D-23

GARY R. SINISCALCO (STATE BAR NO. 64770)
grsiniscalco@orrick.com
ERIN M. CONNELL (STATE BAR NO. 223355)
econnell@orrick.com
KATHRYN G. MANTOAN (STATE BAR NO. 239649)
kmantoan@orrick.com
ORRICK, HERRINGTON & SUTCLIFFE LLP
The Orrick Building
405 Howard Street
San Francisco, CA 94105-2669
Telephone: +1 415 773 5700
Facsimile: +1 415 773 5759

JESSICA R. PERRY (STATE BAR NO. 209321)
jperry@orrick.com
ORRICK, HERRINGTON & SUTCLIFFE LLP
1000 Marsh Road
Menlo Park, CA 94025
Telephone: +1 650 614 7400
Facsimile: +1 650 614 7401

Attorneys for Defendant
ORACLE AMERICA, INC.

FILED
SAN MATEO COUNTY
MAR 06 2019
Clerk of the Superior Court
By DEPUTY CLERK

17-CIV-02669
DIO
Declaration in Opposition
1688925

SUPERIOR COURT OF THE STATE OF CALIFORNIA

COUNTY OF SAN MATEO

RONG JEWETT, SOPHY WANG, XIAN MURRAY, ELIZABETH SUE PETERSEN, MARILYN CLARK AND MANJARI KANT, individually and on behalf of all others similarly situated, Plaintiffs, v. ORACLE AMERICA, INC. Defendant.	Case No. 17CIV02669 **DECLARATION OF ALI SAAD, PH.D. IN SUPPORT OF DEFENDANT ORACLE AMERICA, INC.'S OPPOSITION TO PLAINTIFFS' MOTION FOR CLASS CERTIFICATION** Date: May 31, 2019 Time: 9:00 a.m. Assigned for all purposes to the Honorable V. Raymond Swope Department 23 Trial Date: Not Set Date Action Filed: June 16, 2017

PUBLIC REDACTED VERSION

FILED BY FAX

Exhibit P-439

I, Ali Saad, declare as follows:

1. I am the Managing Partner of Resolution Economics Group LLC, a consulting firm that performs economic and statistical analyses. As a consultant, I have extensive experience in employment matters including performing statistical and economic analyses in discrimination matters, wage and hour matters, and calculating employment damages.

2. I hold a Ph.D. in Economics from The University of Chicago, and a B.A. in History and Economics from The University of Pennsylvania. I am a member of the American Economic Association.

3. Prior to beginning my consulting practice, I worked as a professor in the economics and finance department of Baruch College of The City University of New York. I have also served as an adjunct professor in the economics department of the University of Southern California. I have taught labor economics, micro and macroeconomics, econometrics, and economic history. I also previously worked at Deloitte & Touche, LLP, where I was a Partner; Altschuler, Melovin and Glasser LLP, where I was a Partner; Price Waterhouse LLP, where I was a Senior Manager; and Olympia & York Companies (USA), where I was an Assistant VP and Senior Economist.

4. As a labor economist, I have extensive experience providing statistical and economic analyses in connection with company pay equity studies, evaluations of compensation systems, and class action employment cases. I have previously performed a number of consulting analyses involving the California Fair Pay Act and have significant experience in analyzing complex data for the purpose of litigation. I have been qualified as an expert witness in both federal and state courts.

5. I have been asked by counsel for Oracle America, Inc. to respond to the expert report submitted by Dr. David Neumark in support of Plaintiffs' motion for class certification, and Dr. Neumark's subsequent deposition testimony.

///

///

///

///

6. Attached hereto as **Exhibit A** is a true and correct copy of the report I wrote containing my findings in response to Dr. Neumark's report, methodology, and analysis.

I declare under penalty of perjury under the laws of the State of California that the foregoing is true and correct.

Executed in Los Angeles, California on March 4, 2019.

ALI SAAD, Ph.D.

EXHIBIT A

JEWETT, ET AL. V. ORACLE AMERICA, INC., **CASE NO. 17-CIV-02669**

SUPERIOR COURT OF THE STATE OF CALIFORNIA, COUNTY OF SAN MATEO

EXPERT REPORT OF ALI SAAD, PH.D.

MARCH 4, 2019

CONTENTS

ASSIGNMENT

1. I was retained by counsel for defendant Oracle America, Inc. to respond to the expert report submitted by Dr. David Neumark on behalf of Plaintiffs in the case of *Rong Jewett, Sophy Wang, and Xian Murray, et al. v. Oracle America, Inc.* Plaintiffs allege that "Oracle has discriminated against its female employees by systematically paying them lower wage rates than Oracle pays to male employees performing equal and substantially similar work under similar working conditions."[1] I was provided with electronic human resources data, payroll data, performance review system data, and other documents related to Oracle, including depositions and company policy documents, in order to conduct my assignment. My report responds to the analyses and opinions summarized in Dr. Neumark's report and associated backup materials, as well as his deposition testimony. I may supplement this report at a later date if additional relevant information is made available to me.

QUALIFICATIONS

2. I am the Managing Partner of Resolution Economics Group LLC, a firm whose activities include performing economic and statistical analyses in connection with litigation and other consulting matters. Before beginning my consulting career I was in academia as a member of the faculty of the economics and finance department at Baruch College of The City University of New York. While there I taught labor economics, micro and macroeconomics, econometrics,

[1] Fourth Amended Class Action Complaint, in the matter of *Rong Jewett, Sophy Wang, Xian Murray, Elizabeth Sue Petersen, Marilyn Clark, and Manjari Kant, individually and on behalf of themselves and others similarly situated, v. Oracle America Inc.*, Superior Court of the State of California, County of San Mateo, filed September 7, 2018, p. 2.

and economic history. In connection with my consulting, I have extensive experience providing statistical and economic analyses in connection with company pay equity studies, evaluations of compensation systems, and class action employment cases, including employment discrimination and wage and hour matters. I have also published and lectured on these topics. A particular focus of my work has involved economic and statistical analysis related to claims of systemic gender discrimination, as well as claims brought under the California Equal Pay Act. I have also performed a number of consulting analyses involving the California Equal Pay Act for companies interested in understanding their circumstances relative to the new law. In the litigation context, I have significant experience in analyzing complex data for the purpose of assisting counsel in evaluating class certification and liability. I hold a Ph.D. in Economics from The University of Chicago, and a B.A. in History and Economics from The University of Pennsylvania. I have been qualified as an expert witness in both Federal and State Courts. My resume, including all publications and testimony over the past four years, is attached to this report as Attachment A. My firm bills for my services at my current hourly rate of $750 per hour.

DATA AND DOCUMENTS

3. I was provided by Counsel with databases, depositions, and other documents. In addition, I collected publicly available data, and relied on additional secondary materials. The materials I considered in connection with my analysis and opinions are listed in Attachment B.

INTRODUCTION AND SUMMARY OF FINDINGS

4. I have been asked to respond to the statistical analyses and opinions of Dr. David Neumark as presented in his expert report dated January 18, 2019, and during his deposition that took place on February 8, 2019. In sum, it is my professional opinion that Dr. Neumark's report does not provide support for his conclusions regarding the relationship of gender to pay at Oracle, because his regression model does not compare employees performing substantially similar work from a labor economics perspective. The key flaw in Dr. Neumark's statistical work is that by uncritically adopting the Oracle job codes found in the data, he has not constructed an analytical approach that statistically controls for the nature of the work different employees at Oracle perform. I show below that there are very substantial differences between employees within the same job code and job grade. In addition, because Dr. Neumark's analysis does not compare employees doing substantially similar work, there are large and unexplained variations in pay between ostensibly equivalent employees. In addition to using inappropriate statistical controls for the nature of the work being performed, the large variation I find in employee level statistical outcomes is compounded by the fact that Dr. Neumark's analysis is aggregated in a single regression model that combines all employees working at all Oracle locations and in all years.

5. The putative class in this case is incredibly diverse, spanning hundreds of different job titles from entry-level to Executive Vice Presidents. There are 15,324 unique employees in Dr. Neumark's data, of whom 4,132 are putative class members.[2] The putative class members have

[2] Dr. Neumark reported 4,201 class members but he did not account for employees who move between job functions in his count, i.e. in Dr. Neumark's report, employees who move between the Support job function and the Product Development job function are counted twice. He has also incorrectly limited his data and includes many years not worked by employees within the proposed putative class as defined by Plaintiffs (i.e., years worked outside of California, or

occupied 180 different job codes, across 23 California locations (not counting those who work from home) and have worked in hundreds of different organizations in the company on hundreds of different products over the course of more than five years. There are 265 job code/job grade combinations[3] across all employees in the data for 2013-2018, for which annual base salaries for full time, full year employees range from [redacted] and for whom total annual compensation ranges from [redacted].

6. Against this backdrop, Plaintiffs' expert, Dr. Neumark, presents a series of multiple regression compensation models aggregated across this entire employee population in order to measure the average difference in several measures of pay between all of these men and all of these women at Oracle. A regression model can almost always be estimated if a data set is large enough; it is simply a mathematical way to look at the average relationship between female status and pay while taking other factors into account. But that does not mean a regression always yields analytically meaningful results. For the results to be meaningful – in the current context, for the estimated average gender pay difference to form the basis for a reasonable inference of discrimination – the model has to be designed to group together, by virtue of the factors used in the statistical analysis, employees working in the same or similar positions, and control for the appropriate set of other bona fide factors that influence pay.

7. Dr. Neumark regresses various pay measures on what he calls "job characteristics," which are statistical control variables including job code and job grade, zip code, line of business

---

outside of the three job functions). However, in much of my analysis, I will take Dr. Neumark's data and analytical approach at face value and use them to demonstrate the problems inherent in his approach.

[3] Dr. Neumark controls for job code/job grade combinations in his analyses. Job codes are associated with standard job titles. Job grade indicates what salary band a job title is associated with. Multiple job titles may share a grade and salary range. (Deposition of Kate Waggoner (PMK) Volume 1, July 26, 2018, Exhibit 24, Bates No. 00000407-00000409.) In the data, a job code is assigned one grade per fiscal year but may be reassigned to a new grade over time.

(defined by the head of that line of business), and part time or hourly status. Several of these controls are not particularly important. For reasons that are not explained, he uses home zip code to identify the "work location" for employees who report working at home. Part time and part year flags are not the most important variables in thinking about jobs because they reflect employee decisions about how much to work at Oracle in a year. The real action in Dr. Neumark's statistical models comes from the job code, job grade, and line of business head control variables. Thus, the fundamental question is whether job code, job grade, and line of business head are sufficient, from a labor economics perspective, to group employees doing substantially similar work. My conclusion is that they are not.

8. My analysis of Dr. Neumark's data and model leads me to believe these variables are not sufficient measures to group employees doing substantially similar work, and therefore Dr. Neumark's results do not establish any sound basis to infer that women are paid lower than substantially similar men. In addition, the statistical outcomes under Dr. Neumark's model are highly variable, and do not provide a consistent or statistically common picture of the pay outcomes for women at Oracle. On the contrary, there are substantial differences between the thousands of women. Dr. Neumark also analyzes the relationship between prior pay and starting pay at Oracle. His analysis of the relationship of Oracle starting pay to applicants' prior pay is flawed on so many levels as to be completely unreliable. Finally, there are a number of data and analytical errors in Dr. Neumark's work, which I discuss below.

9. I have several major categories of critiques of the work Dr. Neumark has presented in his report. First and perhaps most fundamentally, his model is misspecified in that it does not contain adequate controls for type of work performed, as the data and other detailed information available in this case makes clear.

a. There are enormous differences in pay within the same job title/job grade "bucket." This suggests to me – and I believe would suggest to any reasonable labor economist – that very different types of work are being performed by what can be thousands of employees sharing a common job title and grade. Yet Dr. Neumark asserts, with no further analysis or support, that "[i]ncluding this highly-detailed set of [job title] controls in my regression model allows me to compare women's and men's pay within *very narrowly defined jobs* (emphasis added)."[4]

b. The data and documents available for analysis in this case include large amounts of detailed information describing the different work performed by employees in different positions within the job codes he uses in his models, but Dr. Neumark did not utilize any of that information. This information includes thousands of job requisitions with detailed descriptions of work, as well as thousands of detailed hiring manager notes and performance reviews that would have allowed him to refine his measures of work and to test his assumption that job title "*very narrowly define[s] jobs*."

c. Because his model does not make comparisons that are "apples to apples" in terms of employees doing substantially similar work, Dr. Neumark's inference that the statistical relationship between gender and pay identified in his analysis is consistent with unequal pay for women is not scientifically supported.

10. Second, as a methodological issue, it is not clear that Dr. Neumark's regression approach is answering the question that I understand to be relevant at this stage of this case – namely, is there statistical evidence consistent with an inference that the many individual women at Oracle share a common circumstance of being underpaid relative to men who are performing substantially similar work? I use his data to predict "expected" base pay for each female and male employee using his model but excluding gender, so that gender is not among the factors predicting pay based on job and worker characteristics (such as tenure). I then compare each employee's actual pay to the benchmark pay predicted by Dr. Neumark's model. I find that:

a. Overall, about 2.2% of women earn statistically significantly <u>more</u> than his model predicts and about 3.5% earn statistically significantly <u>less</u> than his model predicts. Another 51.8% earn less than predicted but not to a statistically significant extent, and the remaining 42.5% earn more than predicted though again, not to a statistically significant

---

[4] Expert Report of David Neumark in the Matter of Jewett et al. v. Oracle America, Inc. January 18, 2019, pp 13-14, paragraph 27.

extent. The difference between actual and predicted pay ranges from actual pay 120% higher than predicted to actual pay 56% lower than predicted.

b. In addition, the spread between the predicted and actual pay for women within the 94.3% of women whose individual pay did not differ significantly from what was expected is quite large – the difference between actual and predicted pay for this group of women ranges from actual pay 29% higher than Dr. Neumark's model predicts to 23% lower.

c. Roughly 1,100 to 1,400 women each year earn more than Dr. Neumark's model predicts. They comprise 43%-45% of all women depending on the year. Relative to the benchmark set by Dr. Neumark's regression model, there are a substantial number of women who do not appear to have been underpaid, according to the factors he decided to include and control for.

d. What Dr. Neumark's regression model cannot answer is whether any particular woman could identify a man doing substantially similar work who is paid more than she is. To better understand this, I randomly selected a woman in each of the ten largest female job titles in the data who have at least two male comparators as defined by Dr. Neumark's model. The "comparators" for purpose of this exercise are employees who are similar in all of his regression model variables: experience within two years, Oracle tenure within two years, job tenure within two years, job code and grade, part time and hourly statuses, zip code and line of business head. I observed a variety of outcomes. In some groups, women are both the highest and lowest earners. In other groups, women earn less or more in general, and yet other groups exhibit no pattern by gender. Women do not appear to be systematically lower earners, even within groups defined by Dr. Neumark's model as comparators.

e. These outcome variations are the rule, not the exception. Taken as a whole, these are the highly variable and inconsistent outcomes I would expect to see in a model that does not compare pay among employees doing substantially similar work.

