- based on decisions made by Oracle -- are biased in any way, then the estimated gender disparities controlling for job descriptor are understated.

Table 2 of my report presents clear evidence that the measure of compensation disparities for Asian employees is not affected by education or tenure controls, showing that Asian employees are comparable to white employees in these characteristics at least in terms of their effects on compensation. When I add age (for example, column 3 of Table 2(a)), the measured Asian disparity falls by about ten percentage points, implying that Asian employees are younger and therefore have less experience than white employees. The disparities in compensation for Asian employees, after adding the age control, remain highly statistically significant and are generally over ten percent. In contrast to the gender disparity, the racial disparity for Asian employees does not change when job descriptors are added to the analyses. Asian employees are comparable to white employees of the same experience and education in their areas of specialization. The measured Asian disparity, after controlling for the exogenous characteristics of education, experience, and area of specialization, is between 10 and 18 percent and highly statistically significant in every year between 2013 and 2018.

Table 3 of my report indicates that compensation disparities for African American employees are not affected by education or tenure controls, showing that African American employees are comparable to white employees in these characteristics, at least as weighted by the effects of the characteristics on compensation. As for Asian employees, however, the addition of age (for example, column 3 of Table 3(a)) results in a drop in the measured African American disparity of about a third, implying that African American employees are also younger and therefore have less experience than white employees. As with women, the addition of job descriptors decreases the estimated racial disparity. The measured racial disparity for African American employees, after controlling for the exogenous characteristics of education, experience, and area of specialization, is between 22 and 32 percent and statistically significant in every year between 2014 and 2018. The disparity in 2013 is not statistically significant. As explained in my July 19, 2019 report, there are simply too few African American employees at Oracle to permit precision in statistical analyses of compensation disparities between African American and white employees.

The main reasons for the differences in the estimated effects of gender and race on compensation between my report and Dr. Saad's report are differences in the employee characteristics or controls. Dr. Saad decreases the statistical power of his analyses by adding hundreds of control variables and by dividing employees into separate, smaller groupings. Specifically, Dr. Saad includes several hundred control variables for Oracle's assignment of employees to organizations and job titles (endogenous variables), but includes no controls for education (exogenous variables). Dr. Saad's studies answer a different question from the question I address in my report. Dr. Saad studies gender and racial compensation differences within a job, and does not evaluate compensation differences arising from gender or racial differences in Oracle's promotion or initial assignment decisions. The gender and racial compensation effects of job assignments cannot be analyzed with a statistical analysis that controls for job assignments (effectively assuming from the start that no such gender and racial differences exist). Dr. Saad shows that most of the gender and racial differentials in compensation are due to gender and racial differences in job assignments. I concur with this finding, as shown in my July 19, 2019 report. He assumes without any scientific testing, however, that *all* of these differences in jobs (at the level of organizational names and standard job titles) are due to the unobserved systematic productivity inferiority of women, Asians and African American employees who are otherwise identical in age, tenure, education, and job descriptors to their male or white counterparts.

I discuss endogenous and exogenous characteristics or controls in greater detail below. The next section reviews the concepts of endogeneity and exogeneity. The subsequent section reviews the reasons why endogenous controls cannot be included in any analyses, including analyses of race and gender discrimination. In the following section I discuss the endogenous and exogenous characteristics or controls used by Dr. Saad and by me and present some additional analyses that clarify the roles of these characteristics in determining gender and racial disparities.

The Concepts of Endogeneity and Exogeneity

A characteristic is considered to be "endogenous" (i.e., determined "inside") if its value is determined, at least in part, by the process (or the behavior) the statistical analysis is describing. Alternatively, a characteristic is considered to be "exogenous" (i.e., determined "outside") if its value is determined outside the process (or the behavior) the statistical analysis is describing. For example, the educational attainment of an employee is not determined by an employer's decision, so educational attainment is an exogenous attribute, determined outside of Oracle. The job title of an employee is assigned by the employer. Current job title is the accumulation of the employer's initial job title assignment and the employer's subsequent job assignments of promotion decisions. So, job title is an endogenous attribute, determined, at least in part, by Oracle's employment processes.[2]

Endogenous Characteristics Cannot Be Included

Both Dr. Saad and I want to compare women, Asian, African American, men and white employees who are "similarly situated." For these comparisons, Dr. Saad defines similarly situated persons as those whom Oracle has defined as similarly situated (endogenously), that is, as those whom Oracle has assigned to the same job and same tasks (or clusters). He does not use the clearly exogenous (to Oracle's decision-making) characteristic of education (a characteristic which employees, and not Oracle, control) to define "similarly situated."

Dr. Saad's approach to deciding which employee characteristics should be included in his analysis is circular because his approach requires the assumption that Oracle does not discriminate in job assignments, as a condition or premise for his test of whether Oracle discriminates in compensation. If there were discrimination at Oracle, then that discrimination would quite likely affect how women, Asian, African American, men, and white employees were assigned to jobs; that is, discrimination would affect how they were promoted and assigned to jobs and tasks at hire. If there were discrimination, women, Asian, African American, men, and white employees with the same relevant

[2] The standard approach to these issues in the economics of discrimination literature is discussed in David E. Bloom and Mark R. Killingsworth, "Pay Discrimination Research and Litigation: The Use of Regression," *Industrial Relations*, 21:3, (Fall 1982). They explain that only "pre-hire" characteristics of employees are "not affected by practices of the present employer…[and] not subject to the kinds of difficulties that arise in the context of analyses… which in effect control for at-hire or post-hire variables denoting job level or job type at one's present employer." (p. 326). Later, at p. 329, "The essential point is that both pay and [having a particular current job placement] are outcomes that depend on decision of the employer, i.e., they are 'endogenous.'"

skills would be assigned to different jobs and tasks. An analysis of discrimination that assumes from the start that such work assignments are nondiscriminatory (or exogenous and not endogenous) begs the question. Dr. Saad's inclusion of endogenous attributes means that his analyses are biased toward finding no discrimination when discrimination truly exists.

My approach to the inclusion of employee characteristics is fundamentally different. I make no assumption, one way or the other, about whether Oracle discriminates. I use exogenous employee attributes *that are not the result of Oracle's decisions*, but are the result of employee's decisions, to define similarly situated individuals. Women, Asian, African American, men, and white employees are similarly situated when they come to Oracle with equivalent education and work experience, characteristics that are not the result of Oracle's decisions. Although some of my analyses control for Oracle's endogenous job assignments, I perform them only to parse out the specific sources or practices that yield differential compensation by gender or race, such as compensation differences within-job versus compensation differences arising from promotion versus compensation differences arising from the initial job assignment.[3] Full and complete analyses of gender and racial differences in compensation require that there be no assumption that Oracle does not discriminate; full and complete analyses of differences in outcomes require that the statistical analysis use exogenous characteristics and not be biased by including endogenous characteristics of employees (those characteristics that are the result of decisions by the employer). There

---

[3] I also use some of Oracle's broader assignments of job (job descriptors) as measures of the field or area of education and prior experience. Implicitly, I then assume for the sake of argument that there is no discrimination in this level of assignment of employees at Oracle. If these assignments were to be affected by gender or race, then I have underestimated the compensation differentials by gender or race.

can be no prejudicial assumptions that Oracle does not discriminate in fair and accurate statistical analyses, or tests, of whether they discriminate.

As described above, an endogenous characteristic is one affected by the process under investigation, regardless of the direction of the effect. If the endogenous attribute at issue is also “tainted”—that is, the direction of the effect is clearly adverse to women or Asian or African American employees—then including that effect results in biased underestimates of the extent of the true gender and racial differences.

Identifying Exogenous Characteristics

So how do these issues affect the list of characteristics that should be included in an analysis of gender and racial disparities in compensation and initial job assignments? Because I examine whether there are unexplained gender and racial disparities that are consistent with discrimination among employees who are “otherwise the same,” I require data capable of identifying which employees are “otherwise the same” that are *exogenous* or *not potentially tainted by Oracle’s gender or race discrimination*.

In the next subsection, I describe the most obvious exogenous characteristics as used in my analyses included in my report of July 19, 2019. The second subsection discusses the characteristics that Dr. Saad used in his analyses included in his report of July 19, 2019 and that I did not.

Education, Age, and Tenure

I use educational attainment and years of experience prior to coming to Oracle,[4] as well as time at Oracle,[5] as metrics, which are unlikely to be affected by any potential discrimination by Oracle, to identify similarly situated employees. I am not using educational attainment or years of non-Oracle experience primarily as measures of productivity differentials among employees in the same job. While there is evidence that education and work experience acquired with other employers affect productivity levels within a job,[6] that is not how I use them in my analyses. I use education and non-Oracle experience along with other characteristics, including time at Oracle and job descriptors (not as job controls, but as measures of the field or area in which education and prior work experience occurred) as independent or exogenous measures of employee attributes that Oracle does not control. These measures, which are not affected by the very discriminatory behavior that we are trying to detect, define similarly situated, or "otherwise the same," employees of different races and genders at Oracle. The

---

4 I use age (and age squared) along with controls for highest degree attained and for Oracle tenure as a proxy for experience before coming to Oracle.

5 Dr. Saad criticizes the OFCCP analyses supporting SAC for not considering leave of absence time in calculating the amount of tenure at Oracle. My computation of time at Oracle, as used in my July 19, 2019 report, did remove leave time in calculating time at Oracle. I discuss below the reasons why Dr. Saad's techniques for measuring tenure at Oracle are flawed.

6 Dr. Saad's discussion of the correlation between age and compensation within a job and global career level (Software Developer IC4) at pp. 106-108 of his report is misleading because it is based on the well-known "ecological inference fallacy." As software developers (or workers within any skill category) age, the more successful ones move to higher global career levels and the less successful ones stay at lower levels. Similarly, the youngest software developers who are more talented are more likely than the less talented to have attained global career level 4. As a result, the naïve correlation of age with compensation *within* the software developer 4 job ignores the larger ecology of how movement in and out of the particular job and global career level interacts with age and compensation. The youngest software developers within a global career level are the most talented (and therefore more highly compensated) of their age group and the oldest are the least talented (and therefore less compensated) of their age group. As a result, the observation of a flatter age-compensation curve reflects the selection into and out of the job, and not the relationship of age, other things being equal, to compensation or productivity.

education measures are statistically significant, with the expected effects, in my analyses of gender and racial disparities in compensation. Table R1, for example, shows the estimated effects of various educational attainments on compensation in 2018, in the Information Technology, Product Development, and Support job functions for men and women and in Product Development job function for Asian and white employees. The coefficients are of magnitudes entirely consistent with the expectations of labor economics.[7]

Variables Included in Dr. Saad's Compensation Analyses

Dr. Saad adds four variables that I did not include in any column of my Tables 1, 2, and 3. These include data on patents, cumulative leaves of absence, time in current job, and organizational names of job assignment. Time in current job and the organizational name of current job are both clearly endogenous variables set by Oracle. They are characteristics determined by Oracle's decision-making. As endogenous variables, they cannot be used to measure the gender and racial disparities in compensation. I will discuss these two variables in more detail below as endogenous variables.

Patents. Dr. Saad uses the data on whether an employee has ever received a bonus from Oracle for receiving a patent as a control variable in his compensation analyses. As employees who develop patents are more productive than those who, given the same assignments, do not, and the innovativeness represented by patent attainment is arguably an exogenous variable to Oracle, patents, especially patents awarded before hire at

[7] Table R1 provides the education estimates for the results reported in Column 4 of my Tables 1b and 2b from my July 19, 2019 report. Please note that the estimation technique measures the effects of these degrees relative to a bachelor's degree.

Oracle, may be reasonably included in an analysis of gender and racial compensation disparities. It does not appear, however, that data on prior patents awarded were considered. Rather, Dr. Saad uses compensation data indicating whether an employee has ever received a bonus for receiving a patent as an Oracle employee. If there were no evidence of racial or gender differences in assignments to project teams, in whether members of a project team are included on a patent, and if all patent holders receive a bonus,[8] then receiving a bonus for a patent is an exogenous and therefore appropriate variable to include. If Oracle were to differentially assign women, Asian, or African American employees to project teams developing patents, then the patent bonus variable should not be included as a control in analyses of gender and racial compensation disparities. Another way to say this is, if Oracle were to assign women, Asian, or African American employees to teams responsible for cutting edge products and services at a different rate than were men and white employees, then the patent bonus variable is endogenous and should not be included. If women, Asian, or African American employees were less likely to be included in patent applications by their project teams, then the patent bonus variable should not be included as a control in analyses of gender and racial compensation disparities. If there were gender or racial differences in whether Oracle employees listed on patents are awarded a bonus, then the patent bonus variable should not be included as a control in analyses of gender and racial compensation disparities. I have produced extensive evidence of differential assignments by gender and race among Oracle employees. I cannot accept the fact that an employee at some time

[8] Oracle states that these awards are at the discretion of the Oracle patent department (see ORACLE_HQCA_0000414169-70.pdf, ORACLE_HQCA_0000414368-71.pdf, and ORACLE_HQCA_0000414372.ppts for example). I have not seen any data that allows me to determine how frequently such discretion is used.

received a patent bonus as endogenous (unaffected by Oracle's decisions) in the absence of evidence that the above standards have been met. I also note that Dr. Saad could have obtained data on patents prior to Oracle employment, a clearly exogeneous variable, from the applications materials. He did not do so.

I add this patent variable to the analyses presented in the 7th column of Tables 1(a) and 2(a) of my July 19, 2019 report. The effects of including patents on the measurement of gender and racial disparities within jobs appear on Tables R2 and R3.

Adding a control for having received a patent bonus decreases the gender disparity. To determine this, I compare the gender coefficients of column 2 (which adds a control for having received a patent bonus) with column 1 (which has the same controls with the exception of the patent bonus) of Table R2. The gender disparity decreases by about two percentage points, or by about 15 percent. The gender compensation disparity, after controlling for patents, is between 9 and 13 percent and 6 to 8 standard deviations.

Adding a control for having received a patent bonus also decreases the Asian compensation disparity. As with gender, I compare the race coefficients of column 2 (which adds a control for having received a patent bonus) with column 1 (which has the same controls with the exception of the patent bonus) of Table R3. The racial disparity decreases by between two and two and a half percentage points or by about 15 percent. The Asian compensation disparity, after controlling for patents, is between 10 and 16 percent and 5 to 8 standard deviations.

Cumulative leaves of absence. In my July 19, 2019 report, I controlled for cumulative leaves of absence by reducing the time employed at Oracle by the cumulative leave time. Time at Oracle quantifies the experience within the firm that each employee has. This experience within the firm allows employees to get more on-the-job training and therefore become more productive. Taking a leave of absence, while keeping the employee in touch with the company, removes that employee from exposure to on-the-job training. Dr. Saad does not adjust work experience measures for cumulative leaves of absence, but adds a new control variable, cumulative leaves of absence. Leaves of absence affect compensation because they reduce work experience. Women are more likely to take leaves of absence for maternity and child care leave. Dr. Saad's decision to account for leaves of absence as a separate control variable, rather than by adjusting experience controls appropriately, amounts to justifying discrimination against mothers. Women have higher cumulative days of leave of absence because they take parental leave. It is reasonable to account for any ensuing differences in exposure to on-the-job training. The use of a separate variable effectively "marks" mothers and downwardly biases the gender disparity.

Table R4 presents two panels for each of the job functions illustrated in the graph on page 86 of Dr. Saad's July 19, 2019 report. The table first reports Dr. Saad's regression results from the computer backup that he sent to explain his graphics. The graphics show that the only year with a statistically significant gender disparity is 2013 for the PRODDEV job function. The regression model details that yielded that result is reported in the first row of Table R4. The gender coefficient appears in column 1.[9] Column 2 shows the coefficient on his cumulative leave variable, which is -0.0479. Columns 3, 4, and 5 report Dr. Saad's coefficients for Oracle USA tenure, previous

[9] Dr. Saad appears to have adjusted his regression coefficients throughout his report to yield the precise percentage difference, so the coefficient of 0.0177 becomes 1.75%

experience, and total Oracle years (includes experience in acquired companies and Oracle companies that are not Oracle USA).

I note two important aspects of these various types of experience coefficients in Dr. Saad's regressions. First, the coefficients are negative, meaning that employees with more experience receive *lower* compensation than those with less experience. Second, the years of cumulative experience (effectively a motherhood control) is also negative, and at a magnitude that is a multiple of the other experience measures. If experience of any type does not increase compensation, then why would taking a leave of absence (which reduces experience) have a negative effect, and such a large one, on compensation? Why does the cumulative leave in years control have an effect that swamps all of the other experience effects? These results are consistent with the hypothesis that the cumulative leave in years control is not measuring productivity effects of taking a leave but is identifying mothers who receive less compensation. In this case, the coefficient on cumulative years of leave reflects a compensation disparity for mothers. Such a variable should not be included as a control in an analysis of gender disparity.