11. Third, Dr. Neumark's analysis of prior pay is methodologically flawed. The issue of the correct benchmark to use in making comparisons is also evident here. Dr. Neumark does not claim there is any statistical evidence that Oracle specifically relied upon prior pay when setting starting pay. Instead, he states that the gender gap in starting pay "reflects" the gap in prior pay, and that his evidence "is consistent with" the gender gap in annual pay being related to gaps in starting pay. These are not statements about causality, only correlation. (A popular and instructive example of this phenomenon is that higher ice cream sales in the summer are

correlated with higher murder rates in the summer. This does not mean that ice cream causes homicidal behavior.)

a. Dr. Neumark fails to note that his observation of a "strong relationship" between prior and starting pay is fully expected. In the economy at large, the extent of correlation between prior and starting pay is also very high. This is fully expected. Dr. Neumark fails to establish the correct benchmark against which to test the Oracle finding, and simply concludes that the correlation between Oracle starting pay and prior pay is so strong that such a correlation has less than a one in a billion probability of occurring by chance. Such a statement compares to a hypothesis of zero correlation, which is of course absurd given that prior pay and starting pay would be expected to co-vary, even if a new employer had no knowledge at all of applicants' prior pay.

b. My analysis of national data on people changing jobs shows high correlation rates between prior pay and starting pay across the labor market. That the two measures of pay are highly correlated at Oracle is not unique to Oracle or even to the technology sector.

c. The data Dr. Neumark relies on for his prior pay analysis is also seriously flawed. This is because the prior pay variable in the hiring data is not standardized; some people list annual base salary, others list total compensation, and yet others list some mix of the two, and it is not always clear which is which. As Dr. Neumark states on page 26 of his report, "[...] I attempt to use prior base pay whenever base pay is explicitly reported (425 employees). However for most employees (2,358), it is ambiguous whether the salary number given is base pay or total compensation." In spite of this observation, Dr. Neumark does not limit his analysis to the sample of 425. He instead uses the full data sample of 2,783, even though he cannot tell if he is comparing "apples to apples" for 85% of the data.

d. When I follow Dr. Neumark's advice to compare "apples to apples," and I restrict his analysis to the 425 employees and use his data as is, the gap in starting pay remains similar to the gap in his report: women's starting pay is 2.16% less than men's, based on his model. However, the gender gap in prior pay is much larger: women's prior pay is 4.91% less than men's using Dr. Neumark's model. The difference between starting pay and prior pay regressed on the same controls shows that women actually do 2.74% better than men when moving to Oracle, according to Dr. Neumark's model. Thus, Dr. Neumark's suggestion that the two pay differences are essentially the same falls apart, and with it, his argument that Oracle is carrying into the company pay disparities from the outside labor market in some consistent, formulaic way.

e. Furthermore, Dr. Neumark's dataset of 425 contains obvious data errors, including typos and interpretative errors for the measure of prior pay. Dr. Neumark's prior pay model, which covers hires over a six year period, also does not statistically account for year – a serious oversight considering the pace of change year-to-year in the technology labor market and significant changes over time in Oracle's annual performance as a business. Fixing Dr. Neumark's data errors within the sample of 425 and controlling for year leads me to estimate that the gender gap in prior pay is under 1% and is not statistically significant. In short, using the corrected data for the sample of 425 and adding one important variable shows that there is no gap in prior pay that could then explain any aspect of starting pay for female Oracle employees.

f. By his own admission Dr. Neumark cannot rely upon the prior pay variable for 2,358 of the 2,783 hires he analyzes due to data problems, and the sample of 425 he states contain consistent data do not support the conclusions he reaches using the flawed larger sample.

12. Dr. Neumark's starting pay and prior pay analyses are not the only fragile estimates he relies upon. Rather, many of his analyses suffer from serious data construction errors; among these are that he includes employee data from outside the putative proposed class, fails to correctly or completely exclude college hires, and overstates compensation for employees who worked only part of a year.

13. Of particular note, Dr. Neumark's total compensation pay measure is completely uninterpretable because of the inconsistent and mistaken way he valued the stock award part of total compensation. Regarding the analysis of base pay (which is not impacted by that particular error):

a. When I analyze base pay, even if I fully aggregate across all jobs and years like Dr. Neumark did (without agreeing that approach is correct), adding a few readily available variables and fixing a couple of mistakes he makes with the variables he does include, immediately cuts the measured pay difference he reports almost in half. This is without doing anything to further differentiate within job codes, but simply using them just as Dr. Neumark does.

b. If one were able to obtain further detail on the specific nature of the work being performed by employees – as well as additional differentiators that individual managers used in making pay decisions – and incorporate this information, it is my professional judgment that the balance of the measured pay difference would likely disappear. Dr. Neumark did none of this work.

OVERVIEW OF STATISTICAL METHODS

14. The primary method used by Dr. Neumark in his work, and by me in my response to him is multiple regression analysis.[5] Multiple regression is a technique that is used to understand the impact of a number of factors, typically called variables, on some phenomenon of interest. The phenomenon, or variable of interest is referred to as the "dependent variable" and the factors used to explain this variable are referred to as independent variables. In this case, various measures of pay constitute the dependent variable, and the independent variables used to explain pay are things like job tenure, job code, age, and other factors. There are a number of different types of regression methods, but they all share the structure of independent variables being used to explain a dependent variable. If a statistician has correctly measured the dependent variable, and has accurate measures of all the independent variables needed to explain the dependent variable, then the results of a regression analysis are typically straightforward to understand and interpret.

15. The result of running the regression procedure will be a set of "coefficients," which represent the *average* quantitative magnitude of the impact of each independent variable on the dependent variable. For example, having one additional year of work experience may increase pay, on average, by 4.9%. To take another example, holding job A relative to a base job may be associated with 13.5% higher pay. In a multiple regression context, we interpret each coefficient as having the measured effect, on average, *holding constant all other variables in the model.* In a gender pay context, the coefficient on female would represent the average difference between male and female pay, holding all other factors in the model constant. Note that the regression

[5] This discussion is by necessity brief. For a more thorough definition, there are many econometric textbooks that describe the methodology in great detail. See, for example, Greene, W. (1993) *Econometric Analysis*, 2nd Edition, NY: Macmillan Publishing Company.

coefficient represents only the average effect – we have to probe further to understand the variations surrounding that average. A coefficient can be positive, meaning women earn more, all other things held constant, or it can be negative, meaning women earn less. While "more" and "less" are indicated by the magnitude and sign of the coefficient, we also want to know whether the coefficient is what is referred to as "statistically significant." In other words, the effect measured by the regression may be due simply to random fluctuation. We determine if a coefficient is meaningful by conducting a test of statistical significance. This test will tell us if a variable should be considered "meaningfully" related to the dependent variable. The results of these tests are typically noted by the analyst.

16. Of course, in the real world, it is not always easy to measure any variable correctly, whether dependent or independent, and it is not always the case that the analyst knows which independent variables are the ones that should be included in a regression model. For example, if an analyst left measures of relevant work experience out of a pay regression model, this is likely to create problems because the coefficients on correlated variables that *have* been included will be biased because they include some of the correlation properly associated with the omitted variable. In the context of using regression methods to study a gender equal pay claim, there is a particular problem that the statistician must deal with. If there is a variable that does in fact relate to pay and is left out of the model – i.e., it is "omitted" – then the question is how this affects the magnitudes of the *other* coefficients in the model. It turns out that many variables are correlated to each other, such that omitting a variable from a regression, or including it when it was previously omitted will change the value of the coefficients on other variables. For example, if women had more work experience than men on average, and you omitted work experience, the measured effect on pay of being female would be biased upwards by that

omission. This is what is called **"omitted variable bias."** This is a persistent issue in multiple regression analysis in the real world, where it can be difficult to know what factors matter, and difficult to obtain measures for variables you know are important.

17. Now, suppose we want to see how well our regression model predicts pay. We can use the commonly derived set of average impact regression coefficients together with each employee's individual values for the model's variables. Because the regression coefficients are common, they represent the average impact across all employees, and thus there is one set of regression coefficients that is applied to each employee in the data to compute each prediction. We compute the pay the model predicts for each employee, and compare that predicted pay to their actual pay. If the model is well specified, meaning we have captured most or all important factors that impact pay and we have measured them correctly, the model should more or less predict what an employee actually earned. If we have left out variables, or measured them poorly, we will not get a good set of predictions, and there could be wide discrepancies between the actual and predicted pay. This procedure is called "in-sample prediction," or analysis of residuals, because we are using the model computed on a sample to predict values within that same sample. This is a common way to assess the quality of a regression model.

18. Another problem in applying regression analysis in real world situations is that the measures of both dependent and independent variables are not always accurate. Below I discuss at some length problems with Dr. Neumark's use of Oracle's job codes. In this case Dr. Neumark attempts to use regression methods to compare men and women who he claims are performing substantially similar work, and to then test to see if women are paid differently than men. If the variables that are critical to this are overly broad, and include many types and levels of work within the same job code, use of these job codes with no further refinement can lead to

misleading and biased conclusions regarding female pay. Dr. Neumark also analyzes mismeasured dependent variables, for total compensation and for prior pay. The consequence is that neither analysis can be viewed as reliable.

VARIABILITY

19. In this section, I use the dataset created by Dr. Neumark for his analysis to show that his "one size fits all" regression models gloss over significant variation in job responsibilities, scope, and skill demands among employees at Oracle. I will elsewhere document the shortcomings in this data, but for now, I will set aside those issues and use exactly the data he relies upon in reaching his conclusions. Dr. Neumark has proposed a model that he represents or that Plaintiffs assert is adequate to demonstrate a statistically common circumstance among the members of the putative class, and it is that hypothesis that I examine. My conclusion is that the highly variable and inconsistent outcomes I observe are not consistent with a model that compares pay among employees doing substantially similar work.

<u>The proposed class is quite diverse</u>

20. As noted above, according to Dr. Neumark's data, the proposed class contains 4,132 women employed at Oracle in three job functions (Product Development, Information Technology, and Support) from June 2013 through the present. There are 265 job code/job grade combinations, for which annual base salaries for full time, full year employees range from [redacted]. The range in total annual compensation for these full time, full year employees is [redacted]. The proposed class members in Dr. Neumark's data include everyone from hourly employees working as Business Services Representatives to

Product Development Executive Vice Presidents overseeing organizations[6] employing thousands of professionals. The proposed 4,132 class members directly reported to 3,409 managers; of the 2,190 managers for whom the data contains gender information, 514 are women.[7]

21. Oracle is a large company that develops and markets a wide variety of business software and data management technologies.[8] My understanding is that their scope includes both maintaining legacy software solutions and creating new software solutions and software delivery systems in response to market demand. My further understanding is that there are employees who build the backbone of such systems, as well as those who build applications on top of that backbone, and yet other employees who design the customer interface and work with clients to implement the software and iron out problems.[9] Some employees arrive at Oracle straight from school, others previously worked at one of 72 companies acquired by Oracle as reflected in the data, and still others arrived at Oracle after starting their careers elsewhere. The latter category worked at a variety of companies, such as IBM, Apple, Bank of America, Gap Inc., and Herbalife, or were self-employed. The college majors of putative class members in the data include not just engineering and computer science but also art, biology and philosophy.

22. In the Product Development job function, the data contains employees in Individual Contributor career level 0 ("IC0"), earning [redacted] on average in total compensation (all salary,

[6] It is my understanding from the Steven Miranda declaration and from the correspondence between Mantoan and Finberg regarding the data that products and services are correlated at least roughly with a variable in the data called organization. See Declaration of Steven Miranda in Support of Defendant Oracle America, Inc.'s Motions for Summary Judgment or, in the Alternative, Summary Adjudication, January 17, 2019, paragraphs 3 and 8; August 17, 2018 letter to James Finberg, [Oracle] Mantoan ltr to [Jewett] Finberg in resp to data Qs 21, 22, 26.pdf, page 3.

[7] This is based on supervisor ID in the assignment data. Gender is not available for all supervisors, as they do not always fall into the population definitions used for data production where gender is recorded (for example, because they worked outside California). There are 1,219 direct managers for whom gender information is not available.

[8] Oracle Fact Sheet (http://www.oracle.com/us/corporate/oracle-fact-sheet-079219.pdf).

[9] See Miranda Declaration, paragraph 2.

no bonus or stock) up through a handful of employees in career level Manager 8 ("M8") with average total compensation of [redacted]. In the Support job function, total compensation ranges from an average of [redacted] to an average of [redacted]. In the Information Technology job function, total compensation ranges from averages of [redacted] (comprised only of base salary) to [redacted]. Indeed, the composition of total compensation is markedly different by career level, with [redacted].

In the Individual Contributor Career Levels, Total Compensation is [redacted]

- Dr. Neumark's Data, 2013-2017, Full Time Full Year Employees, Individual Contributors By Career Level -

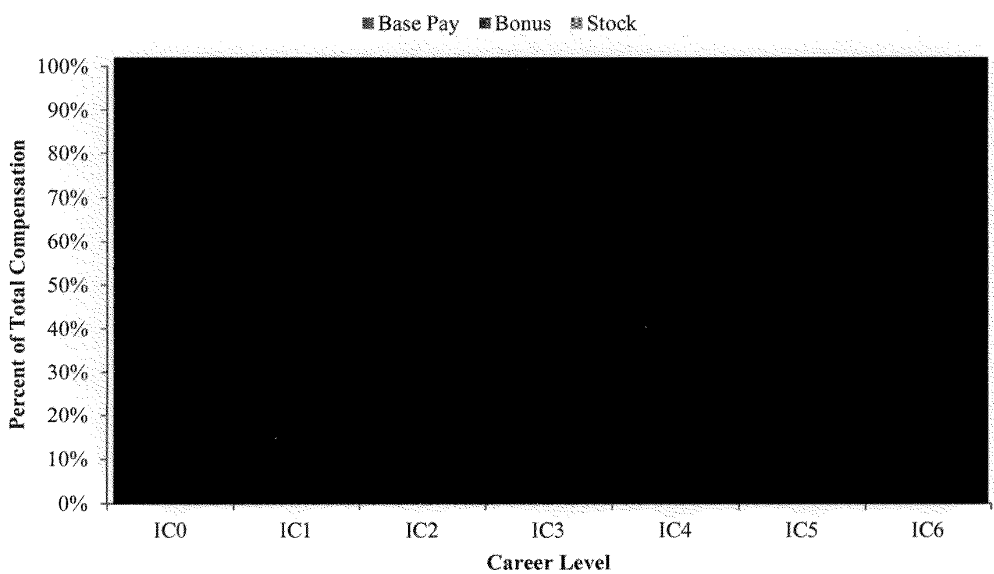


Exhibit 1

At Higher Manager Career Levels, [redacted]

- Dr. Neumark's Population, 2013-2017, Full Time Full Year Employees, Managers By Career Level -

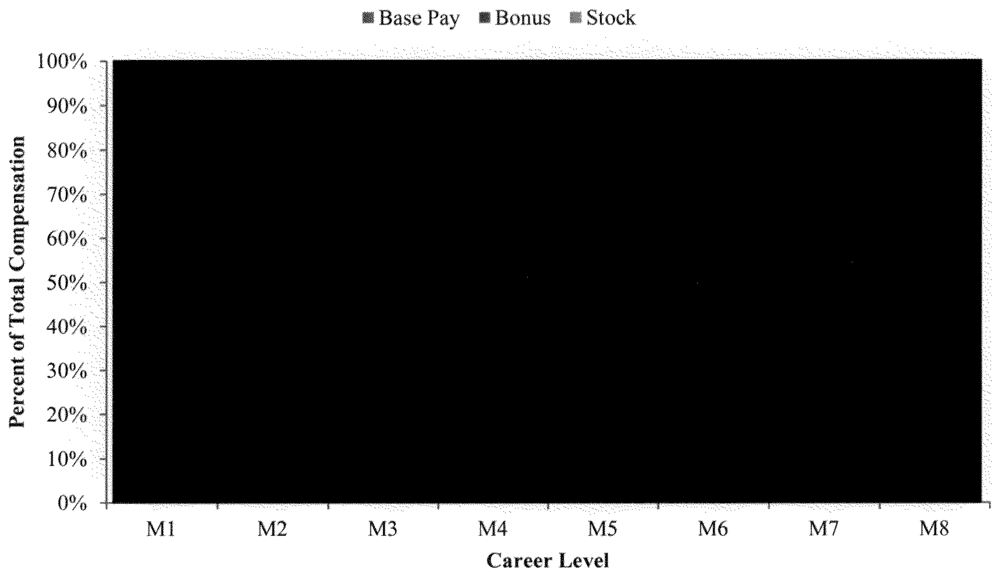


Exhibit 2

23. It is my understanding based on materials produced in the case that the kinds of jobs at issue in this matter draw on a wide range of skills and experience, not just various types of computer programming. For example, job postings for some Layout/Mask Designers indicate that they create and model integrated circuits related to work on Sun Systems Hardware, a computer hardware company which was acquired by Oracle in 2010.[10] Job postings for some

[10]The requisition data includes a posting for an IC2-level Layout/Mask Designer: 100520.Layout/Mask Designer 2.PRODEV.ENGSVCS.IC2, Vacancy name IRC1353472.: "Develops and prepares multi-dimensional layouts and detailed drawings of the semiconductor devices from schematics and related geometry provided by Design Engineers. Work may be completed through the use of CAD equipment. Checks dimensions, writes specifications and verifies completed drawings or digitized plots. Performs and may plan mask design work." "Preferred skills include: proficient in SRAM related layout knowledge and skills related to

Technical Analysts indicate that they assist customers with their technical issues.[11] Technical Writers may have majored in English in college, and work on manuals and publications.[12] They may all have college degrees, but their specialties are quite different.