Dr. Saad's regression analysis underestimates the gender disparity due to the inclusion of a motherhood control. The second panel in Table R4 repeats Dr. Saad's analysis, but corrects his treatment of cumulative years of leave. This panel takes the cumulative years of leave and subtracts it from tenure at Oracle. The re-estimation, then, eliminates the cumulative years of leave variable, reformulates the tenure at Oracle variable and computes the gender disparity, which is the gender coefficient in column 1. This was done for each year, 2013 through 2018. As a result, the gender coefficient for employees in the PRODEV job function increases in absolute value and becomes statistically significant at 3 to 4 standard deviations in each year.

The next two panels on the table perform the same analyses for employees in the INFTECH job function. The results parallel those for PRODEV. The number of women (and of men) employed in INFTECH is much less than in PRODEV, however. As a result, the coefficients of all control variables are less precisely measured. Note that many of the experience controls are not statistically significant. As for PRODEV, the gender coefficient increases in absolute value when the motherhood control is removed. The gender disparity becomes statistically significant in each year except 2016. This is remarkable for a regression that includes over 100 control variables (see column 6) and only 124 to 143 women (see column 7). As I discuss below, Dr. Saad's analyses frequently reduce statistical significance by partitioning the analyses into subdivisions that lead to imprecision.[10]

In summary, education, prior experience, and tenure are exogenous variables. Having obtained a patent could be an exogenous and records of bonuses for getting a patent may reflect that characteristic. Given the evidence of gender and racial disparities in Oracle's job assignments, however, it is likely that this variable is endogenous. Cumulative leaves of absence are also exogenous, but must be considered by adjusting experience controls and not by labeling women taking parental leave.

---

[10] I also note that the experience coefficients are quite similar for PRODEV and for INFTECH, further illustrating why it is inappropriate to partition the analyses and lose precision.

Endogenous Characteristics

Oracle's assignments of employees to specific jobs, including job titles, global career levels, organizational names within Oracle and time in current job are endogenous because Oracle sets the values for these characteristics. If Oracle were truly to discriminate, such discrimination would affect the values of these characteristics, as well as affecting compensation.

In my July 19, 2019 report, I include job titles and career levels in the last analysis reported in the last columns of each panel in Tables 1, 2, and 3. I perform these particular analyses to determine the sources of compensation disparities. I calculate the share of overall compensation disparities arising from pay differences within the same job and differences in job assignments. The analyses controlling for Oracle's job assignments allow me to determine disparities within the same job. The analyses controlling only for exogenous employee characteristics provide me with the complete gender and racial compensation disparities for employees who came to Oracle with the same education and prior experience and who have the same tenure at Oracle. By subtracting the disparity within job from the total disparity, I assess the roles of compensation disparities arising from differences in job assignments and compensation disparities within jobs in creating total disparities. Both women and Asian employees of the same experience and education experienced compensation disparities within job title and global career level, although the size of the disparity was substantially smaller after taking account of the gender and racial differences in Oracle's job assignments.

D-449
449.22

In my analyses of the effects of current job assignments in my July 19, 2019 report, I did not control for time in current job or for the job's organizational name within Oracle. Dr. Saad includes these two characteristics in his analyses.

Time in Current Job.

I do not control for time in current job because the variable is a measure of promotion timing. Adding this variable to the analyses of gender and racial compensation disparities means that the estimated disparities do not include any consideration that promotions may take longer for woman, Asian, and African American employees, with the same experience and time at Oracle. As long as we recognize that this is an endogenous variable, set by Oracle, and therefore potentially reflecting discrimination, however, it can be included in an analysis to measure the effect of differences in current job assignments on gender and racial disparities.

Table R2 shows the effects of adding time in current job to the measurement of the gender disparity at Column 4. As discussed above, this table is adding variables to the analyses shown in Table 1(a) of my earlier report. By comparing the coefficients in Column 4 to those in Column 3, we can see that the time in current job has virtually no effect on the gender disparity, reducing it by between zero and 0.3 percentage points. In all years, the gender disparity within current job is between 4.2 and 5.3 percent and remains highly statistically significant at four standard deviations.

Table R3 shows the effects of adding time in current job to the measurement of the Asian-white disparity at Column 4. As discussed above, this table is adding variables to the analyses shown in Table 2(a) of my earlier report. By comparing the coefficients in Column 4 to those in Column 3, we can see that the time in current job reduces the Asian compensation disparity generally by about ten percent, or between 0.2 and one percentage points. In all years, the Asian disparity within current job is between 2 and 6 percent and remains statistically significant at two to four standard deviations for all years but 2013.

Organizational Name.

I also did not control for organizational names in my analyses in my July 19, 2019 report. While I think it reasonable to include time in current job for measuring the extent of compensation disparities coming from within job differences, I do not think it reasonable to include the organizational name for each job. There are four reasons not to include these controls, even when using endogenous controls. First, there is no reason to place equally qualified women, Asians, or African Americans who are in the same job in lower paying organizations within Oracle. Second, labor economic theory indicates that there is no reason for employees to accept less compensation because Oracle makes less money from the product produced at their organization than for the product produced at another organization by identically qualified employees. Third, because employees work in multiple organizations within the same year, organizational names are questionable indicators of productivity differences among employees.[11] Fourth, controls for organizational name add hundreds of variables to the regression analyses undermining the precision of the estimates of gender and racial compensation disparities. I discuss these reasons in more detail below.

[11] At paragraph 116 of his July 19, 2019 report, Dr. Saad describes organization (and his computer backup shows that this is organization name) as correlated "at least in a general way" with products and services worked on. Oracle represented that organization names were cost centers used for tracking various financial outcomes. Oracle went on to say that not every product and service team had an organization name identified with it (Letter of June 29, 2018 from Jinnifer Pitcher to Laura C. Bremer page 8.) If Dr. Saad wants a control for product or services produced, he should use a control variable that actually represents them. Organizational name does not.

In the absence of discrimination, I do not expect that Oracle systematically assigns women, Asians, or African American employees to those organizations within Oracle that yield less profit or are lower paying than those organizations employing men or white employees in the same job with the same education and experience. There is no reason for women, Asian, or African American employees of the same education and experience (exogenous characteristics) as men or white employees to be located in organizations within Oracle that pay them less.

I do not understand why organizational name should lead to any compensation differences among equally skilled employees. Dr. Saad correctly states that pay is a function of productivity. Productivity determines the willingness of employers to pay, or as an economist would explain, the demand for labor at a given pay level. Demand for labor alone, however, does not determine pay. Actual pay also depends on the willingness of employees to work at a given pay level, or as the economist would explain, the supply of labor at a given pay level. While the revenue or profit from a particular product affects the willingness of Oracle to allocate money to pay wages in producing the product, the observed pay to employees also depends on the willingness of Oracle employees to accept lower pay *only because they are creating a less profitable product*. There is no reason for an employee of a given skill level and ability to accept lower pay producing product "A" when the same skills are paid higher for producing product "B." Pay is determined by the intersection of demand for, and supply of, workers. Dr. Saad is ignoring universally accepted theory in labor economics. Labor economists agree that companies selling their product at less profit than do other companies hire fewer workers (because their demand for labor at each wage level is less than for more profitable companies). However, these companies must still pay the workers hired the "competitive wage" (due to the supply of labor being the same for them and for more profitable companies). According to labor economics, any lower profitability translates into fewer workers, but not into lower wages.

Organizational name is a fluid characteristic. Virtually all employees within the job functions included here work in more than one organizational name between 2013 and 2018. Table R5 lists the distribution of the number of different organizations in which the 8,658 employees worked between 2013 and 2018.

Finally, the inclusion of organizational names in the composition analyses compromise the precision with which the effects of gender and race, as well as all other variables in the analysis can be determined. I understand that it might well appear to the lay observer that it is better to be more inclusive; that is, to include all characteristics that might reasonably be expected to influence the compensation, promotions, or job assignments of individual employees at Oracle. Social scientists widely accept, however, that is simply not the case, for two main reasons:

> First, we must consider the *power* of the statistical analyses; that is, the capacity of the data available (number of observations or employees) to measure accurately the effect of each specific characteristic, as the number of characteristics included in the analyses increase while the number of observations (employees) stays the same. There is no "free lunch" in adding thousands more employee attributes to the analyses.[12] The studies or analyses should be designed

---

[12] Statistics textbooks warn against putting a large number of variables in any analysis. For example, see Mario F. Triola, *Elementary Statistics*, 9th ed. (Boston: Pearson/Addison-Wesley, 2004): pp. 545-546; Peter Kennedy, *A Guide to Econometrics*, 4th ed. (Cambridge, MA: The MIT Press, 1998): p. 95;

to provide accurate and precise statistical estimates of the effects of gender and race. Adding characteristics that do not matter (in that they do not differ by gender or race after other characteristics are included) decreases the precision, or "accuracy," of the measurement of gender and racial effects.