Job code and grade mask differences in the work employees do, contrary to Dr. Neumark's assumption that job code and grade "very narrowly" defines work

24. As noted, Software Developers, who are spread among career levels 1 to 5, comprise the largest job category in the data. From a labor economics perspective, my examination of the requisitions that describe openings leads me to believe that there are differences in the focus and demands of their work. There are Software Developers tasked with designing and guiding new

---

28nm process technology and/or beyond; excellent understanding of layout dependent parameters (LDP), DFM and ERC; self-motivated, team player with good interpersonal, leadership and communications skills; familiar with Sun environment, P; R and SKILL scripting are great pluses; intimate working knowledge of followings: Cadence: Virtuoso, Virtuoso-XL; Mentor: Calibre DRC/LVS; Unix environment"

[11] 90120.Technical Analyst 1-Support.SUPP.PRODSUPP.IC1. Vacancy Name IRC1895672. "As a member of the Support organization, your focus is to deliver post-sales support and solutions to the Oracle customer base while serving as an advocate for customer needs. This involves resolving post-sales non-technical customer inquiries via phone and electronic means, as well as, technical questions regarding the use of and troubleshooting for our Electronic Support Services. A primary point of contact for customers, you are responsible for facilitating customer relationships with Support and providing advice and assistance to internal Oracle employees on diverse customer situations and escalated issues." "Work involves some problem solving with assistance and guidance in understanding and applying company policies and procedures. As this is an entry-level technical position at Oracle, a technical degree is preferred i.e., BS Computer Science/Management Information Systems/Science/ Engineering/Math/Physics/Chemistry with a 3.0 GPA. OR functional degree + technical higher degree or in lieu of degree may substitute 4 years professional experience & professional certification (i.e., CNE, MCSE, CPA, Oracle, etc.)."

[12] 13520.Technical Writer 2-ProdDev.PRODEV.TECHWR.IC2. Vacancy Name IRC2007631. Technical Writers document the products; an IC1-level job requisition describes it as: "Creates, develops, plans, writes and edits operational, instructional, maintenance, test or user manuals for paper, multimedia or web-based publications. Contributes to the timely design, production and delivery/completion of product documentation and document sets." The position requires certain skills: "Knowledge of information development and publication tools, such as FrameMaker and VisioExcellent; research skills; Ability to understand and document complex concepts; Ability to write clearly and concisely using correct grammar, spelling, and punctuation about highly [...] Undergraduate degree in English, journalism, or computer science."

products, and others charged with maintaining existing systems and products, and those who test existing products. For example, this Software Developer 1 is working on a new product within Oracle and will need to bring artificial intelligence and machine learning skills to bear. The starting base pay for the person hired through this requisition was [redacted]

> "Software Developer IC1: Oracle Intelligent BOTS is a **newly formed group** within Oracle working on solving some really hard problems like a platform for computer programs that leverages machine learning, artificial intelligence to enable natural conversations with people. We are **like a start-up inside a large company with a big charter and lot of creative freedom**. We have assembled some of the smartest people in the industry and are growing this team. If you are looking for an exciting new opportunity, **build new products from the ground up with state-of-the-art technology**, high level creativity, then we are the right fit for you. [...] You will get the opportunity to **apply your knowledge in AI, Machine Learning, NLP** etc. to our Chatbot Platform. You will also work on **designing and creating** a new highly available, scalable, and performant architecture [...]."[13]

25. Other job descriptions for entry level Software Developers call for no particular programming skills. The starting pay for the individual hired through the following Software Developer 1 requisition was [redacted][14] and the description emphasized testing existing products:

> "Software Developer IC1: "As a QA Analyst in Configuration Engineering, you will be part of a team that is a collaborative, global organization consisting of developers and analysts with a deep understanding of Oracle Hardware Products and Oracle's E-Business Suite. You will **develop and execute test plans to ensure quality of configurator software releases**. You will also be responsible for test planning, test plan generation, hands-on **testing and problem reporting. No prior experience with Oracle Configurator or other Oracle products is required. Due to the specialized nature of the work, training on essential tools and processes is provided. Some basic programming experience is helpful.**"[15]

26. This example of two Software Developer 1s, and many other such examples in the data I reviewed, undermines the idea that these individuals are performing substantially similar work; instead, they likely differ not just in terms of the content of their work at Oracle but in their outside opportunities in the market. The "price" that their skills carry may differ depending on

[13] Vacancy IRC3537583. Emphasis added.
[14] In 2017 dollars, using Dr. Neumark's CPI conversion.
[15] Vacancy IRC2485625. Emphasis added.

the other companies in the industry that are competing in those product areas and thus also competing for workers. For example, a legacy software product might have smaller future profit opportunities from a discounted cash flow perspective, and therefore employees working on these products might earn less relative to employees working on new products for which the company has or anticipates high current and/or future profits.[16] And the "price" of these different skills would be expected to vary over time, as both the demand for and supply of those particular skills fluctuate in the market because of changes to interest in a particular technology, the number of universities or other training programs teaching individuals those skills, the number of companies participating in (and thus trying to hire in) the market for those technologies, and so on.

27. The control variables in Dr. Neumark's regression model do not ensure appropriate comparisons of employees performing substantially similar work from a labor economics perspective. For example, in Dr. Neumark's data, there are 2,517 unique employees who held the title of Software Developer 4. The question is whether all 2,517 employees are indeed doing substantially similar work or whether Dr. Neumark has failed to incorporate control variables that would account for meaningful differences in their skills or responsibilities. In his data, annual base salaries for full time full year Software Developer 4s range from [redacted]

[16] In labor economics, wages in the short run are influenced by wages in the market, the demand for the company's product and the structure of the market they compete in. (See for example, Cahuc and Zylberberg, *Labor Economics*, Cambridge: The MIT Press, 2001, p. 175, though the principle is much older than that.) Dr. Neumark concisely summarized the relationship between productivity and pay in prior testimony, noting that "pay is based on a worker's productivity. In fact, their marginal productivity [...] the worker adds this much revenue to the firm, the firm isn't going to pay you more than that because then they'd lose money on you. And they're probably, in equilibrium, not going to pay you less than that because somebody else would pay you that much." (Deposition Testimony of Expert David Neumark, *Rabin and Chapman et al. v. Pricewaterhousecoopers, LLP*, United States District Court Northern District of California San Francisco Division, Case No. 16-cv-02276-JST, January 12, 2018, 100:16-101:1.)

Narrowing to employees within Mr. Kurian's line of business, for example, and to employees with 8 to 10 years of Oracle tenure, the differences are [redacted], and [redacted] for base and total compensation respectively. Likewise, among employees in Mr. Kurian's line of business with 4 to 6 years of tenure in their jobs, the differences are [redacted], and [redacted] for base and total compensation respectively. A pay range that wide suggests that Software Developer 4s do in fact engage in very different kinds of work. From a labor economics perspective, it would be very surprising if employees with substantially the same set of skills, duties, and responsibilities as others were willing to accept [redacted] the pay of others performing substantially the same work.

28. Job requisitions for specific Software Developer 4 positions in the Taleo Application system reflect those differences. For example, one such requisition describes the position this way:

> "As a <u>Software Test Automation Engineer</u>, one is responsible for developing backend automation tests cases (from scratch). Will perform a wide variety of testing from performance, functional, load and reliability testing. Engineers will report, analyze, troubleshoot bugs and work with development team for resolution. Must be highly passionate about tearing software apart and finding defects bugs."

The posting further specified: "We are seeking seasoned engineers with <u>a minimum of 3+ yrs software development automation and testing experience</u>."[17] Starting pay for the individual hired into this position was [redacted].[18]

29. In contrast, a posting for a different Software Developer 4 position for whom the successful candidate earned a [redacted][19]) indicated that:

> "As a <u>Sr. Principal Software Architect Engineer</u> you will own and lead software architecture and development for major components of Oracle's cloud infrastructure. You should be a distributed systems generalist, able to architect broad systems interactions,

[17] See Taleo requisition 140009JE.
[18] In 2017 dollars.
[19] In 2017 dollars.

while being very hands-on, able to dive deep into any part of the stack and lower level system interactions." This position also called for "8+ years' experience delivering and operating large scale, highly available distributed systems." [20]

30. Again, both positions were for a Software Developer 4 in the line of business headed by Executive Vice President Thomas Kurian.[21] Dr. Neumark controls for line of business head in his models, which he describes as the "reporting chain of command,"[22] only by controlling for the highest-level manager in that line (in this case, Mr. Kurian) – but Dr. Neumark does not claim to have studied these lines of business, or whether there are further differences in the type of work performed within sub-divisions or sub-units of these lines of business.[23] Dr. Neumark further testified that he did not study job postings, narrative information from performance reviews, promotion justifications, or hiring justifications, to support his assumption that the variables in his model effectively grouped employees doing substantially similar work.[24] He apparently ignored this readily available information. In my professional opinion, one should at least evaluate this additional detailed information in order to test the assumption that job code/grade "narrowly" defines work.

The representativeness or applicability of an average depends on the extent of the variation around it

31. By estimating only a single, aggregate regression model over all employees, all years and all jobs, Dr. Neumark has assumed that women at Oracle are a cohesive group best summarized

[20] See Taleo requisition 140014C9.
[21] My understanding is that Thomas Kurian is no longer employed by Oracle (and thus is no longer a line of business head), but that he was through the end date of the data produced in the case. See https://www.linkedin.com/in/thomas-kurian-469b6219/.
[22] Neumark Deposition, 116:7-8.
[23] "Q. ·Does the line of business structure at· Oracle relate in any way to the products or services· on which an employee is working? ··A. ·You've already asked me that.· I -- I -- I· don't know, sitting here, what the relationship is.· It wouldn't surprise me if there was some· relationship." Neumark Deposition, 123:12-18.
[24] Neumark Deposition, 62:19-64:24.

by a single common model that averages over all their experiences.[25] As noted above, regression coefficients describe aggregate average differences in pay. When there is not much underlying variation in the data and associated analytical outcomes, then an average is potentially a useful summary measure. If there is a lot of variation - both in the ranges of pay levels across or within job categories, or in the individual statistical outcomes - an average can be misleading.

32. To illustrate this point, in a descriptive statistics context, imagine a group of 100 employees who earn $100,000 on average. If all of these employees earn between $90,000 and $110,000, then the average of $100,000 is a good summary measure of income in that group and the difference, or "error" between that average and any individual's actual pay when using the average to estimate the pay of a particular person will be relatively small. Now imagine instead that half of the 100 employees earn $50,000 a year and half earn $150,000. Average earnings are still $100,000 but as a summary measure, the average does not capture "typical" earnings very well and the single statistic obscures the fact that there are two distinct and very different earnings groups in the population. At issue in a setting with thousands of employees is whether the average captures something meaningful for all the members of the group, such that extrapolating the mean experience to everyone is reasonably accurate. In the first example,

[25] It is my understanding that the California Equal Pay Act (Labor Code section 1197.5) was amended effective 1/1/2016. My understanding is that, prior to 1/1/2016, the act applied to pay differences between employees performing "equal work on jobs the performance of which requires equal skill, effort, and responsibility, and which are performed under similar working conditions" in the "same establishment." After that date, my understanding is that the language changed to apply to pay differences between employees performing "substantially similar work, when viewed as a composite of skill, effort, and responsibility, and performed under similar working conditions" without the establishment-based limitation. From a technical statistical perspective, these are two different analytical standards for a statistical analysis. To a labor economist "equal work" would suggest that the analysis must focus more narrowly. Dr. Neumark has not addressed this issue in his report, and because I am responding to him, I have not addressed statistically how one could or should break the analytical approach in this case into two parts. But because I conclude that his model does not have sufficient controls to compare employees performing similar work, if some portion of the class period is subject to a stricter comparator standard, his model necessarily would fail there as well.

where everyone earned close to $100,000, extrapolating the average to everyone would be reasonably accurate. Extrapolating from the average of $100,000 to all members of a group where some earn $50,000 and some earn $150,000 would lead to highly inaccurate estimates for all individuals in that group.

33. In the regression context, the single regression coefficient on female represents the average difference in pay between women and men, taking into account all of the characteristics in the model. This average difference may not characterize the experience of a substantial number of women in the analysis. Dr. Neumark's model is estimated by averaging gender differences across all employees, even though gender differences in pay among hourly employees may be quite different from gender pay differences among Executive Vice Presidents who have considerably more advanced skills and more complex responsibilities, and whose pay combines salary, bonuses and stock awards.

34. His aggregate model also only averages the effect of other variables' impact on pay. For example, a regression coefficient on years of Oracle tenure will represent the average impact of Oracle tenure over all employees and all jobs. However, the underlying relationships of the employees to the impact of tenure might vary widely; for example, tenure may be more relevant for someone working on a legacy product than for someone working on a team trying to design a new and innovative product. Consequently, the computed average effect of a single regression coefficient does not speak to the extent of variation in the underlying data used to estimate that average.

A single coefficient on gender in a pay regression is only an average and does not in and of itself answer the question of whether, from a statistical perspective, the circumstances of pay outcomes of female employees at Oracle are amenable to common analytical treatment across all class members

35. Multiple regression statistical methods are used to study whether there is on average a relationship between pay and gender once the factors that the model uses to reflect other characteristics that determine or influence compensation are taken into account. The estimated coefficient on gender can be characterized by its sign, its magnitude, and its statistical significance. If women are paid less on average than men holding similar jobs, defined as work involving similar skills, effort and responsibility, and performed under similar working conditions, then the regression coefficient on gender will be negative: In other words, women would be observed being paid less *on average* than expected based on their work-related characteristics. If the coefficient on gender is positive, it indicates that women are paid more on average than men doing similar work. The size of the coefficient speaks to the *practical significance* – in other words, the substantive impact or real-world implication – of the relationship between pay and gender.[26] Finally, if the regression coefficient is not statistically significant, then sex is not statistically related in a meaningful manner to pay on average, holding

[26] "Practical significance means that the magnitude of the effect being studied is not *de minimis*—i.e., it is sufficiently important substantively for the court to take notice. For example, if the average wage rate is $10.00 per hour, a wage differential between men and women of $0.10 per hour is likely to be deemed practically insignificant because the differential represents only 1% ($0.10/$10.00) of the average wage rate." Rubenfeld, Daniel, "Reference Guide on Multiple Regression," *Reference Manual on Scientific Evidence: Third Edition* (p. 318), Washington D.C.: The National Academies Press. There is also a sizeable statistics and econometrics literature on the issue of "practical significance." See McCloskey, Donald N. (1985). The Loss Function Has Been Mislaid: The Rhetoric of Significance Tests. *American Economic Review Papers and Proceedings*, 75(2), pp. 201-205. Also see Leamer, Edward, *Specification Searches: Ad Hoc Inferences with Non-Experimental Data*, New York, Wiley, 1978. Also see Piette, Michael J. and Paul F. White (1999). Approaches for Dealing with Small Sample Sizes in Employment Discrimination Litigation. *Journal of Forensic Economics,* 12(1), pp. 43-56.

work-related characteristics constant.[27] A coefficient can be statistically significant without being large in any meaningful, practical sense; similarly, a coefficient can be large but not statistically significant, particularly if the population on which the estimate is based is small.

36. All that said, a regression coefficient is simply an average. Some women will have been paid less than the model predicts based on their non-gender characteristics such as education, experience, and job title. Some women will be paid about what the model predicts, and some will earn more than the model predicts based on their individual characteristics. An average always *can* be estimated; that in and of itself does not mean it is necessarily the best summary statistic to describe the data. A sink with separate hot and cold taps will produce warm water on average, but neither tap is accurately described as warm.

<u>The extensive variation in outcomes among employees who Dr. Neumark's model considers observationally similar suggests that the model is misspecified and does not compare employees doing substantially similar work</u>

37. It is my understanding that the legal issue currently before the Court is to consider whether class certification is appropriate, which from a statistical perspective would ask whether there are patterns in pay outcomes that are "common" among the individual members of the putative class. In this section, I use Dr. Neumark's data and variables to examine pay outcomes and thereby gauge how sensitive his conclusions are to being aggregated into a single number with a single conclusion regarding the pay of women at Oracle. One way to examine this variability is to study employees' actual earnings relative to what Dr. Neumark's model predicts for each person. The statistical software itself essentially automatically predicts pay for

---

[27] The t-statistic typically used to calculate the statistical significance of a coefficient is based on part on the underlying variability of the data but speaks to whether the coefficient's size could be explained by chance (assuming the correct model has been estimated). It does not address whether the data are appropriately analyzed in one big group or if it would be more sensible to model subsets of the data separately.

everyone in the data as part of its calculations that generate the regression results. It is a simple matter to modify Dr. Neumark's computer code to retain and view each employee's predicted pay.[28]

38. In developing these predictions, the other adjustment to Dr. Neumark's computer code is to remove the gender variable from the model that predicts pay. The idea here is to predict pay based only on job and employee characteristics *other than gender*. What would an employee earn regardless of gender based on their characteristics? Thus, I re-estimated Dr. Neumark's regression model, dropping gender as a control variable, and then examined each person's actual and predicted pay.

39. The graph below plots actual base pay for each female employee on the vertical axis and what their predicted base pay would be based on the non-gender variables in Dr. Neumark's analysis on the horizontal axis. Because Dr. Neumark aggregated everyone from every year into a single model, each dot in the graph indicates a person-year. The dashed line indicates where actual pay equals predicted pay. Dots above the dashed line indicate employees who are paid above what his model predicts; dots below the line indicate employees who are paid less than his model predicts. By design, because regression models estimate the average effect, roughly half of all the points should be scattered randomly above the line and half below. However, if women are systematically underpaid, they should "leap out" visually on the graph by being predominately below the dashed line. The graph has only female data points.

---

[28] Dr. Neumark uses the inverse hyperbolic sine transformation for base pay and total compensation because in his analysis of [redacted] Also, I run a linear least squares model instead of the "reghdfe" method he uses to absorb multiple fixed effects. These technicalities make virtually no difference to the results.

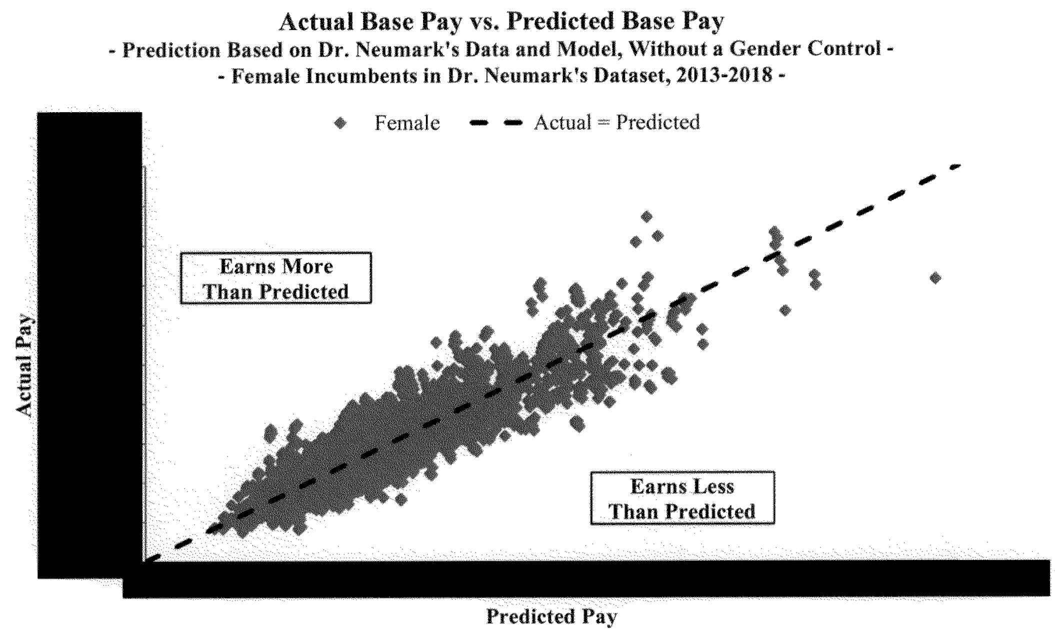


Exhibit 3

40. Dr. Neumark's regression model cannot explain wide pay differences in employees it considers similar. For example, take the point along the horizontal axis at [redacted], which is where predicted pay equals [redacted]. If one were to draw a straight line vertically from that point upwards, which intersected with a dot for an employee below the dashed line, that would indicate someone whose actual pay was below the predicted amount of [redacted]. If one were to continue that same line up from [redacted] and intersect it with an employee dot above the dashed line, that is someone whose actual pay was higher than the predicted [redacted]. Both of those dots represent women who based on their observable characteristics, were predicted by Dr. Neumark's model to be paid [redacted], but one woman is paid more than the expected [redacted] and the other woman is paid less. The employees in the data – both men and women – who are predicted to earn about [redacted] using Dr. Neumark's model actually earned between [redacted]

and [redacted][29] This wide variation in actual pay between employees that the model considers similar is unexplained by the regression model, because the model makes the same average prediction for all of them.

41. The next graph displays the same information but portrays it somewhat differently. As before, an employee whose actual pay is greater than her predicted pay is plotted above the horizontal axis and an employee whose actual pay is less than her predicted pay is plotted below the axis. The height of the bar measures for each female employee, the percentage by which actual pay differs from predicted pay.[30] Employee outcomes are sorted from highest to lowest. If most or all women were adversely affected by Oracle's pay policies and practices, they would largely appear below the horizontal zero axis – i.e., their percentages would be negative when comparing actual to predicted "should have been paid" pay. The graph shows instead that 43.6% of women are not systematically adversely situated relative to men, using Dr. Neumark's model; the point at which the bars flip from positive to negative is near the middle of the graph, not over toward the left.[31] That the height of the bars ranges from roughly positive 120% to negative 56% shows that a one size fits all regression model is likely inappropriate, and that a single regression coefficient is only a summary measure that masks a great deal of variation in what Dr. Neumark claims are tightly circumscribed regression-controlled outcomes in the underlying data for women.[32]

---

[29] No one was predicted to earn exactly [redacted]; for the purposes of discussion, I looked at employees predicted to earn between [redacted]

[30] This is calculated as (exp(residual)-1)*100.

[31] Technically, each point represents a woman-year observation because Dr. Neumark aggregates all years together in his model.

[32] Economists (including Dr. Neumark) have argued for using benchmark regression models combining both genders because they requires less restrictive assumptions about employers' decisions about marginal product and pay. See Neumark, David (1998). Employers' Discriminatory Behavior and the Estimation of Wage Discrimination. *Journal of Human Resources* 279-295. Cotton, Jeremiah (1988). On the Decomposition of Wage Differentials. *The*

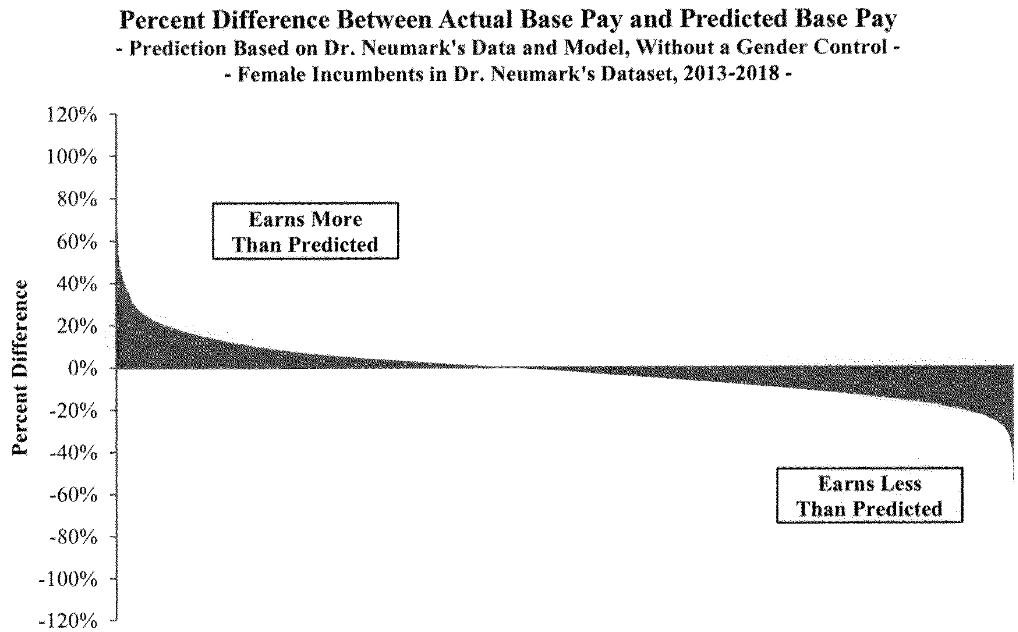
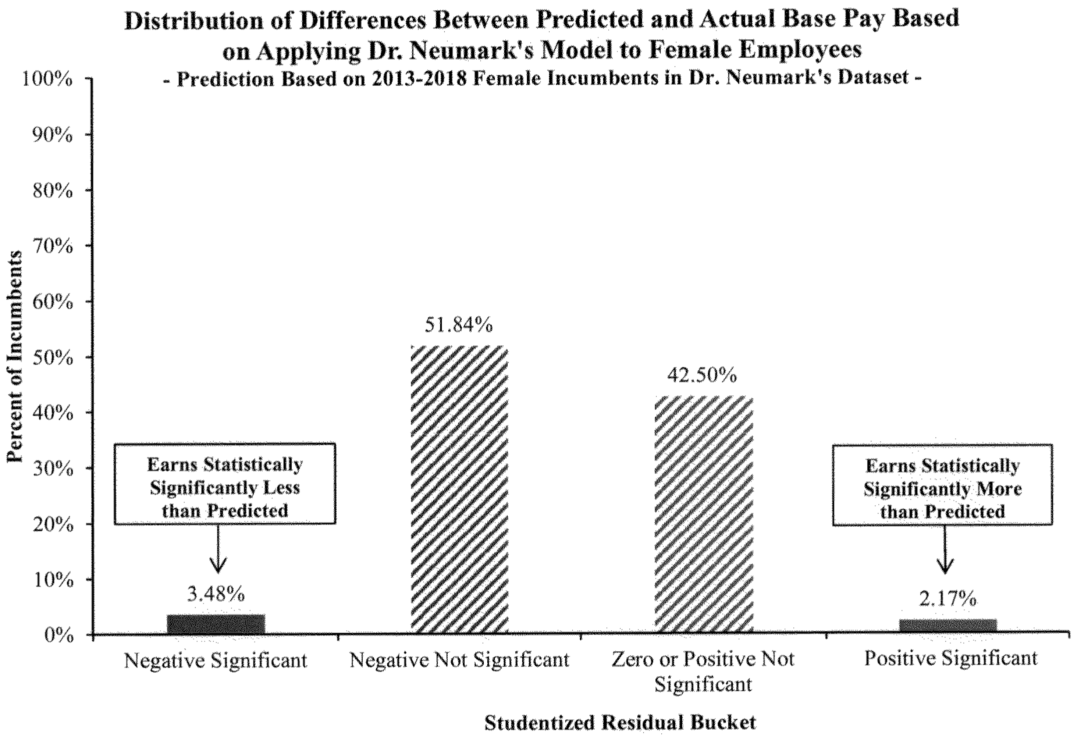


Exhibit 4

42. Finally, whether each woman is paid more or less than expected according to Dr. Neumark's own data and model can be tested statistically, to determine whether they are paid statistically significantly more than expected, about what was expected, or statistically significantly less than expected based on his model. The graph below shows that over all years, just 3.5% of women earn statistically significantly less than predicted and that 2.2% of women earn statistically significantly more than predicted. Another 51.8% of women earn somewhat less than predicted but their actual pay was not statistically significantly less than predicted, and 42.5% whose actual pay was above their predicted pay but not to a statistically significant degree. And even within the range that is not statistically significant, actual pay can still be as much as 29% higher than predicted or 23% lower than predicted.

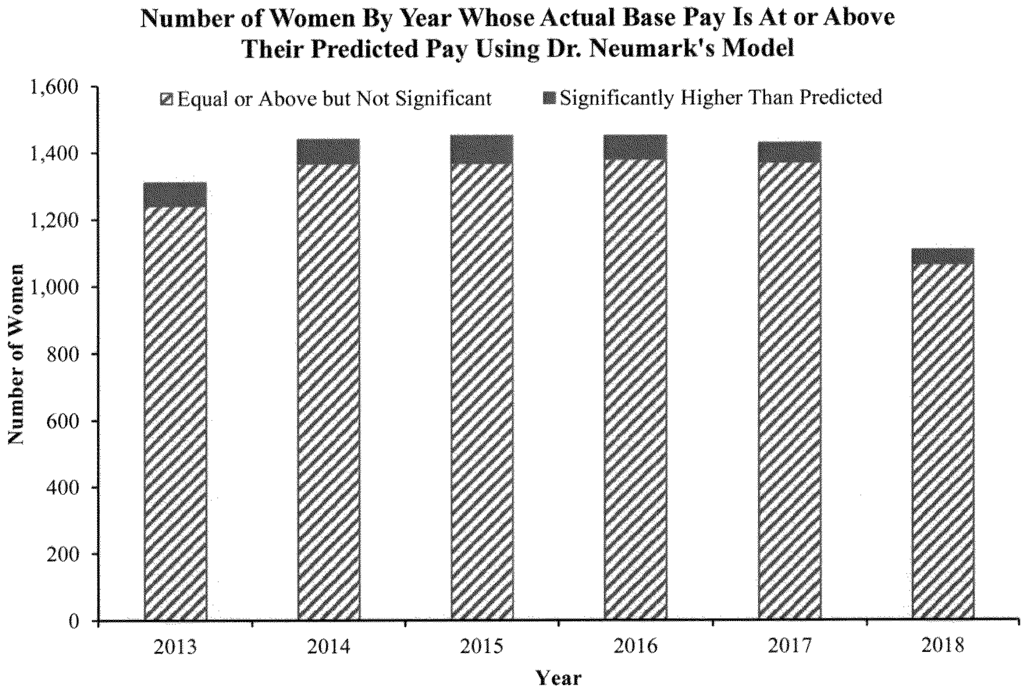
Review of Economics and Statistics, 236-243.



Note: Dr. Neumark's regression model was estimated by year without gender controls.

Exhibit 5

43. Looking at the data year by year shows that over 1,100 women each year earn more than the predicted amount based on Dr. Neumark's data and model. In 2013, the 1,309 women were 43% of all women in the data. Every year thereafter, 45% of women in the data earned at or above the predicted amount. Those numbers and percentages would be expected to rise if the flaws in his model were corrected, provided that variables related to the work being performed are distributed differently by gender.



Note: Dr. Neumark's regression model was estimated by year without gender controls.

Exhibit 6

It is not clear what conclusions can be drawn from a regression model for individual women

44. The statistical issue in this case is whether women doing similar work to men are paid less. A regression model answers this question by predicting pay based on an individual's job and personal characteristics (as reflected in the variables the analyst chooses) and then comparing that prediction to actual pay. In that sense, everyone predicted to earn say, $160,000, is doing equally "valuable" work when considered as the combination of factors included in the model. This approach hinges on having the correct control variables, because otherwise it is not comparing "apples to apples." What the variability charts show is that relative to the average benchmark set by the regression model, women can be paid well above what Dr. Neumark's

model predicts or well below that amount, and his report does not address *why* their pay diverges so much from that benchmark if they are supposedly doing similar work.

45. Is the average (i.e., predicted pay) the correct benchmark? The chart below takes a narrow slice of data – employees predicted to earn between [redacted]. As a group, the commonly applied model predicts the company should value and pay these employees equally. It is hard to know what conclusion to draw: can everyone earning less than the highest earner in this group claim that they are underpaid, even if they themselves earn more than the model predicts? Does it matter that the highest earner (a woman) is a Product Management/Strategy VP, and the other employees in the chart include not only Product Management/Strategy VPs but also IT Senior Directors, a Hardware Development Senior Director, and Software Developer-Architects?

Actual Base Pay vs. Predicted Base Pay:
Predicted Pay Range [redacted]
- Prediction Based on Dr. Neumark's Data and Model, Without a Gender Control -
- Incumbents in Dr. Neumark's Dataset, 2013-2018 -

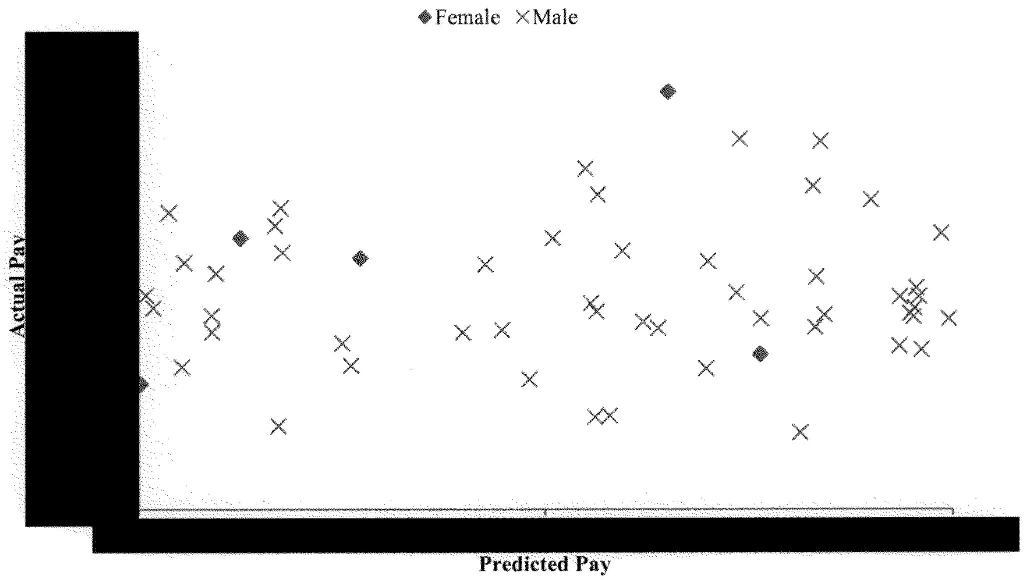


Exhibit 7

46. The proper benchmark question is not one that can be addressed simply by adding whatever variables are at hand to his aggregate model, though certainly the model could do a better job of capturing substantive differences in the skill requirements and responsibilities of various jobs as described in the job postings, performance reviews and hiring manager narratives. Similarly, disaggregating the model into various separate subgroups is still only estimating the average effect and gauging women's outcomes relative to a predicted average. Instead of one overall average for women, as Dr. Neumark has provided, there would be 3 averages or 5 averages or so on. Attachment C contains additional graphs similar to those in Exhibit 3 above, restricted to various subgroups in the population. What they show is that

explaining women's pay outcomes at Oracle is not as simple as dividing the data into some number of subgroups and re-estimating the same model. Fundamentally, his regression model is not an effective approach for answering the question of whether women at Oracle are paid less than men doing substantially similar work. This is to be expected if his model does not adequately control for the skills, effort and responsibility related to the work being done.