Second, we must consider whether each *characteristic is "endogenously" determined;* that is, whether the values of attributes included in the analyses might be affected by the very discriminatory behaviors that the statistical analyses are meant to detect. If an attribute is endogenous, then it should not be included in the analyses.

I explain these concerns in more detail below.

In statistical terms, Dr. Saad's analyses include large numbers of organizational name characteristics, including many that are irrelevant, which "use up" the observations on compensation for women, Asian and African American employees to estimate hundreds of irrelevant effects, resulting in too few observations (employees) left to estimate the effects of gender and race. Dr. Saad has added too many controls, or characteristics of workers, to the model for the effects of the characteristics of gender and race to be estimated precisely. The large number of characteristics included in his analyses arises from his decision to include full job title and organizational name and to obtain different measures of the effects of each characteristic within each job function in his analyses of gender disparities. One standard textbook on social science research, for example, reports that most researchers would recommend at least 100 observations for a

Eric A. Hanushek and John E. Jackson, *Statistical Methods for Social Scientists*, New York: Academic Press, 1977): pp. 93-94.

statistical estimate of *one characteristic* and notes that this value increases when reliable estimates for a subgroup, such as African American employees in this case, are sought.[13]

Table R6 reports the number of estimated effects of characteristics or controls and the number of women, Asian or African American employees in each of Dr. Saad's compensation analyses. The number of characteristics or controls that Dr. Saad includes in his analyses far exceed the standards of the literature, given the number of observations and of women or minority employee observations in particular. Dr. Saad's analyses "wash out" gender and racial effects by taking the relatively small numbers of women, Asian and African American employees, distributing them across the large number of irrelevant effects of attributes to be estimated, yielding too few left to measure gender and racial effects with precision.

Table R2 shows the effects of adding organizational name of current job to the measurement of the gender disparity at Column 5. As discussed above, this table is adding variables to the analyses shown in Table 1a of my earlier report. By comparing the coefficients in Column 5 to those in Column 4, we can see that the organizational name of current job reduces the gender disparity by widely varying amounts over the years. The large variation in gender coefficients across years arises from the imprecision introduced by adding over 500 additional variables to analyses including only about a thousand women. Adding organizational name of current job reduces the estimated disparity arising within the current job by between 7 and 57% or between 0.3 and 2.6

[13] See Royce A. Singleton, Jr., and Bruce C. Straits, *Approaches to Social Research* Third Edition (New York: Oxford University Press, 1999), pp. 166-169.

percentage points. In all years, the gender disparity within current job is between 2 and 4 percent and is statistically significant in only three years, 2015, 2017, and 2018.

Table R3 shows the effects of adding organizational name of current job to the measurement of the Asian disparity at Column 5. As discussed above, this table is adding variables to the analyses shown in Table 2a of my earlier report. By comparing the coefficients in Column 5 to those in Column 4, we can see that the organizational name of current job reduces the racial disparity by widely varying amounts over the years. Adding organizational name of current job reduces the estimated disparity arising within the current job by between 11 and 100% or between 0.7 and 3.6 percentage points. The Asian disparity within current job becomes statistically insignificant in all years.

Clusters.

Dr. Saad also implies that differences in job descriptions for the same job title might explain gender and racial differences in compensation. As with the job title, such descriptions are also endogenous variables that are controlled by Oracle and therefore inappropriate to use as controls for statistical tests of whether Oracle discriminates. Dr. Saad describes the differences in words used to identify tasks for employees in the Software Designer 4 job title. In addition to the endogeneity of clusters of word descriptions to Oracle decision making, there are three other problems with considering Dr. Saad's cluster analyses of Software Developer 4 job descriptions as relevant to evidence of the presence or absence of discrimination. First, Dr. Saad's analysis of clusters is descriptive and does not appear to meet standards for scientific explanations. Second, there is no basis for assuming that variations in descriptions within job titles vary systematically by race or gender. Third, Dr. Saad fails to relate these clusters to gender and racial compensation disparities. I discuss each of these problems in more detail below.

Dr. Saad claims that his cluster algorithm created the 24 clusters he identifies for Software Developer 4 job descriptions. The computer backup that he provided does not demonstrate that to be the case. While the sorting of job descriptions into a cluster was done by the computer algorithm, he appears to have arbitrarily determined that 24 clusters should be used. The basis for that determination is not clear to me. His computer backup shows he used a command to set the clusters at 24.[14] Normally, the programmer plots the word correlations on a graph and then assesses the number of clusters that best fit the data. I could not find any evidence of this having happened. Furthermore, there is evidence that Dr. Saad tried different alternatives for the number of clusters. His computer output lists alternatives for 10, 15, or 30 clusters, in addition to the 24 he reports. The bottom line is there is no quantitative or scientific basis for the number of clusters he identifies.

There is no basis for assuming that men and women, Asian and white employees, and African American and white employees in the same job title (and, in my analyses with the same educational attainment and experience) would systematically differ by race or gender in word clusters formed for the same job title. Furthermore, it is not only a race and gender difference in the distribution across clusters that matter, but the differences must also be tied to compensation. Dr. Saad is implicitly assuming that women, Asian, and African American employees systematically select into narrower job descriptions that also systematically differ in compensation from men and white employees who are the

[14] Line 47 of Dr. Saad's program uses the CUTREE function which sets (or forces) the number of clusters at 24

same in experience, education, and job title. He provides no basis for this assumption of the gender or racial inferiority of Oracle employees.

Dr. Saad's data on the Software Designer 4 job descriptions that he sorted, apparently arbitrarily, into 24 clusters include 521 men and women and 491 Asian and white employees. Table R7 shows the results of regressing race or gender alone, then race or gender and education, then race or gender and cluster, then race or gender and cluster and education, on compensation. I report the race or gender coefficients and their significance for each regression, as well as the adjusted R^2 for each regression analysis. The cluster control variable has no effect on the measurement of the racial disparity for Asian employees. The cluster variable does decrease the disparity for women by about a third.

Summary

The exogenous control variables for employee education, experience, and tenure are appropriate to include in an analysis to evaluate gender and racially discriminatory compensation practices. Being listed on a patent at Oracle may be exogenous (although job assignment evidence suggests otherwise) and, if so, appropriate to include as a control. It is less clear that getting a patent bonus is exogenous. Cumulative years of leave is not appropriate as a separate control, but should be used to adjust experience measures.

Endogenous variables that reflect Oracle decisions about employees are relevant to parsing out the sources of gender and racial compensation disparities, but bias any evaluation of their existence. Job titles and global career levels, and potentially getting a patent, describe Oracle's job placement decisions. Organizational names are problematic even as endogenous variables because they involve the addition of hundreds of control variables that undermine the precision of statistical analyses, among other problems. Dr. Saad suggests forming clusters within job titles, but does not connect them to gender and racial disparities in compensation. For the Software Developer 4 job title, clusters have no effect on the observed racial disparities and a small effect on the observed gender disparity.

COMPENSATION, INITIAL ASSIGNMENTS, AND PROMOTIONS

My analyses of gender and racial differences in compensation began with an analysis that compares men and women, and Asian or African American and white employees, who have attained the same educational degrees, are the same age, have the same amount of time (tenure) with the company, and are in jobs with the same descriptors. I use job descriptors as indices or proxies of the substantive or content areas of an employee's education and prior work experience. As explained in my earlier report, my analyses test for the total compensation disparities among employees resulting from compensation differences within job and from different jobs (due to promotion and initial assignment differences) for employees who are comparable in the characteristics that employees control and that are not the results of any potential decisions -- or potential discrimination -- by Oracle.

These analyses clearly established that there were gender and racial disparities in compensation after comparing, or grouping, employees of the same education and experience. I then developed a series of analyses to quantify the role of initial placements in the compensation differences. I analyzed initial and current job assignments.

I found gender and racial disparities in initial assignments. I found that about half of current gender differences in compensation arise from gender differences in job assignments at hire for employees of similar experience and education. I found that differences in assignments after hire as well as current compensation differentials with similar job assignments account for the other half of current compensation differentials by gender.

I found that current Asian-white differences in compensation arise almost entirely from differential job assignments by race for employees of similar experience and education. Additional differences in compensation arise from different compensation for employees with similar current job assignments.