47. More detailed information about what these jobs require in terms of skill, effort and responsibility is contained in the thousands of requisitions, hiring manager comments, and resumes produced in this case. To understand whether two jobs are similar, one needs to understand what it means to compare jobs in which, as the general portion of the postings themselves state, "[w]ork is non-routine and very complex, involving the application of advanced technical/business skills in area of specialization."[33] Dr. Neumark has not established that a Software Developer 3 job that requires a "BS or MS degree or equivalent experience relevant to functional area. 4 years of software engineering or related experience"[34] is comparable to a Software Developer 3 job asking for "BS/MS/PhD in computer science or other relevant technical degree; 5+ years of experience in user interface development for web applications; Deep knowledge of HTML, JavaScript, CSS (SASS a plus), DHTML, DOM, Ajax and Java; [...] Expertise using Spring MVC and other frameworks; Familiarity with JavaScript frameworks such as Ext JS and jQuery; Expertise troubleshooting cross-browser and cross-platform issues; Familiarity with XHTML, XML and XSLT; Familiarity with Agile Scrum or similar methodology a plus; Familiarity with SASS a plus."[35]

[33] See Vacancy ID 2456850 for Software Developer 3 in job requisition data.
[34] See Vacancy ID 2456850 for Software Developer 3 in job requisition data.
[35] See Vacancy ID 2491842 for Software Developer 3 in job requisition data.

48. The hiring manager comments are another possible source for understanding the extent to which specific skills are called for, as opposed to general programming knowledge or broad educational background. For example, one candidate for one QA Analyst 3 position was selected over two other candidates because the person "has very good Java, Database, j2ee and plsql and strong Web development skills [...] proficient in unix shell scripting [...]. working in lead role and has good exposure to SDLC [...]" while another candidate "[h]as very minimal JAVA/J2EE experience [...] expertise does not suit our requirement," and the third "[h]as only manual testing experience and lacks product development lifecycle knowledge [...] does not have any development experience."[36] This hiring manager sought particular kinds of experience; it is likely that a different hiring manager hiring someone into the same job title but to work on a different team might prioritize different skills. These differences observable in the data illustrate what Steven Miranda, Executive Vice President of Oracle Applications Product Development, stated: **"Oracle's wide array of products and services translates to a similarly diverse set of skills, duties, and responsibilities among Oracle employees depending upon the product or service (or the component of a product or service) on which an employee works.** Stated another way, just as the technologies themselves differ, so do the skills, duties and responsibilities needed to develop, enhance, modify, support or service those products and services. This can be true whether or not employees share the same job title."[37] Dr. Neumark's model glosses over these differences without ever testing their importance or relevance.

[36] These comments come from OFR APPROVAL COMMENT HISTORY in ORACLE_JEWETT_00007304_native.xlsx for vacancy ID 1719823.
[37] Miranda Declaration, paragraph 3. [Emphasis added.]

There is evidence that employees sharing a job code do a wide variety of types of work, paying widely differing amounts

49. That a single job title may encompass a wide variety of skills and responsibilities is further evident from a statistical perspective when Dr. Neumark's regression model results are restricted to a single job. The next two exhibits show Dr. Neumark's results for female Software Developer 4 employees only. Again, there are a substantial number of women earning more than his model predicts.

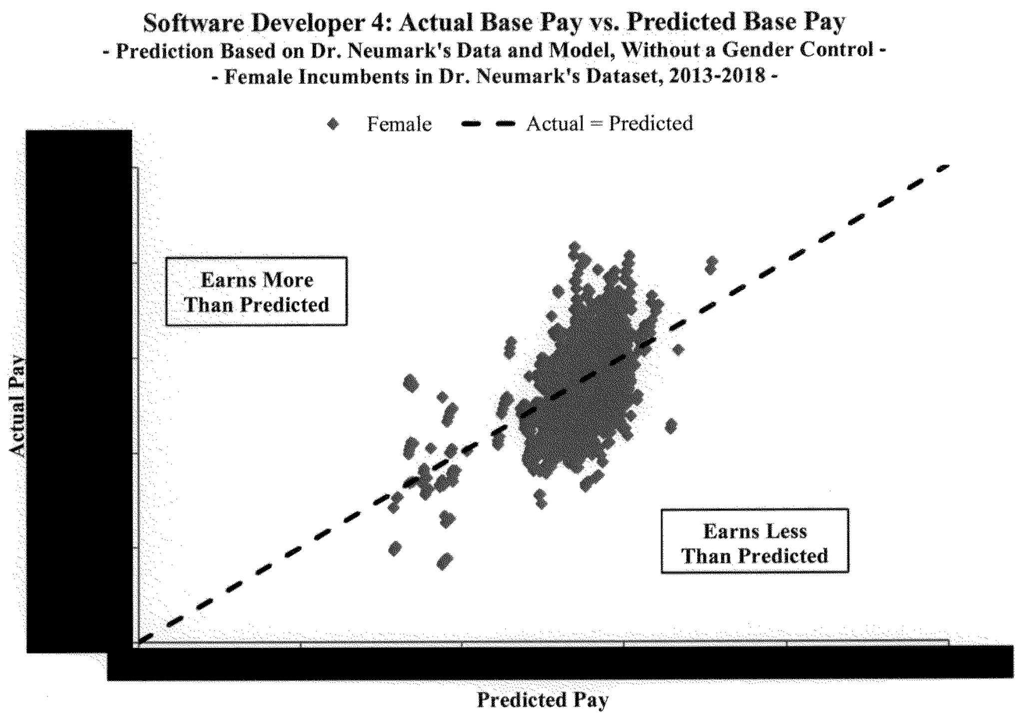


Exhibit 8

Software Developer 4: Percent Difference Between Actual Base Pay and Predicted Base Pay

- Prediction Based on Dr. Neumark's Data and Model, Without a Gender Control -

- Female Incumbents in Dr. Neumark's Dataset, 2013-2018 -

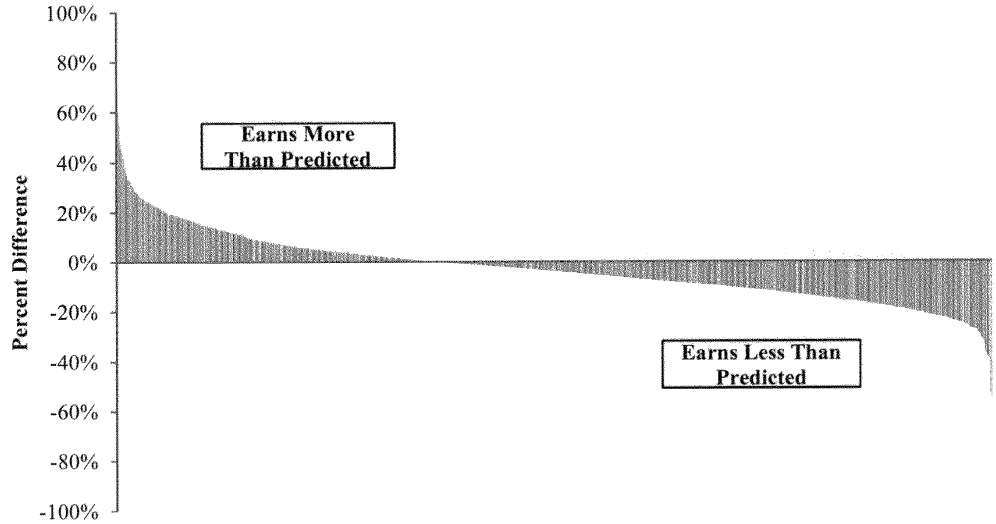


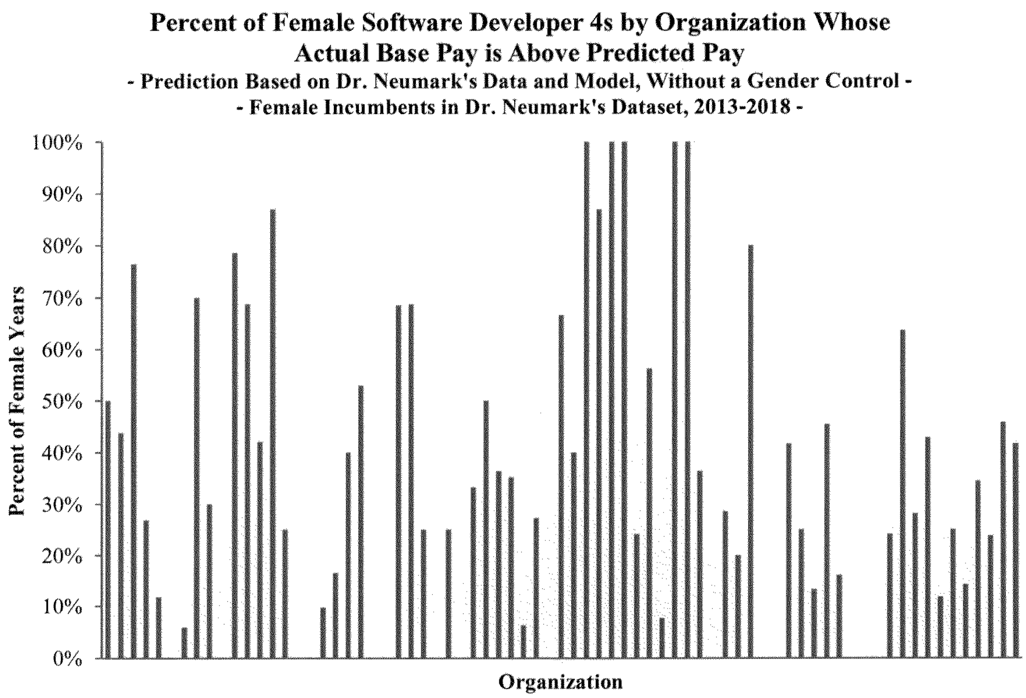
Exhibit 9

50. As illustrated below, outcomes differ a great deal by organization for Software Developer 4s, suggesting that the experiences of women in one part of the company may not explain what is experienced by women in another part of the company.[38] This calls into question the use of Dr.

[38] Organizations indicate cost centers. "Oracle organizes its business, teams, and employees through a financial and accounting hierarchy. This financial and accounting hierarchy mirrors the managerial hierarchy at a high level but often diverges from the managerial hierarchy at a more granular level. That divergence occurs because managers may oversee more than one product team, as that team is defined for the purposes of the financial and accounting hierarchy. Conversely, what is a single product team for financial and accounting purposes may have multiple managers. **At the most granular level of the financial and accounting hierarchy, "cost center" (sometimes called "organizations") are used for purposes of tracking budget and other financial outcomes. A cost center can encompass a single product or service team, but not every product or service team** has its own cost center." Miranda Declaration, paragraph 8.

This is also discussed in the data correspondence between Ms. Mantoan and Mr. Finberg: "**Cost**

Neumark's fully aggregated model as a basis for the pay practices of Oracle as they relate to women. The exhibit below shows the variation in the percent of women within organizations who earn more than Dr. Neumark's model predicts for them.



Note: Chart is limited to organization units with at least 10 female Software Developer 4 observations in the base salary regression population and accounts for 62.6% of female Software Developer 4 observations.

Exhibit 10

---

centers are developed, altered, or deleted in partnership between finance, the business, and HR. These groups work together to organize jobs by product or service, and use the resultant cost centers for purposes of tracking budget, allocating pools of money that can be used for salary increases or bonuses, and tracking other financial outcomes. Not every product or service team at Oracle has its own "Organization_Name," however." August 17, 2018 letter to James Finberg, [Oracle] Mantoan ltr to [Jewett] Finberg in resp to data Qs 21, 22, 26.pdf, p. 3.

51. The characteristics of the differences between actual and predicted pay for women, and whether that difference is statistically significant, can be depicted by organization. The pie chart below looks, by organization, at the sign and significance of unexplained pay differences that come directly from Dr. Neumark's model. I restrict the analysis to organizations of at least ten employees and two women for convenience, but there is no issue with small sample sizes: the power of the statistical tests depends on Dr. Neumark's model and data, not the number of employees in an organization.[39]

52. The results show that in most organizations, all women earn about what Dr. Neumark's model predicts (i.e., the difference between actual pay and the pay predicted by his aggregated model is not statistically significant).[40] This is shown in the light blue slice of the pie chart below. The small, somewhat darker blue indicates organizations in which there are equal numbers of women who earn significantly more than predicted and who earn significantly less than predicted. The darkest blue slice represents organizations in which more women earn statistically significantly above the model's prediction than there are earning significantly below predicted. The red slice indicates the share of organizations in which more women earn significantly less than predicted than earn significantly more. These results are generated using Dr. Neumark's model with its flaws included, but even in that model, it is apparent that organizations with more negative than positive results for women are in the minority.

---

[39] His regression is estimated by year so that statistical tests can be run without having to be concerned about multiple observations on the same person. Once the results are estimated by year, the data are aggregated again so that the pie chart covers all years.

[40] The high spread of the data and that most are not statistically significantly different from predicted is not incompatible. There is an expected 95% of normally distributed data within the confidence interval around a mean, but it is notable that men and women do not have wildly different proportions, and that the error between actual and predicted pay is large for many observations.

Organizations: Base Pay Outcomes for Women Classified by Relationship of Actual to Predicted Base Pay
- Dr. Neumark's Model Applied to Dr. Neumark's Data -

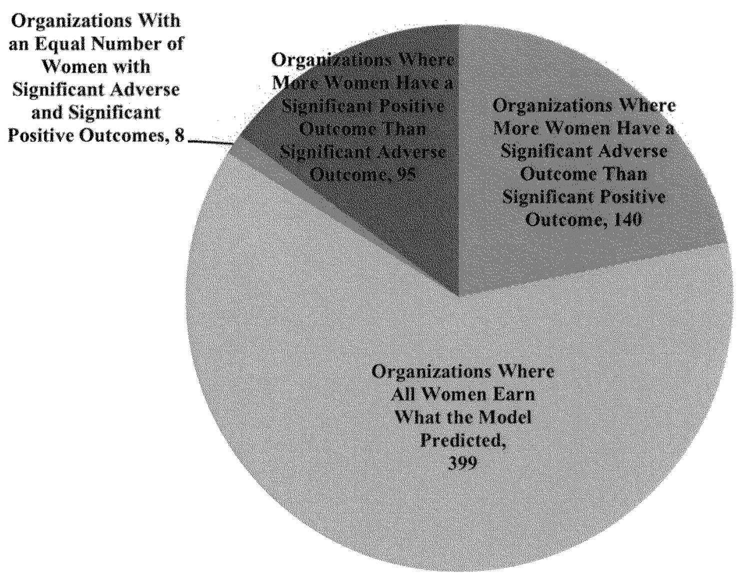


Chart is limited to organization units with at least 10 employees and 2 females, accounting for 96.4% of female employees.

Exhibit 11

53. When the pie chart is instead redrawn to show the percent of women instead of the percent of organizations in which more women do better or worse than Dr. Neumark's model predicts, it remains the case that fewer than half of women are in organizations where more women are paid significantly less than Dr. Neumark's model predicts.