As discussed above, if gender and racial discrimination were to exist, the gender and racial differences in compensation for employees working in the same job are expected to be substantially smaller than the compensation effects arising from gender and racial differences in promotion and initial job assignment. Because gender and racial differences in compensation within the same job would be more apparent to everyone, including employees and management, they are smaller or less likely to occur. In my analyses, I observe that gender and racial differences in compensation within the same job are smaller than racial differences in compensation arising from differences in initial assignments.

Dr. Saad performed some direct evaluations of gender and racial disparities in initial assignments and promotions. I address those studies below.

Initial Assignments

I agree with Dr. Saad that the actual jobs in which individuals are placed at various levels at Oracle have detailed, and often very specific, education and job experience requirements. All applicants of the same age, educational attainment, and specialization area are obviously not equally qualified for all of these varied positions. If one were designing a statistical analysis to assign each individual applicant to each job, each of these detailed requirements for jobs and the specific skill set of each individual would have to be included. A statistical model for assigning individuals to particular jobs would be rather silly because many of these requirements do not lend themselves to quantification and the numbers of hires are too few to allow reliable estimation of the effects of the large number of characteristics that such a model would have to include. Fortunately, I am not developing statistical analyses for assigning individuals to jobs. Rather, I am designing analyses to evaluate statistically whether Oracle systematically assigns women, Asian and African American hires to job title and global career levels in a way that is different, and inferior to, the assignments of men or white hires. For this purpose, I do not have to include all of the characteristics by which individuals, or jobs, differ. In this case, we only need to include the characteristics by which the genders or races differ.

Dr. Saad's analyses of initial assignments of new hires by gender and race do not provide the information needed to evaluate whether gender and racial disparities in job assignments at hire account for gender and racial disparities in current compensation. The problems with his analyses include:

- Dr. Saad's studies include too few of the initial job assignments of relevant employees to draw any conclusions about how initial assignments affect the compensation of the much larger groups of employees we both analyze.
- Dr. Saad's studies do not control for exogenous characteristics that plausibly differ by race or gender, including education and job descriptor, and instead include the endogenous characteristics determined by Oracle.
- Dr. Saad's analyses of whether a newly hired employee's global career level assignment was the same, higher, or lower than that of the job requisition do not control for the global career level of the requisition. When this control is added, there is evidence of gender and racial disparities in the global career level of the initial assignment relative to that in the job requisition.

I discuss each of these problems in more detail below.

Dr. Saad's analyses of initial assignments include a small subset of employees. Dr. Saad's analyses of gender and racial differentials in assignments at hire include only a minority of the assignments at hire for men and women employed in the Information Technology, Product Development, and Support job functions, or for Asian, African American, and white employees in the Product Development job function between 2013 and 2018. Table R8 reports the number of employees whose initial job assignments are analyzed by Dr. Saad. The Table also reports the total number of initial assignments made by Oracle between 2013 and 2018, indicating that Dr. Saad analyzes fewer than two-thirds of these assignments. In my analyses linking current compensation differences by gender or race to initial assignments, I include all men and women employed in Information Technology, Product Development, and Support job functions and all Asian, African American and white employees in the Product Development job function between 2013 and 2018. My analyses of current compensation and initial assignments show that initial assignments account for about half of the current gender compensation differences and the majority of the current compensation disparities for Asian employees. As Table R8 indicates, Dr. Saad's study of initial assignments includes only 20 to 27% of the initial assignments for these employees. Simply, Dr. Saad's studies of initial assignments include far too few of the relevant employees' initial assignments to determine either the extent of gender or race differences in initial assignments at Oracle, or the effects of those assignments on current compensation.

Dr. Saad does not include the relevant control variables determined by employees, but includes control variables determined by Oracle. Dr. Saad does not consider the effects of education on initial assignments of employees. Education is a characteristic determined by the employee (and not affected by Oracle's decision-making) that affects initial assignments. Dr. Saad does not use any measures of education in his analyses of initial assignments, but instead uses global career level and standard job title (both defined and used by Oracle) as non-discriminatory measures of employee qualifications. Dr. Saad's use of these controls in a study whose purpose is to test for gender and racial disparities cannot be justified. The use of these characteristics or controls as indicators of employee qualifications requires an assumption of no discrimination by Oracle. Because the purpose of the test itself is to measure discrimination, such an assumption cannot be justified in testing for discrimination in initial assignments. All of Dr. Saad's analyses of initial assignments are compromised by the failure to include education and the unjustified inclusion of Oracle's decisions on employees.

Dr. Saad fails to include a critical control variable in his analyses of global career level assignments at hire. Dr. Saad analyzes the global career level assignments by race and gender for a subset of the hires. The subset includes experienced hires who matched an Oracle job requisition. Dr. Saad argues that prospective employees generally apply for one particular job and, if hired, simply get the job for which they applied. Applicants may be offered a different job, in either a lower or higher global career level, than requested on the application.

In particular he reports that women and Asian hires applied for lower global career level jobs than did men and white hires. He also reports that women, Asian, and African American hires were equally likely as were men and white hires to be assigned the global career level of the job for which they applied.

Dr. Saad fails to take the next step, however, of determining whether jobs advertised at lower global career levels were more likely to be filled at different global levels than those at higher levels, and, if so, whether there were racial or gender differentials in the initial assignment when hired for advertised jobs at the same global career levels. Had he done so, Dr. Saad would have found evidence, for this subset of hires, that women and Asian employees received lower initial global career levels.

Charts R1 and R2 use the graphics and statistical tests that Dr. Saad used in his analyses of "Actual vs. Applied for Job Level" by gender and race, but control for the job's global career level. The charts include the three largest global career levels, IC3, IC4, and IC5, into which employees were hired.

Chart R1 shows that, for job openings at IC3, women were more likely than were men to receive a lower global career level than in the requisition, but less likely to receive a higher level. These gender differences, in isolation, are not statistically significant. For job openings at IC4, the same pattern occurs, but the gender disparity is more striking and is statistically significant in isolation. For job openings at IC5, no women received a higher level (although 6.5% of men did). With only 46 women hired into these jobs, the statistical test for difference lacks precision and is not statistically significant in isolation. Oracle hired over ninety percent of the women in individual contributor jobs, and about eighty percent of women in any job in Dr. Saad's hire dataset, into jobs advertised as IC3, IC4 or IC5. Women's disadvantage increases as the global career level of the advertised job increases. Higher global career level jobs pay more.

Chart R2 repeats the same analyses, comparing Asian and white hires. For job openings at IC3, Asian hires were more likely than white hires to receive a higher level than advertised, and less likely to receive a lower level, but the racial differences were not statistically significant. For job openings at IC4 and IC5, higher paying jobs, the racial pattern is reversed. Asians are less likely to get a higher level than the advertised job for which they applied. These racial disparities are statistically significant in isolation. Oracle hired over ninety percent of Asians in Dr. Saad's hire analysis data set into jobs advertised at IC3, IC4, or IC5. The racial disadvantage of Asian hires increases as the global career level of the advertised job increases. Higher global career level jobs pay more.

Finally, I use regression analysis to test for differences in initial assignments controlling for the "job applied for." I analyze the starting salary for the hired employees whom Dr. Saad matched to a requisition. I control for the exogenous characteristics of race or gender, age, education, and hire year, as well as the job descriptor. I also control for the global career level of the job applied for, as listed on the job requisition. The first column of Table R9 reports the gender disparity in starting pay for women relative to men with the same race, age, educational attainment, hire year, job descriptors, applying to job requisitions with the same global career level. Women average 3.8 percent less starting pay, a gender difference of 3.63 standard deviations. The second column performs the same analysis for Asian employees relative to white employees and finds Asian employees average three percent lower salaries, a racial difference of 2.52 standard deviations. The third column performs the analysis for African American employees. Because there are so few African American employees, the statistics are quite imprecise, but the average salary disparity is seven percent for African American employees of the same gender, age, educational attainment, hire year, job descriptors, applying to job requisitions with the same global career level, as white employees.

In summary, the statistical evidence on initial assignments shows disparities in the salary and the global career levels given to women, Asian, and African American hires. My July 19, 2019 report showed differences in starting salaries arising from differences in starting assignments of global career levels and from differences in starting salaries within the same job and global career level. Once I modify Dr. Saad's analyses of the small subset of hires with job requisition data available to include exogenous characteristics, such as education, and to control for the global career level of the job applied for, the evidence is consistent with gender and racial disparities in initial assignments.