Base Pay Outcomes for Women by Organization Classified by Relationship of Actual to Predicted Base Pay
- Dr. Neumark's Model Applied to Dr. Neumark's Data -

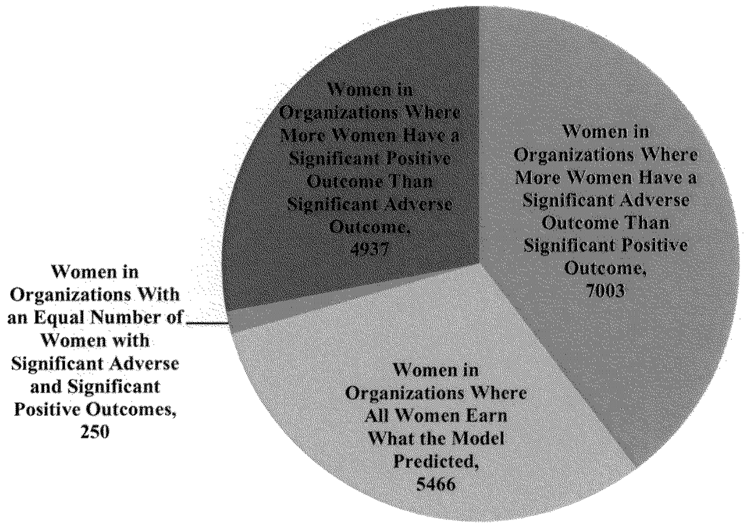


Chart is limited to organization units with at least 10 employees and 2 females, accounting for 96.4% of female employees.

Exhibit 12

54. When the same exercise is conducted for total compensation, the results are even more pronounced. Only a small sliver of organizations have more women who earn significantly less than predicted than who earn significantly more than predicted.

Organizations: Total Compensation Outcomes for Women Classified by Relationship of Actual to Predicted Base Pay
- Dr. Neumark's Model Applied to Dr. Neumark's Data -

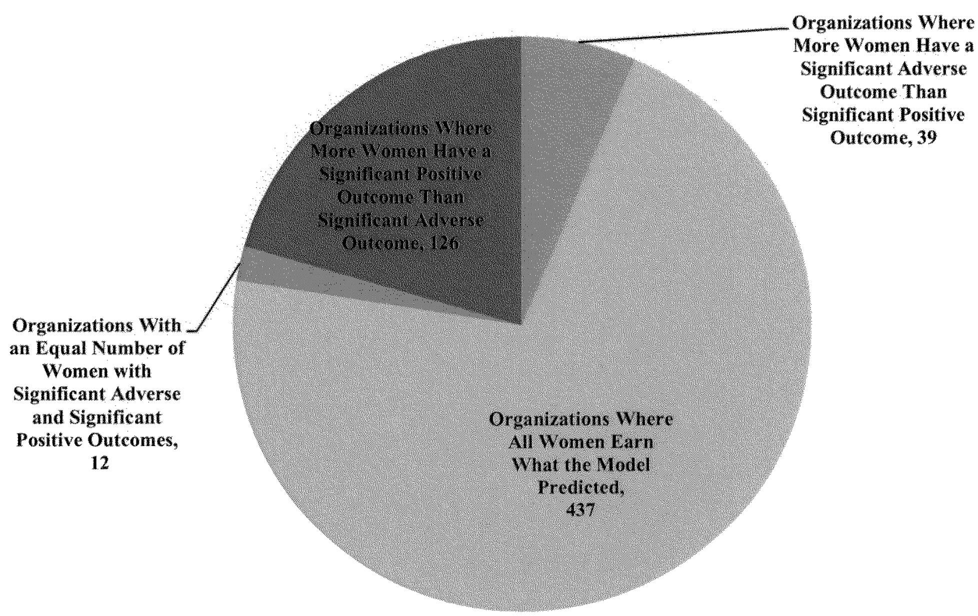


Chart is limited to organization units with at least 10 employees and 2 females, accounting for 95.8% of female employees.

Exhibit 13

Total Compensation Outcomes for Women by Organization Classified by Relationship of Actual to Predicted Base Pay
- Dr. Neumark's Model Applied to Dr. Neumark's Data -

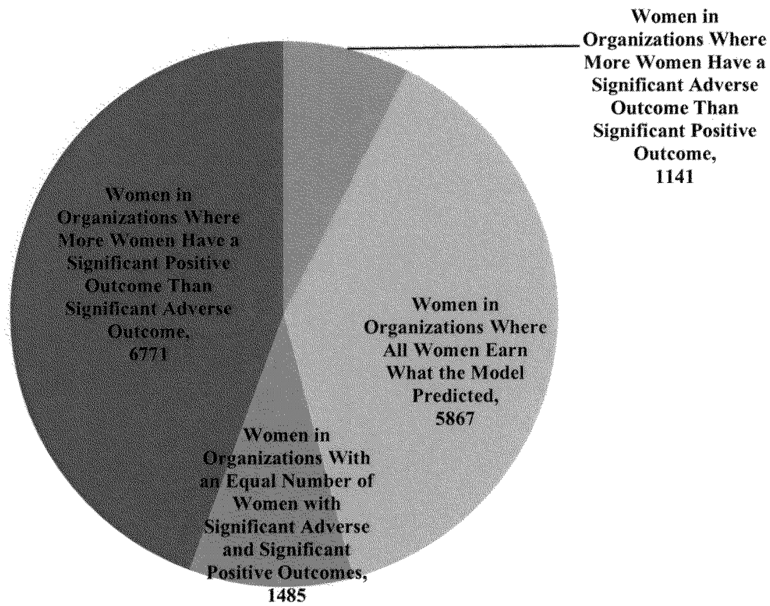


Chart is limited to organization units with at least 10 employees and 2 females, accounting for 95.8% of female employees.

Exhibit 14

55. This is true even when the pie chart is redrawn to show the percent of women who work in organizations in which more women do significantly worse than predicted (according to Dr. Neumark's model) than do significantly better. This is not a surprise, considering the wide underlying variation in total compensation. The question is what drives that variation.

56. The data suggests that specific skills are being called for, not general experience (like the age minus 22 measure of experience used by Dr. Neumark). Consider the relationship between

age and starting pay for Software Developers at various Career Levels.[41] In many industries and occupations, one sees pay rising with age and experience. However, as the graph below shows, starting pay for Software Developer 3s ranges between roughly [redacted] and average starting pay bears no clear and obvious relationship to age.[42]

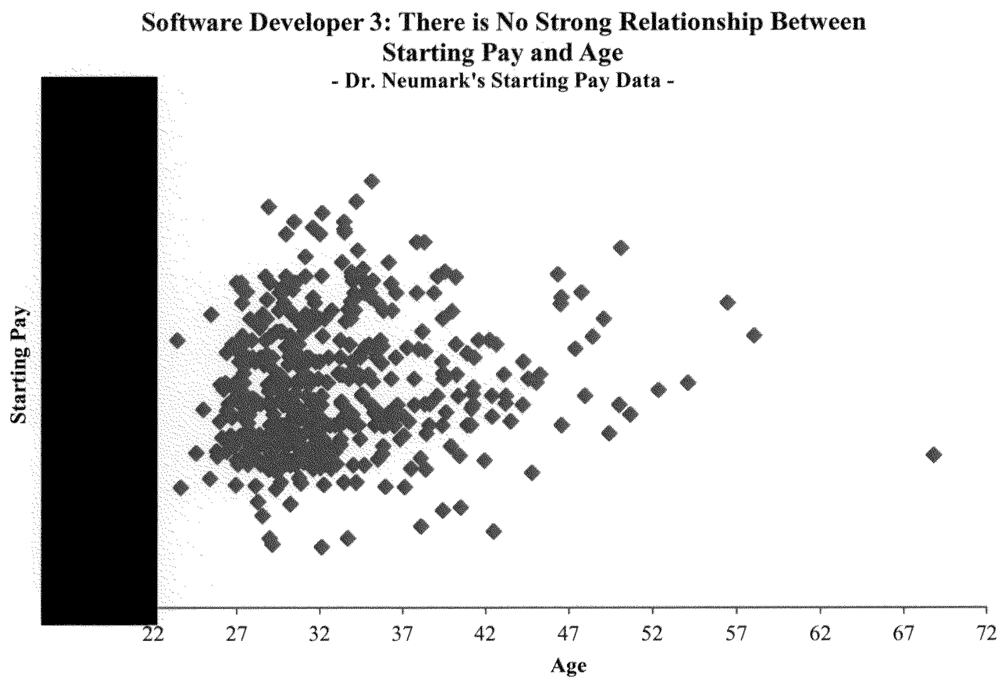


Exhibit 15

---

[41] Because of the flawed way Dr. Neumark defined college hires, similar graphs for IC2-level employees would be mostly made up of college hires that he failed to remove from his data. Thus, I do not present these charts.

[42] A regression of starting pay on age reveals a coefficient of 0.002, meaning that starting pay is associated with a 0.2% increase for each year of age. While statistically significant, a coefficient this small indicates that there is no meaningful relationship between age and starting pay. The model explains less than 1% of the variability in starting pay.

57. The same patterns hold for Software Developer 4s and Software Developer 5s.[43] Fifty year old applicants being hired do not tend to earn meaningfully more than 25 year olds. Rather, an alternative hypothesis is that new hires of all ages are paid differently based on their particular skills rather than based on the number of years they have worked in the labor market.

Software Developer 4: There is No Strong Relationship Between Starting Pay and Age
- Dr. Neumark's Starting Pay Data -

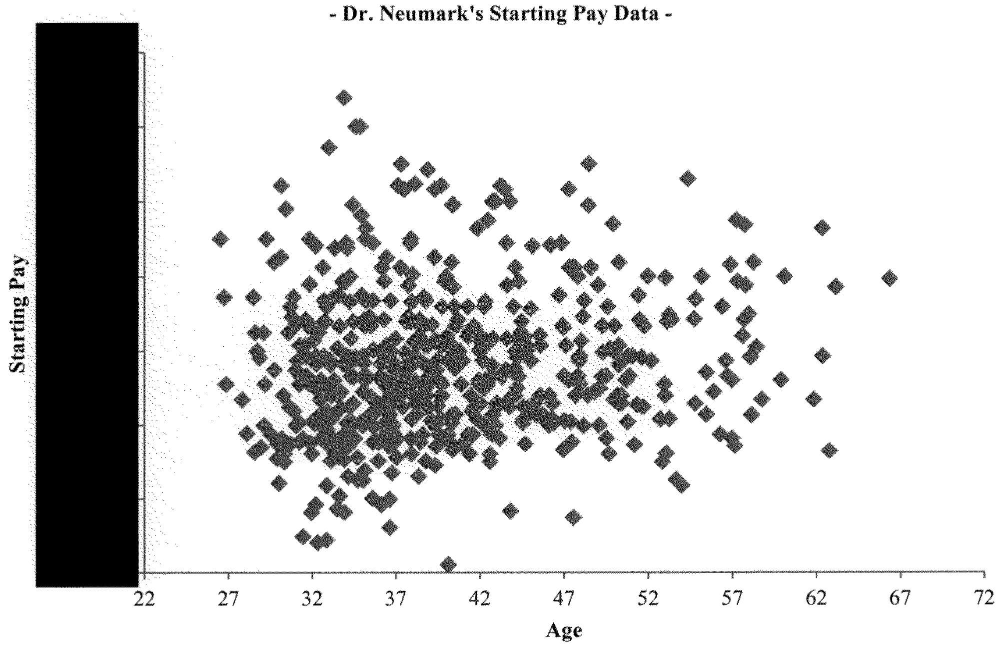


Exhibit 16

[43] Among Software Developer 4s, the regression coefficient on age is 0.002, for a 0.2% change in starting pay for each year of age. Like the Software Developer 3 results, the coefficient is statistically significant but not practically significant in a real world sense. For Software Developer 5s, the coefficient on age is -0.0001 and is not statistically significant.

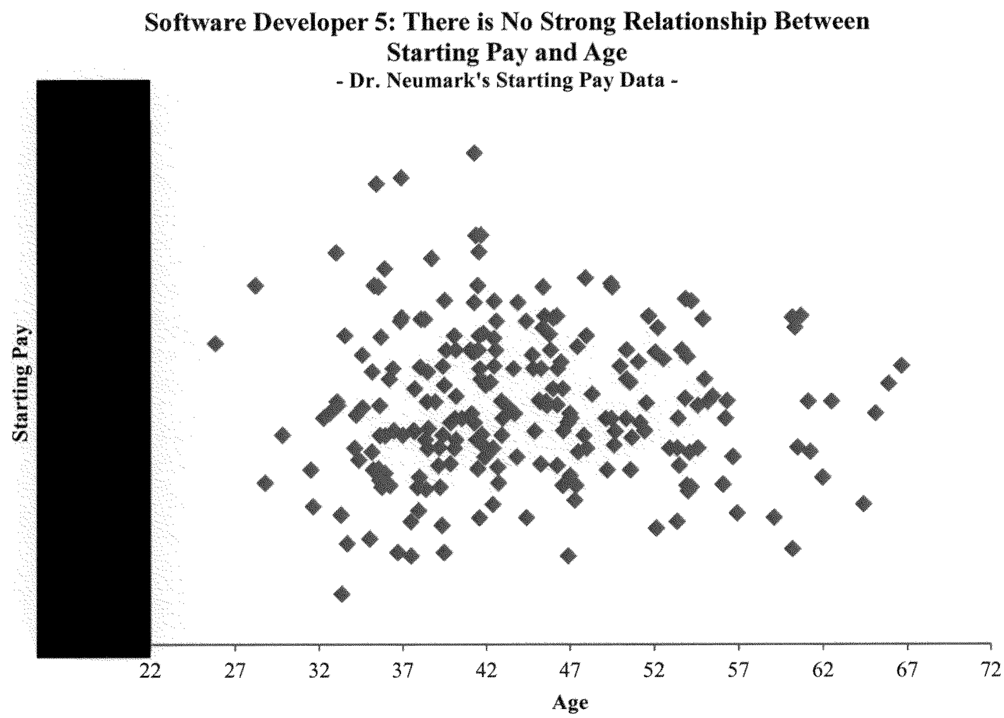


Exhibit 17

58. That specific skills might be more important than simply age and general years of experience is suggested by the graph below which color codes observations for Software Developer 4s who were hired into two different organizations. The red dots indicate those hired into OCI Development. The blue dots indicate those hired into Corp Architecture – Development.

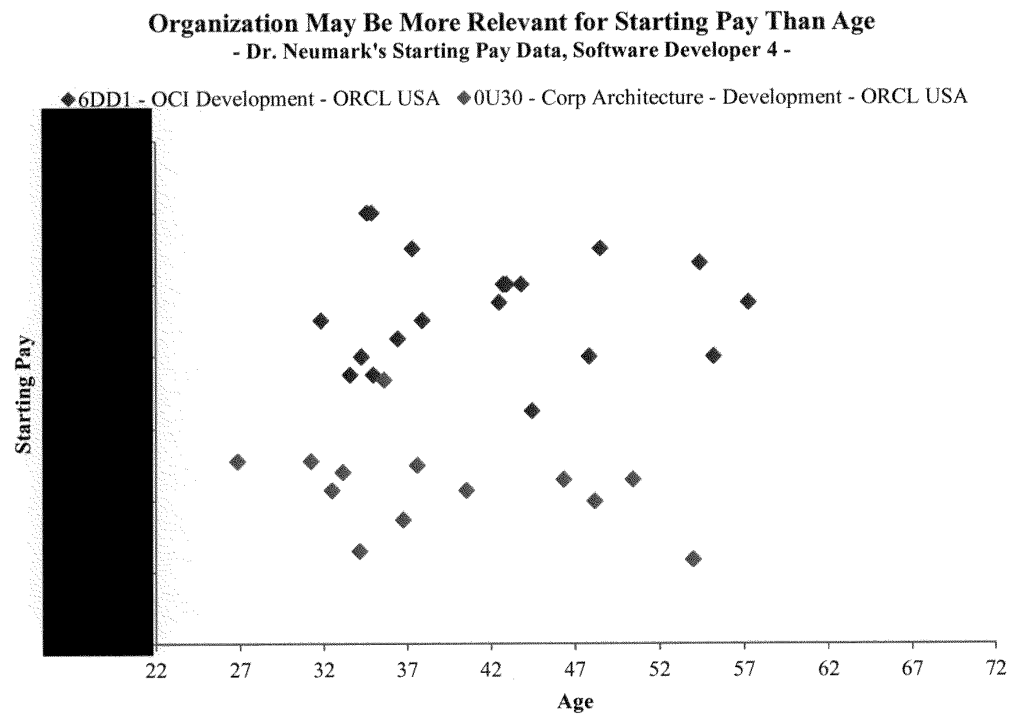


Exhibit 18

59. To view these highly varying outcomes using Dr. Neumark's model another way, I used a statistical algorithm to randomly sample one woman in each of the ten largest job titles in the data who have at least two male comparators and plotted their base pay amounts in 2016 using Dr. Neumark's data. The comparators are closely matched on his regression model variables: experience within two years, Oracle tenure within two years, job code tenure within two years, job code and grade, part time and hourly statuses, zip code, and line of business head. The horizontal axis is job tenure, and the vertical axis is base salary for 2016. The red dot is the randomly selected woman. The blue Xs are men with regression-model characteristics matched

to the selected woman. The blue dots are other women with similar regression-model characteristics. I show four charts below but all ten are included in Attachment D.

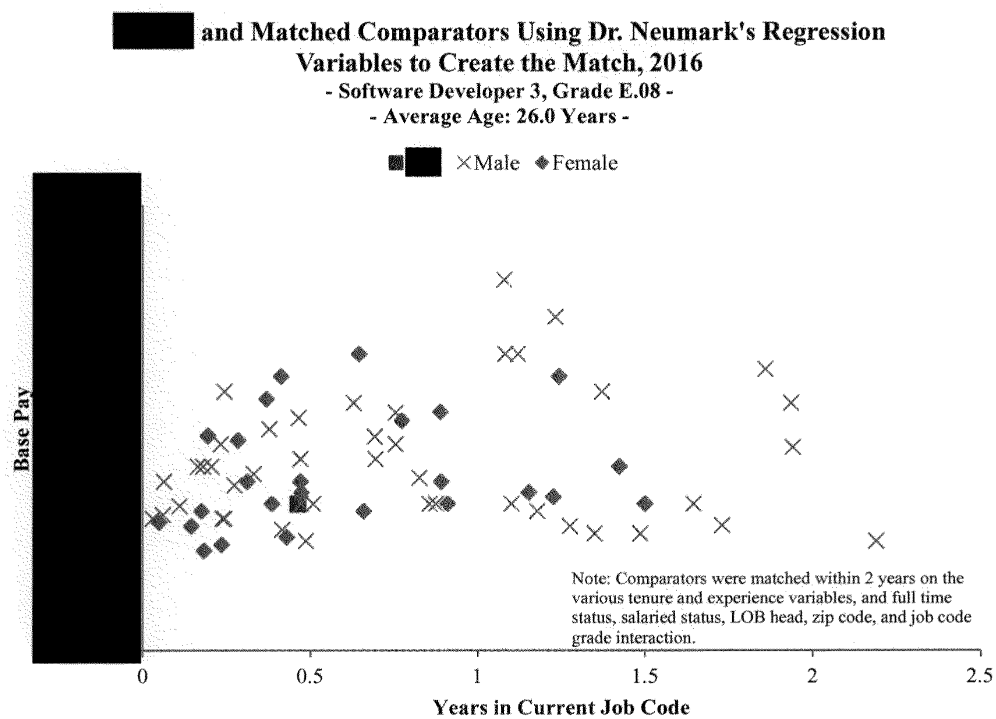


Exhibit 19

[redacted] and Matched Comparators Using Dr. Neumark's Regression Variables to Create the Match, 2016
- Product Manager/Strategy 5-ProdDev, Grade E.011 -
- Average Age: 43.2 Years -

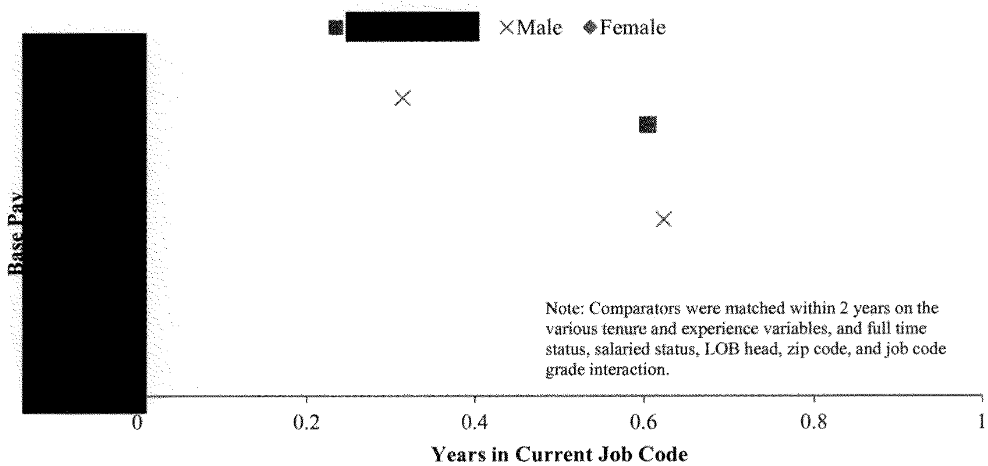


Exhibit 20

[redacted] and Matched Comparators Using Dr. Neumark's Regression Variables to Create the Match, 2016
- Technical Analyst 4-Support, Grade E.12 -
- Average Age: 52.1 Years -

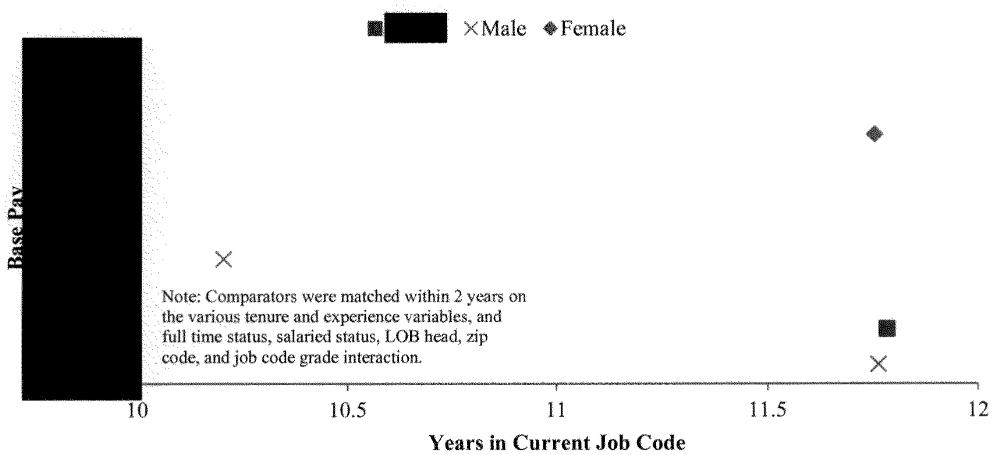


Exhibit 21

[redacted] and Matched Comparators Using Dr. Neumark's Regression Variables to Create the Match, 2016
- Software Developer 4, Grade E.9 -
- Average Age: 39.6 Years -

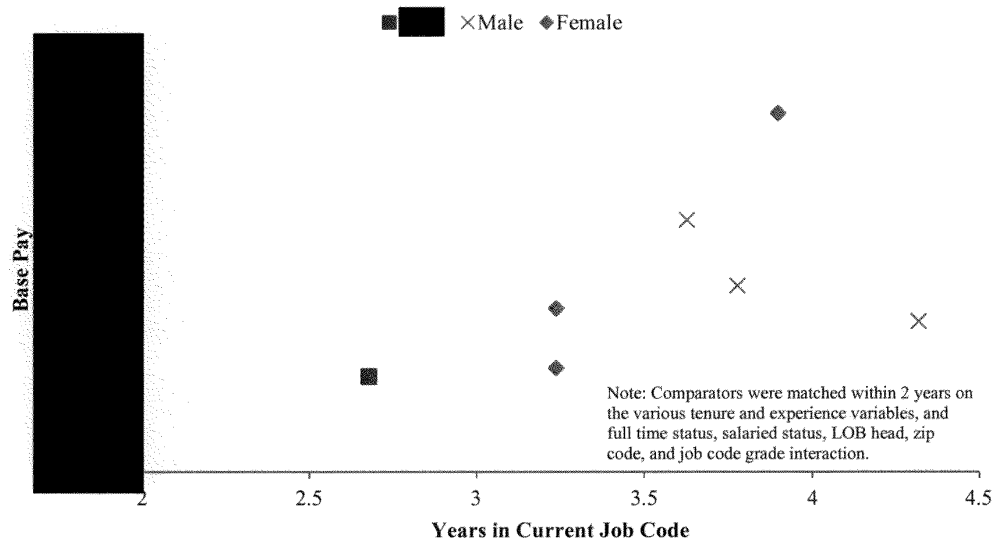


Exhibit 22

60. What we see are a variety of outcomes: pockets where women are both the highest and lowest earners, others in which women earn less or more in general, and yet others with no pattern by gender. Women do not appear to be systematically lower earners even in these groups defined by Dr. Neumark as comparators. For example, [redacted] earns less than some men with the same or less experience in the job code, but she also earns more than men with more experience in the job code. [redacted] earns less than a man with several months less experience, but earns considerably more than a man with roughly the same experience. [redacted] earns more than a man with similar experience and less than a man with less experience, but the highest earner on the chart is a woman. [redacted] is the lowest earner in her group, but another woman is the highest earner.

61. The question is how much of the observed variation is due to genuine unexplained variation in outcomes among women, and how much is due to deficiencies in Dr. Neumark's regression model – in particular, deficiencies in the way he defines "substantially similar" work. Dr. Neumark agreed that employees with different marginal productivity would understandably receive different pay[44] but beyond job code/job grade controls, he did not account for this other than purportedly in his line of business head control.[45] Yet as the graph below shows, the starting pay for Software Developer 4 hires in Thomas Kurian's line of business during the 2013 through 2018 period varies significantly, from just over [redacted] across different organizations.

Average Starting Base Pay for Software Developer 4 in Thomas Kurian's Line of Business, by Organization Within his LOB

[redacted]

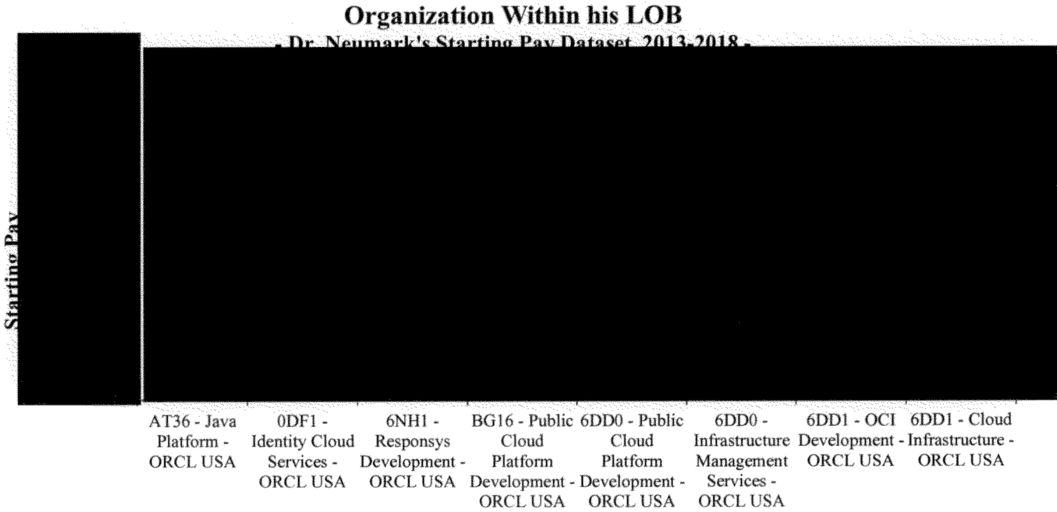


Chart is limited to organizations with at least 10 Software Developer 4 employees in Mr. Kurian's Line of Business in Dr. Neumark's starting pay data.

Exhibit 23

---

[44] Neumark Deposition. 111:16-113:16.
[45] Ibid, 120:21-122:10.

Within job titles, skills and responsibilities vary widely

62. According to Dr. Neumark, work location, job code/grade and line of business head group "similar workers in similar jobs."[46] Although job code and line of business head can be used to segment the data to a certain extent, it appears that employees performing dissimilar work continue to be grouped together using this approach. Organization, or cost center, was used above to group employees as an example but may not be well suited to group employees doing similar work, due to its dual business and accounting function.[47] In order to test Dr. Neumark's theory that job content does not vary within a job code/grade and line of business head, a subset of the new hire requisitions was analyzed to determine whether there are other ways to think about the differences between the job requirements within one of Dr. Neumark's largest groupings.[48]

63. Several studies have used clustering algorithms to extract skill requirements from the text of job requisitions, with a particular emphasis on identifying the specific skills required for different types of IT jobs. Much of this research stems from a need to identify high demand skills in the face of rapid change in the types of skills required by IT jobs.

64. Woweczko (2015) analyzed online job advertisements in Ireland to extract information on skills needs from job descriptions, and presents word clouds[49] showing the top bigrams[50] for

---

[46] Neumark Report, paragraph 34. The notes to his Exhibit 13 provide a list of variables he calls "job characteristics." These include job code and job grade interaction (codes generally map to only one grade per year), zip code, and dummy variables for whether the person was paid on an hourly basis or was part-time.

[47] "At the most granular level of the financial and accounting hierarchy, "cost center" (sometimes called "organizations") are used for purposes of tracking budget and other financial outcomes. A cost center can encompass a single product or service team, but not every product or service team has its own cost center." Miranda Declaration, paragraph 8.

[48] The requisition data contains information relating to job listings and included generic company information, as well as detailed text that described the specific job requirements. The generic text was not analyzed. Rather, the job specific detailed text was analyzed for this analysis.

[49] "Word cloud" is a term of art used to visually depict the importance of each word, where

seven different IT occupations. Woweczko concludes that the skills extracted using this method are more detailed than what would be found in standard occupational descriptions.[51]

65. Litecky, et al. (2010) examined online listings for software engineers on Monster.com, HotJobs.com and SimplyHired.com, finding that "even a brief examination of these tools shows that US job titles vary substantially and that job definitions are often misleading."[52] Their study used cluster analysis of job skill terms found in the listing text and identified 20 IT job categories and associated skill sets. They found that among the advertisements analyzed there were five clusters for software developers: "The software developers group consists of five clusters of traditional non-Web-based development, with moderate demands for programming in general, software development, and object-oriented programming skills, plus specific language skills such as C/C++, Java, or C#. For example, two clusters focus on C/C++ and generic programming skills. The two clusters are distinguished through the supplementary skills required for those jobs. C/C++ programmer jobs focus primarily on programming-language skills, whereas the system-level C/C++ programmer jobs also require skills in general programming, software development, operating systems, security, and Perl. This indicates that the latter cluster undertakes work at the operating systems level as well as supporting traditional Perl-based work."[53] In this case, the word cloud analysis revealed differences in skill requirements for different segments of the software developer job spectrum.

---

importance is measured using word frequency within and across documents calculated by the clustering technique. Less frequent words may appear larger if the algorithm determines they are more important.

[50] A bigram is a pair of consecutive written elements, in this case two consecutive words in a field of text.

[51] Woweczko, Izabella A. (2015) Skills and Vacancy Analysis with Data Mining Techniques, *Informatics*, 2, pp. 31-49.

[52] Litecky, Chuck, et al. (January/February 2010), Mining for Computing Jobs, *IEEE Software*.

[53] Ibid, p. 80.

66. Creating economic variables from text based sources is not new. Economists have a long history of utilizing coded text data in their analyses. One familiar example is the data on workers' occupations and industries collected by the US Census Bureau.[54] The census questionnaire asks respondents "What kind of work was this person doing?" and "What were this person's most important activities and duties?" with a "fill-in-the-blank field" that allows a free-form response. There is no drop down menu option for respondents to choose from. Rather than let respondents decide what their occupational category is, the Census Bureau applies their expertise in the nature of work and what occupation it constitutes to convert free form text descriptions of what people say they do at work to a census OCC code. In the case of the Census, the written responses are then reviewed and coded into standardized occupation classifications, which can then be included as categorical or stratifying variables in quantitative analyses. Similarly, the questionnaire asks about the industry in which one works using both free-form and check-box questions which are then clerically coded by Census Bureau staff.[55] The resulting coded occupations and industries can then be utilized by economists and other researchers in their analyses.

67. I have in my previous work performed conversion of detailed textual descriptive material into job categories. For example, in a hiring case I and my team processed 30,000 handwritten employment applications and created a set of job categories. These categories were then used in statistical analysis of hiring. In another case, I and my team processed tens of thousands of promotion job postings, and converted qualitative material into data that would be subjected to statistical analysis. In short, processing of text and other qualitative material into quantitative or categorical formats is nothing new.

---

[54] United States Census Bureau: Industry and Occupation (https://www.census.gov/topics/employment/industry-occupation/about/occupation.html).
[55] Ibid.

68. Economists and other professionals have increasingly incorporated in their research analysis of text-based data sets to extract and classify textual information.[56] Some of these studies have focused on using textual analysis to examine media sentiment,[57] policy uncertainty,[58] and the health and stability of financial systems.[59] Economists have utilized text data derived from analysis of Google searches,[60] Yelp reviews,[61] and Twitter messages[62] in empirical analyses.