Promotions and Compensation Growth

Dr. Saad discusses the OFCCP studies of growth in pay, which he relates to the SAC. I presented no direct study of pay growth in my July 19, 2019 report. Some of my studies in that report are relevant to pay growth, however. I found gender pay differentials of between 10 and 19 percent.[15] when I controlled for race, age, education, time at Oracle, current job descriptor (to indicate area of education and experience), and management. In addition, I found approximately equivalent gender differences in compensation when I use the job at hire (and its global career level) rather than those characteristics of the current job.[16] When I add current job data to the analysis including job at hire data, however, the measured gender disparity (the gender coefficient) is about half of the gender disparity when only the job at hire is included. This statistical result means that the gender disparities in current compensation decrease by more than half when controls for current job assignments are added to job assignments at hire. These changes in gender coefficients can occur only if compensation decisions subsequent to hire contribute to current compensation disparities. The gender coefficient logically can drop in the latter regression only if pay growth after hire is slower for women, in addition to the disadvantages at initial assignment.

My findings for gender disparities contrasted with my findings for the racial disparity in compensation of Asian employees. For Asian employees, I found that most of the current compensation differentials are due to the original job assignments. Once I have controlled for the job assignment at hire (including global career level), the racial

15 See, for example, Table 1a, columns 6 and 7 from my July 19, 2019 report.

16 See, for example, Table 5a, column 1 from my July 19, 2019 report.

disparity does not change with the addition of controls for current job assignment. The current disparity in compensation for Asian employees must logically arise, then, from the disparities in jobs assigned at hire and to disparities in pay within the current job. Therefore, I found no evidence that there were differences in pay growth for Asian employees, given the initial job assignment.

My findings for racial disparities in the compensation for African Americans were more similar to those for gender than to those for Asian employees. The number of African American employees, however, make it impossible to analyze pay growth with the other controls, which Dr. Saad, or I, include.

Dr. Saad's direct measurement of pay growth, with the correct control variables included, shows *the same phenomena as my indirect approach*. When I revise Dr. Saad's direct measurement of base pay growth as presented in his report at pages 125-127 using the appropriate controls, the results are consistent with the conclusions from my prior analyses, as described above. Dr. Saad regressed the controls listed under each of his pay growth tables on compensation, to get the gender and race coefficients and standard deviations listed in the last two columns of those tables.

Some of the controls he includes undermine the ability of his analysis to measure gender and race effects. First, he effectively controls for the greatest sources of pay changes (which is a problem because that is what he is trying to measure in the first instance), when he adds controls for changes in global career level and job title during the year. Changes in global career level and job title are two of the most important ways for pay to grow. When he adds these controls, his analyses of pay growth no longer include the most important sources of pay growth. He is measuring only the expectedly lower pay growth for those who do not change global career levels or job titles. In technical terms, Dr. Saad placed the "dependent variable" of pay growth or change on both sides of the equation. The explanatory variables must be "independent variables," not measures of the very outcome (dependent variable) the analysis is explaining. Second, Dr. Saad fails to control for pay level at the start of the year. There is generally a statistical tendency for "regression to the mean" (meaning that pay grows the most for the lowest paid and the least for the highest paid). It is also the case that pay growth tends to be greatest for the most recently hired and youngest workers, who are also paid less. In the end, the proof is in the results when this control is included. Prior year pay level is one of the most statistically significant variables in the analysis. The standard deviations on the coefficient for the prior year's pay control range between 7 and 18, far more than for the other 500 controls that Dr. Saad includes in these analyses.

Table R10 shows the results of Dr. Saad's pay growth analyses when we include pay growth from job changes and control for starting pay. The columns follow those in his tables. The first panel compares men and women in Information Technology, Product Development, and Support job functions; the second panel compares Asian and white employees in the Product Development job function. Women of the same experience and education as men had significantly lower pay growth in each year from 2013 through 2016, when measured in isolation. They experienced less growth in 2017 and 2018, but the difference was not significant in isolation. Asian employees of the same experience and education as white employees experienced less pay growth, which is statistically insignificant in isolation, for 2013 through 2017 and equivalent pay growth in 2018.

CONCLUSIONS

I have not changed the general conclusions reported in my July 19, 2019 report. The studies suggested by Dr. Saad, appropriately modified, strengthen those findings.

My statistical analyses are consistent with the existence of a pattern of gender and racially discriminatory compensation at Oracle. The compensation disadvantage of women is in the range of 10 to 15 percent between 2013 and 2018. These salary disparities, summarized in Table 1 of my July 19, 2019 report, are the results of gender disparities in promotions, in level of initial job assignments, and in compensation within current jobs. The compensation disadvantage of Asian employees is in the range of 10 to 18 percent between 2013 and 2018. These salary disparities, summarized in Table 2 of my July 19, 2019 report, are primarily the results of racial disparities in level of initial job assignments and in compensation within current jobs. The compensation disadvantage of African American employees is in the range of one to thirty percent between 2013 and 2018. These salary disparities, summarized in Table 3 of my July 19, 2019 report, are primarily the results of racial disparities in promotions, in level of initial job assignments, and in compensation within current jobs. The wider range of estimated disparities for African American employees is a statistical artifact of their low representation at Oracle, which decreases the precision of statistical analyses.

Janice F Madden

Janice Fanning Madden, PhD

TABLES

Table R1 Education Estimates Used to Obtain the Gender and Racial Compensation Disparities for 2018 Reported in My July 19, 2019 Report, Tables 1b and 2b					
Women-Men			**Asian-White**		
Degree	**Estimate**	**Stan. Dev.**	**Degree**	**Estimate**	**Stan. Dev.**
Doctorate	0.171	3.41	Doctorate	0.153	2.87
Master	0.077	3.44	Master	0.071	2.81
No Bachelor	-0.238	-2.49	No Bachelor	-0.096	-0.68

Table R2

2013 through 2018 Gender Differences in Medicare Earnings at Oracle Headquarters by Year, with Various Characteristics Controlled

			Controls for …									
			Race, Age, Education, Time at Oracle, Job Descriptors, Exempt, Management (1)		Adds Whether Ever Had Patent Bonus (2)		Adds Global Career Level (3)		Adds Time in Current Job (4)		Adds Organization (5)	
Year	Number of Workers	% Women	Gender Coefficient	Stan. Dev.	Gender Coefficient	Stan. Dev.	Gender Coefficient	Stan. Dev.	Gender Coefficient	Stan. Dev.	Gender Coefficient	Stan. Dev.
2013	4327	26.3%	-0.128	-9.21	-0.111	-8.17	-0.049	-4.48	-0.046	-4.23	-0.020	-1.40
2014	4279	26.4%	-0.134	-8.70	-0.114	-7.64	-0.056	-4.70	-0.053	-4.55	-0.037	-1.47
2015	4225	26.1%	-0.105	-7.57	-0.088	-6.49	-0.042	-3.87	-0.042	-3.92	-0.039	-2.32
2016	4273	25.5%	-0.119	-8.23	-0.099	-7.10	-0.046	-4.22	-0.046	-4.34	-0.033	-1.87
2017	4241	25.8%	-0.146	-8.80	-0.124	-7.69	-0.050	-4.16	-0.050	-4.21	-0.039	-3.20
2018	4019	26.2%	-0.151	-8.91	-0.127	-7.76	-0.051	-4.19	-0.049	-4.09	-0.028	-2.22

Table R3

2013 through 2018 Asian Differences in Medicare Earnings at Oracle Headquarters by Year, with Various Characteristics Controlled

			Controls for									
			Gender, Age, Education, Time at Oracle, Job Descriptors, Exempt, Management (1)		Adds Whether Ever Had Patent Bonus (2)		Adds Global Career Level (3)		Adds Time in Current Job (4)		Adds Organization (5)	
Year	Number of Workers	% Asian	Race Coefficient	Stan. Dev.	Race Coefficient	Stan. Dev.	Race Coefficient	Stan. Dev.	Race Coefficient	Stan. Dev.	Race Coefficient	Stan. Dev.
2013	3584	72.5%	-0.123	-7.27	-0.104	-6.35	-0.034	-2.58	-0.024	-1.84	-0.006	-0.32
2014	3534	73.7%	-0.177	-9.32	-0.155	-8.41	-0.070	-4.79	-0.061	-4.20	-0.054	-1.45
2015	3471	74.4%	-0.156	-9.08	-0.135	-8.05	-0.065	-4.84	-0.059	-4.42	-0.023	-0.97
2016	3470	75.9%	-0.125	-6.95	-0.102	-5.86	-0.030	-2.22	-0.027	-2.02	0.004	0.17
2017	3494	76.5%	-0.131	-6.31	-0.106	-5.33	-0.037	-2.43	-0.031	-2.04	-0.017	-1.07
2018	3300	77.4%	-0.138	-6.37	-0.114	-5.51	-0.035	-2.27	-0.033	-2.13	-0.025	-1.55

D-449
449.47

Table R4

Dr. Saad's Compensation Effects of Gender, Motherhood (Years of Cumulative Leaves), and Tenure at Oracle

	Gender (1)		Cumulative Leave in Years (2)		Tenure at Oracle (3)		Previous Experience (4)		Total Oracle Years (5)		# Control Variables	# Women
	Coef.	Stan. Dev.	Coef.	Stan. Dev.	Coef.	Stan. Dev.	Coef.	Stan. Dev.	Coef.	Stan. Dev.		
Dr. Saad's Computer Backup for Graph for PRODDEV on p. 86 of his report												
2013	-0.0177	-2.12	-0.0479	-3.30	-0.0096	-5.70	-0.0034	-6.00	0.0003	0.18	551	1123
2014	-0.0132	-1.39	-0.0602	-3.84	-0.0080	-4.22	-0.0053	-8.20	-0.0045	-2.36	527	1110
2015	-0.0142	-1.43	-0.0786	-4.73	-0.0087	-4.68	-0.0062	-9.32	-0.0059	-3.16	487	1081
2016	-0.0143	-1.48	-0.0916	-5.84	-0.0071	-4.25	-0.0063	-9.70	-0.0056	-3.26	432	1055
2017	-0.0097	-0.93	-0.1044	-6.36	-0.0080	-4.78	-0.0072	-10.29	-0.0080	-4.78	414	1052
2018	-0.0083	-0.76	-0.0910	-5.39	-0.0062	-3.74	-0.0074	-10.07	-0.0084	-4.92	368	999
Dr. Saad's PRODEV Estimation, removing Cumulative Leave in Years and Correcting Tenure at Oracle												
2013	-0.0297	-3.82			-0.0093	-5.53	-0.0032	-5.62	-0.0002	-0.12	550	1123
2014	-0.0283	-3.19			-0.0076	-4.04	-0.0050	-7.77	-0.0050	-2.67	526	1110
2015	-0.0347	-3.78			-0.0083	-4.48	-0.0059	-8.83	-0.0067	-3.56	486	1081
2016	-0.0381	-4.25			-0.0068	-4.08	-0.0059	-9.11	-0.0064	-3.67	431	1055
2017	-0.0365	-3.78			-0.0079	-4.69	-0.0068	-9.75	-0.0063	-3.63	413	1052
2018	-0.0313	-3.09			-0.0061	-3.65	-0.0071	-9.58	-0.0089	-5.19	367	999
Dr. Saad's Computer Backup for Graph for INFTECH on p. 86 of his report												
2013	-0.0340	-1.75	-0.0693	-1.55	-0.0065	-2.01	-0.0013	-1.12	0.0011	0.36	107	124
2014	-0.0348	-1.49	-0.0708	-1.19	-0.0090	-2.34	-0.0027	-1.97	0.0008	0.20	102	124
2015	-0.0367	-1.67	-0.1377	-2.21	-0.0064	-1.88	-0.0028	-2.12	-0.0031	-0.92	119	136
2016	-0.0086	-0.41	-0.1345	-2.14	-0.0030	-2.42	-0.2221	-1.00	-0.0021	-0.71	122	143
2017	-0.0313	-1.33	-0.1520	-2.31	-0.0051	-1.55	-0.0029	-2.10	-0.0030	-0.91	116	132
2018	-0.0589	-2.37	-0.0748	-1.23	-0.0032	-0.95	-0.0023	-1.67	-0.0063	-1.88	125	127
Dr. Saad's INFTECH Estimation, removing Cumulative Leave in Years and Correcting Time in Company												
2013	-0.0457	-2.50			-0.0069	-2.14	-0.0011	-0.96	0.0014	0.44	106	124
2014	-0.0492	-2.36			-0.0098	-2.54	-0.0023	-1.73	0.0014	0.38	101	124
2015	-0.0568	-2.81			-0.0076	-2.23	-0.0022	-1.71	-0.0022	-0.67	118	136
2016	-0.0267	-1.37			-0.0070	-2.25	-0.0025	-2.07	-0.0018	-0.61	121	143
2017	-0.0509	-2.30			-0.0060	-1.80	-0.0023	-1.71	-0.0026	-0.78	115	132
2018	-0.0722	-3.20			-0.0033	-1.00	-0.0022	-1.58	-0.0063	-1.88	124	127

Table R5

Counts of Oracle Employees by Number of Organizational Names of Employment between 2013 and 2018

Number of Organizational Names	Number of Employees
1	31
2	4242
3	2056
4	1254
5	612
6	297
7	105
8	37
9	17
10	5
11	1
12	1

Table R6

Counts of Employees and Control Variables in Dr. Saad's Compensation Regressions

Year	Job Functions	Groups Compared	Number of Men/White Employees	Number of Women/Asian/ African American Employees	Number of Control Variables
2013	INFOTECH	Men/Women	316	124	107
2014	INFOTECH	Men/Women	323	124	102
2015	INFOTECH	Men/Women	420	136	119
2016	INFOTECH	Men/Women	461	143	122
2017	INFOTECH	Men/Women	412	132	116
2018	INFOTECH	Men/Women	394	127	125
2013	PRODEV	Men/Women	2778	1123	551
2014	PRODEV	Men/Women	2762	1110	527
2015	PRODEV	Men/Women	2733	1081	487
2016	PRODEV	Men/Women	2754	1055	432
2017	PRODEV	Men/Women	2764	1052	414
2018	PRODEV	Men/Women	2586	999	368
2013	SUPP	Men/Women	191	42	91
2014	SUPP	Men/Women	178	42	89
2015	SUPP	Men/Women	72	31	63
2016	SUPP	Men/Women	72	23	58
2017	SUPP	Men/Women	65	20	59
2018	SUPP	Men/Women	62	21	57
2013	PRODEV	Asian/White	1037	2746	547
2014	PRODEV	Asian/White	992	2764	524
2015	PRODEV	Asian/White	937	2750	484
2016	PRODEV	Asian/White	881	2778	427
2017	PRODEV	Asian/White	849	2820	412
2018	PRODEV	Asian/White	773	2662	364
2013	PRODEV	African American/White	1037	25	375
2014	PRODEV	African American/White	992	26	359
2015	PRODEV	African American/White	937	25	329
2016	PRODEV	African American/White	881	29	298
2017	PRODEV	African American/White	849	27	289
2018	PRODEV	African American/White	773	27	251

Table R7 Compensation by Race, Gender, and Education of Dr. Saad's 24 Clusters of 521 Software Designer 4 Employees			
Control Variables	Race Coefficient	Standard Deviation	Adjusted R^2
Asian only	-0.033	-1.85	0.01
plus education	-0.038	-2.25	0.13
plus cluster	-0.036	-2.07	0.09
plus cluster and education	-0.037	-2.19	0.19
Control Variables	Gender Coefficient	Standard Deviation	Adjusted R^2
Women only	-0.035	-1.76	0.00
plus education	-0.033	-1.73	0.10
plus cluster	-0.024	-1.23	0.08
plus cluster and education	-0.021	-1.08	0.15

Table R8

Dr. Saad's Study of Initial Assignment Differences, by Gender and Race, Omit Most Employees

	Number of Records Included in Analysis		
	Men/Women	Asian/White	African American/White
Dr. Saad's Initial Assignment Study	1659	1517	338
All Employees Hired between 2013 and 2018	2819	2581	504
% included in Dr. Saad's Study	58.9%	58.8%	67.1%
Unique Employees included in Medicare Compensation Disparity Analyses	6758	5598	1381
% included in Dr. Saad's Study	24.5%	27.1%	24.5%
Unique Employees included in Base Pay Disparity Analyses	7849	6480	1620
% included in Dr. Saad's Study	21.1%	23.4%	20.9%

Table R9

2013 through 2018 Gender and Racial Differences in Starting Pay at Oracle,
Employees Matched to Job Requisitions, Controlling for
Race (Gender), Age, Education, Hire Year, Job Descriptor
and Global Career Level of Job Requisiiton

	Women Employees		Asian Employees		African American Employees
Coefficient	-0.038		-0.030		-0.070
Stan. Dev.	-3.63		-2.52		-1.06
Number	841		766		185

Table R10 Dr. Saad's Pay Growth Analysis, Including Job Changes and Controlling for Starting Pay					
Year	# Obs Used	# Protected Group	Average Pay Growth	Gender or Race Coefficient	Standard Deviation
Women Employees					
2013	4565	6578	[redacted]	-0.0039	-2.07
2014	4528	6542	[redacted]	-0.0039	-2.33
2015	4463	6478	[redacted]	-0.0038	-2.29
2016	4502	6518	[redacted]	-0.0030	-2.20
2017	4441	6458	[redacted]	-0.0019	-1.62
2018	4175	6193	[redacted]	-0.0002	0.11
			[redacted]		
Asian Employees			[redacted]		
2013	3774	2743	[redacted]	-0.0016	-0.74
2014	2745	2761	[redacted]	-0.0017	-0.90
2015	3677	2743	[redacted]	-0.0013	-0.63
2016	3653	2777	[redacted]	-0.0025	-1.58
2017	3666	2817	[redacted]	-0.0011	-0.80
2018	3421	2652	[redacted]	0.0002	0.09

CHARTS

Chart R1
Comparison of Actual vs. Applied-For Global Career Level for Men vs. Women Hires

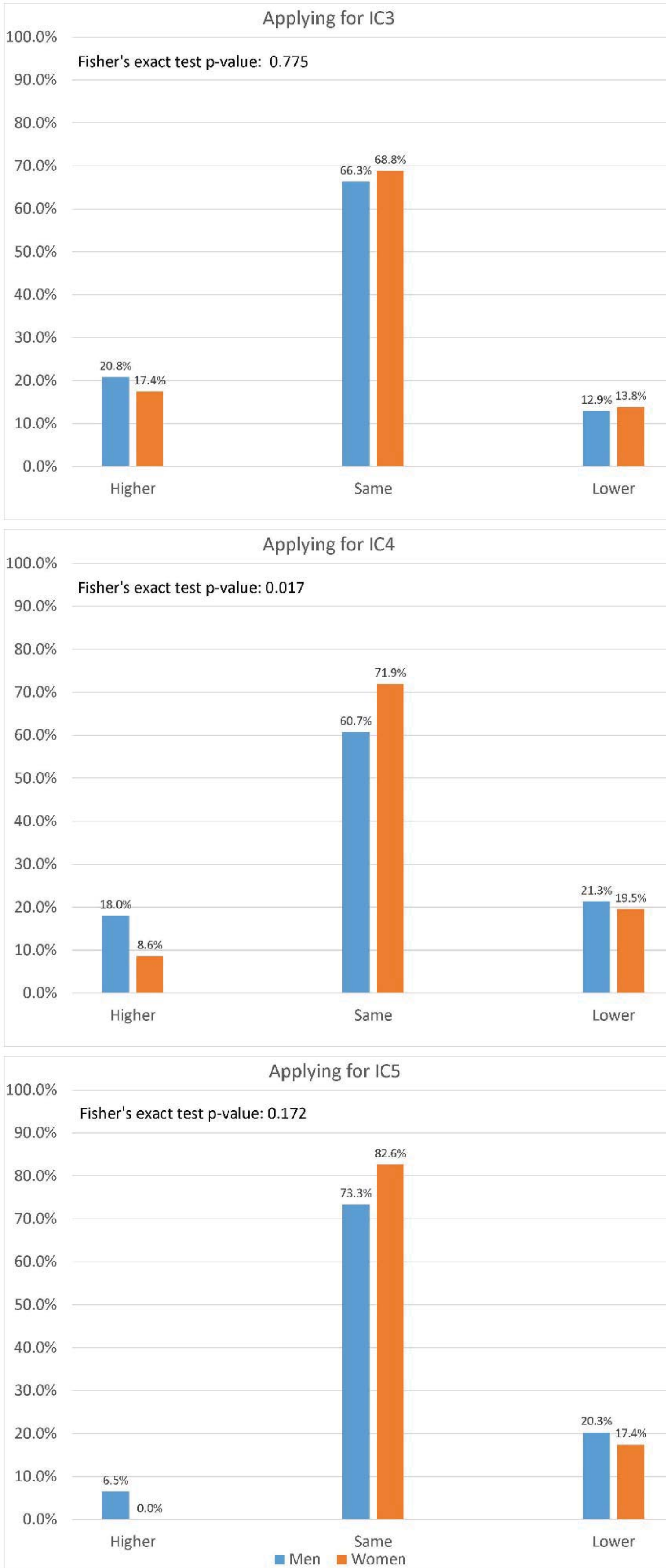
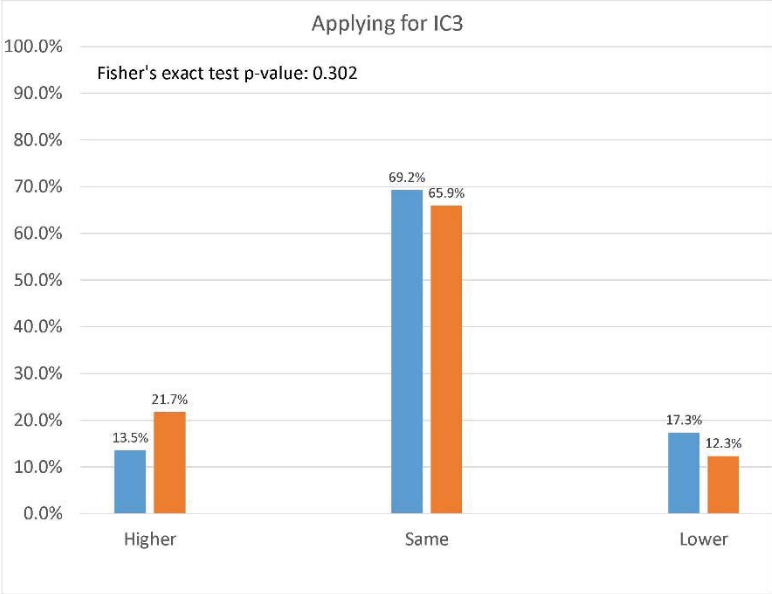
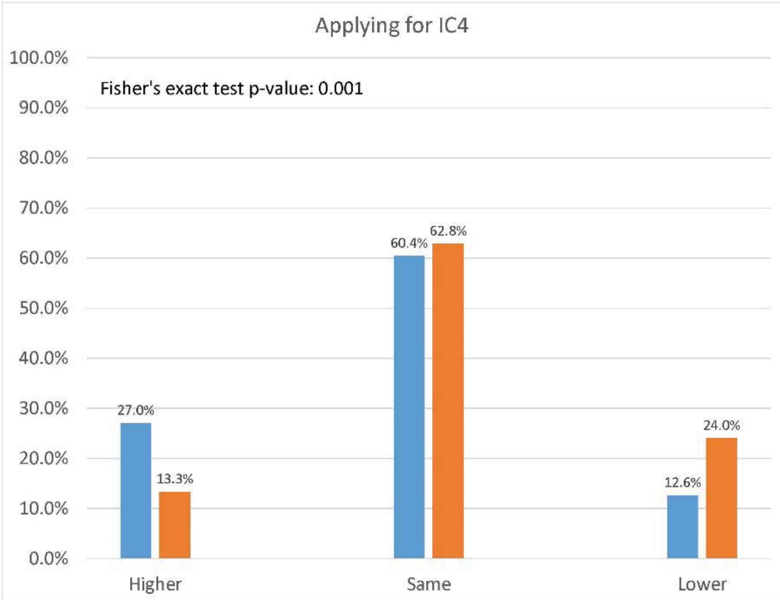
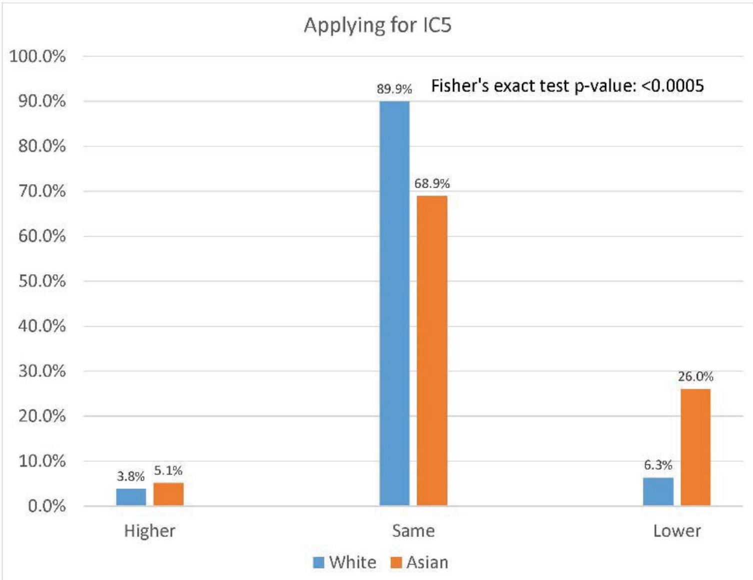


Chart R2

Comparison of Actual vs. Applied-For Global Career Level for White vs. Asian Hires







D-449
449.57