69. Here, I use these techniques to analyze the 1,053 detailed text job requisitions for the Software Developer 4 job code at Oracle. Following methodology that is typical in the application of text processing, the job posting text was prepared for analysis by removing what are referred to as stop words, as well as punctuation and irregular characters that are not useful

---

[56] See, for example: Einav, Liran and Jonathan D. Levin (2014) The Data Revolution and Economic Analysis. *Innovation Policy and the Economy*, 14, pp. 1-24; and Gentzkow, Matthew, Bryan T. Kelly and Matt Taddy. (Forthcoming) Text as Data. *Journal of Economic Literature*.

[57] See, for example: Gentzkow, Matthew, Jesse M. Shapiro and Michael Sinkinson (2014). Competition and Ideological Diversity: Historical Evidence from US Newspapers. *American Economic Review*, 104(10), pp. 3073-3114; Gentzkow, Matthew and Jesse M. Shapiro (2010), What Drives Media Slant? Evidence from U.S. Daily Newspapers. *Econometrica*, 78(1), 35-71; and Groseclose, Tim and Jeffrey Milyo, A Measure of Media Bias. *The Quarterly Journal of Economics*, 120(4), pp. 1191-1237.

[58] See Baker, Scott R., Nicholas Bloom and Steven J. Davis (2016), Measuring Economic Policy Uncertainty. *The Quarterly Journal of Economics*, 131(4), pp. 1593-1636.

[59] See, for example: Romer, Christina D. and David H. Romer (2017) New Evidence on the Aftermath of Financial Crises in Advanced Countries. *American Economic Review*, 107(10), pp. 3072-3118; and Born, Benjamin, Michael Ehrmann and Marcel Fratzscher. (2013) Central Bank Communication on Financial Stability. *The Economic Journal*, 124, pp. 701–734.

[60] See, for example: Chae DH, Clouston S, Hatzenbuehler ML, Kramer MR, Cooper HLF, Wilson SM, et al. (2015) Association between an Internet-Based Measure of Area Racism and Black Mortality. *PLoS ONE* 10(4):e0122963; and Saiz, Albert and Uri Simonsohn (2013) Proxying for Unobserved Variables with Internet Document-Frequency. *Journal of the European Economic Association*, 11(1), pp. 137-165.

[61] Taddy, Matt. (2015) Distributed Multinomial Regression. *The Annals of Applied Statistics*, 9(3), pp. 1394-1414.

[62] Taddy, Matt. (2013) Measuring Political Sentiment on Twitter: Factor Optimal Design for Multinomial Inverse Regression. *Technometrics*, 55(4), Special Issue (November 2013), pp. 415-425.

for analysis.[63] Hierarchical clustering, a type of machine learning algorithm, was applied to the text in the qualifications section of the requisitions data to identify similarities and differences between words used to describe the job requirements of each requisition.[64] The algorithm calculates these similarities and differences found in the text by determining the uniqueness of words using a mathematical equation. No analyst judgement is applied at the requisition level.

70. The measure used here to evaluate the importance of a specific term or word is called "Term Frequency, Inverse Document Frequency" (TF-IDF). The TD-IDF is equal to the term frequency weighted by the fraction of documents the word appears in. Technically, the TF-IDF score of a word equals the frequency of word multiplied by the log of the ratio of the number of documents to the number of documents with that word. The algorithm places a higher value on words that from their frequency appear to delineate required skills within subsets of requisitions – such as "cloud" or "fusion."

71. For example, the word "Oracle" appears in almost all requisitions and thus does not provide any information for distinguishing among requisitions. A word's "importance" is scored by combining the frequency of a word in a document, adjusted by the frequency with which it appears in the other documents. Suppose we have a sample of 100 requisitions. Suppose the requisition we are looking at includes the word "computer" 10 times and the word "manage" twice; assume 97 of the other requisitions for this job code also include the word "computer" and just 9 include the word "manage." We calculate the TF-IDF score of the word "computer" by computing "10 * ln(100/97)" which is equal to 0.274. The TF-IDF score of the word "manage" is calculated as "2 * ln(100/9)" which is equal to 4.816. If a particular term appears in every

[63] Stop words are commonly used words such as "a," "the," "is," etc.
[64] The clustering algorithm was applied to all Software Developer 4 requisitions before restricting the data to Mr. Kurian's line of business.

document then it is not useful for distinguishing between subsets of documents; the TF-IDF score for that word equals zero and it is not given any weight.

72. Ultimately the algorithm clusters similar requisitions into groups that are most similar based on the importance and frequency of the specific terms contained in the descriptions. The analysis applied to the Software Developer 4 requisitions resulted in the creation of 15 unique clusters.

73. The first indication of differences between the clusters can be seen by examining the average starting salary across clusters in the graph below. If one were to place all fulltime Software Developer 4 requisitions from Mr. Kurian's line of business into one group, the overall average starting salary would be roughly [redacted]. However, after clustering the requisitions by the descriptions, it is evident that there are distinct differences in starting pay within the Software Developer 4 requisitions working in Mr. Kurian's line of business. As the chart shows, there is a range of average starting salaries between employees in each of the clusters ranging from an average starting salary of [redacted] in Cluster 9 to an average starting salary of [redacted] in Cluster 14.

Average Starting Salary By Requisition Cluster	
Cluster*	**Average Starting Salary**
1	[redacted]
2	[redacted]
3	[redacted]
4	[redacted]
5	[redacted]
6	[redacted]
7	[redacted]
8	[redacted]
9	[redacted]
10	[redacted]
11	[redacted]
12	[redacted]
13*	[redacted]
14	[redacted]
15*	[redacted]

* No employees from clusters 13 or 15 worked in Thomas Kurian's line of business.

Exhibit 24

74. The differences between the clusters can be seen when the text in the qualifications portion of the requisitions is depicted by importance of words in a cluster in a visual "word cloud."[65] The word clouds for all 15 clusters of requisitions for Software Developer 4s are in Attachment E but I will discuss two clusters here as examples. Each word cloud below presents the 50 most important words per cluster, with the most important terms being presented in large blue or purple font, and the less important terms being presented in small red font. When visually comparing the word clouds, it is evident that there are distinct differences in the importance of terms that appear in each of the clusters.

[65] For the purpose of presenting terms or words in a word cloud, important terms are identified as those with the highest proportion in a cluster minus their proportion across all clusters.

Cluster 7

Exhibit 25

75. The word cloud above is based on the 52 requisitions in Cluster 7, which has an average starting salary of [redacted]. The most term with the greatest weight that appears in Cluster 7 is "test," which suggests that this cluster of requisitions relates to testing applications that are developed by others. Closer manual inspection of the text in the qualifications section of the requisitions supports this finding. As just one example, a portion of the responsibilities section of requisition IRC2487806 in Cluster 7 states, "As a software test automation engineer, one is responsible for developing backend automation tests cases (from scratch). Will perform a wide variety of testing from performance, functional, load and reliability testing. Engineers will report, analyze, troubleshoot bugs and work with development team for resolution. Must be highly passionate about tearing software apart and finding defects bugs."

Cluster 14

Exhibit 26

76. The average starting salary of Cluster 14 is higher, at [redacted]. The chart above shows that the incidence of terms for Cluster 14 is different than those that appear in Cluster 7. The highest weighted term in Cluster 14 is "systems," with "distributed" and "iaas" also being common terms.[66] The prominent terms that appear in Cluster 14 indicate that this group of requisitions is associated with developing Oracle's cloud based infrastructure. For example, an excerpt of requisition IRC3037555 in Cluster 14 states, "Oracle's Infrastructure Cloud Container Service is Hiring Software Engineers Oracle is building a from scratch Infrastructure-as-a-Service (IaaS) Public Cloud, and a large part of that effort is a container service targeting modern

---

[66] My understanding is that IaaS stands for infrastructure as a service and is a form of cloud computing that provides virtualized computing resources over the internet. (See https://www.oracle.com/cloud/infrastructure.html)

DevOps engineers. We are looking for exceptional distributed systems and operating systems engineers to join this effort."

77. The cluster analysis is consistent with the idea that controlling only for job title and line of business head and not more detailed aspects of work does not group employees doing substantially similar work. Because if women are distributed across these clusters differently than men are – for example, women were 29.8% of new hires in Cluster 7 in Kurian's LOB and 7.7% of new hires in Cluster 14 in Kurian's LOB – then not accounting for within-job title differences in skills and responsibilities will lead to omitted variable bias. Because Dr. Neumark does not accurately or fully control for the nature of the work employees are doing, his analysis suffers from measurement error.

PRIOR PAY AND STARTING PAY

Dr. Neumark's prior pay analysis does not show that gender gaps in prior pay *cause* gender gaps in starting pay at Oracle

78. Plaintiffs allege that Oracle has a practice of relying upon applicants' prior pay when setting their starting pay upon joining Oracle. Specifically, in their Class Certification Motion, Plaintiffs claim that "inequities from this illegal practice [of using prior pay to set starting pay] persist. Plaintiffs will prove through common evidence that Oracle's policy of tying salaries to prior pay, and failing to rectify imbalances [that are present and adverse to women generally in the labor market from which Oracle's applicants are derived] violated FEHA and the UCL. Plaintiffs will prove this illegal policy and practice through company documents, testimony of Oracle's persons most knowledgeable, and expert analysis of company data."[67]

---

[67] Representative Plaintiffs' Memorandum of Points and Authorities in Support of Motion for

79. Dr. Neumark presents an analysis of the relationship between prior pay and starting pay. His analysis is flawed and inconclusive, however, because it does not identify an empirical relationship between prior and starting pay reflecting *causality* flowing from prior pay to starting pay. It is entirely possible that both prior and starting pay are associated with employee and employer/job characteristics, and that all Dr. Neumark has done is identify a correlation, and not a causal relationship between prior and starting pay. Dr. Neumark simply assumes that prior pay is collected at Oracle in order to explicitly set starting pay as some sort of direct function of prior pay, but without evidence to support that assumption.[68]

80. Dr. Neumark's Exhibit 39 depicted a scatterplot of prior and starting pay for all applicants for which he states he has the data needed to conduct the analysis. Dr. Neumark then summarizes the results of a simple regression analysis of prior pay on starting pay in his Exhibit 40. In the Table Note at the bottom of his Exhibit 40, Dr. Neumark identifies the results of the regression of Oracle starting pay on applicants' prior pay. In his report, he states the following conclusion:

> "The line [representing the regression line between the two variables] is strongly upward sloping, as is the cluster of plotted points, indicating that starting pay is tightly linked to prior pay. On average, prior pay being higher by $1 predicts that starting pay is higher by $.75. Alternatively, prior pay explains 74% of the variation in starting pay. The likelihood that this strong relationship between prior pay and starting pay occurs by chance is less than 1 in 1 billion, as reflected in a t-statistic on the coefficient on prior pay in the starting pay regression of 89.9 (or an effect on prior pay of 89.9 standard deviations)."

81. Dr. Neumark is careful never to state explicitly that he is inferring that this "analysis" of a correlation between prior and starting pay demonstrates that Oracle relied upon prior pay in setting starting pay, because from a scientific perspective, he cannot. Yet Dr. Neumark states his findings on statistical significance in terms that would make the reader infer that it must be the

---

Class Certification, January 17, 2018, p. 9.
[68] Neumark Deposition, 296:15-24.

case that Oracle relied upon prior pay, because the relationship between prior pay and starting pay is extremely strong. Indeed the relationship is strong, but why? Dr. Neumark's statistical test is against the hypothesis that there is zero relationship. This of course is not the appropriate alternative hypothesis to evaluate the starting pay/prior pay relationship at Oracle. The correct alternative hypothesis, or benchmark, would be a comparison to what would be expected. No one expects there to be zero relationship between prior pay and next observed starting pay for any job, at any employer, in any economy anywhere at any time. That would be an absurd benchmark.

82. According to Plaintiffs' theory, if female prior pay has embedded within it labor market bias against women, such that it follows logically that Oracle has simply embedded that bias in its own initial pay for its female employees. However, Dr. Neumark's results are also consistent with a hypothesis that Oracle does set pay based on the human capital and specific job experience an applicant brings to the job, and that the "disparities" in both prior and initial pay that Dr. Neumark claims to have found in his Exhibit 41 reflect that Dr. Neumark has not included any of the details of prior employment experience or sufficient job controls in his regressions. Dr. Neumark assumes his conclusions, by assuming that somehow an R-squared of 0.74 is high enough to "prove" that Oracle did rely causally on prior pay, and further assuming that the approximately 2% female difference in pay found in both his prior and starting pay regressions shown in his Exhibit 41 reflect outcomes due to being female, rather than gender differences in other characteristics, such as the companies they came from, for example.

83. It is worth examining Dr. Neumark's implied premise, based on his Exhibit 40 that Oracle relies formulaically on prior pay to set starting pay. Dr. Neumark discusses the issue of statistical hypothesis testing in various places throughout his report, and in fact conducts dozens

of hypothesis tests against benchmarks he notes are appropriate for the claims in the case. For example, in each of his pay regressions, he tests the female coefficient for statistical significance. Each of these tests is a test against a benchmark value of 0 – i.e., the test is one where the analyst seeks to know if the numerical value estimated for the gender coefficient differs meaningfully from no value or no relationship at all – that is, that it differs from a value of 0. This is what is referred to as a null hypothesis test. However, there are many other benchmarks which are relied upon in the statistical analysis of employment practices. For example, in a hiring context, one might want to test the percentage of women hired into a job against the benchmark of the rate at which men are hired. Of course this benchmark is seldom equal to zero. Or the analyst may wish to test the rate of promotion for women relative to men, which would mean a female promotion benchmark equal to the rate of promotion for men.

84. While Dr. Neumark carefully describes the many results of the hypothesis tests he performs on a variety of compensation measures using a variety of different regression specifications, he is notably silent on what benchmark is used in his Exhibit 40 comparison of starting pay to prior pay. Clearly, one would not expect this regression coefficient benchmark to be zero. One would also not expect the relationship to be exactly the same (equal to a benchmark of 1). These are the two extremes – either starting pay is completely unrelated to prior pay (practically an impossibility, since employee human capital does not evaporate when moving from job to job, and workers seldom look for jobs completely unrelated to their accumulated human capital), or it is perfectly related to prior pay (also very unlikely, given that job seekers seldom move between jobs identical in every single respect). If neither of these extremes is what one would expect, then what exactly should one expect? Dr. Neumark fails to disclose anything in this regard, and is careful in his wording not to imply it either.

85. Starting pay and prior pay are strongly correlated throughout the economy. This is not unique to Oracle. For example, I reviewed National Longitudinal Survey (NLS) data on prior pay and starting pay for people who changed jobs.[69] The correlation between starting pay and prior pay is 0.75 across all individuals in this data. Dr. Neumark himself appears to understand this.[70] Nothing about Dr. Neumark's simple presentation of the relationship of starting pay and prior pay at Oracle is anything other than what one would expect in any company.

<u>Not only is Dr. Neumark's prior pay-starting pay research design wrong, but Dr. Neumark's analysis is also incorrectly performed, and fixing his errors completely changes the conclusions</u>

86. Dr. Neumark finds a 2.4% gender gap in starting pay at Oracle and a 2.2% gender gap in prior pay, and uses that to argue that "Oracle may be mimicking the gender pay gap reflected in the prior pay of employees who come to Oracle from other employers – especially if the gender gap in prior pay and starting pay is similar."[71] However, the results Dr. Neumark obtains do not hold up when one follows his own advice to use the correct data and focus on "apples to apples" comparisons.

---

[69] The National Longitudinal Surveys (NLS) started in 1997 with 14-18 year olds and surveys them every year about a wide range of topics (the most recent year was 2016). My analysis examines job changes and the difference between the ending pay of the prior job and the starting pay of the new job. After limiting the data to exclude people changing occupation, changing part-time/full-time status, or who have extreme values of the reported hourly rates, I analyze data for 3,488 respondents.

[70] When asked, "Would you expect to see that the prior pay … would be highly predictive of the starting pay …?" Dr. Neumark answered, "[…]the worker has some characteristics, how good they are at the job, and maybe the jobs are related.· They're probably being hired -- at least in -- in a big, complex organization, probably being hired to do a job that has some relationship to the prior job, or the skills they learned on their prior job, and some of those may be -- you know, you take -- your skills go with you, your abilities go with you, so it's not surprising at all." Neumark Deposition, 295:1-24.

[71] Neumark Report, p. 26, paragraph 64.

87. Dr. Neumark warns that the measures of prior pay in the Oracle data can include a variety of possible pay elements: base pay, bonus, and the value of stock. However, for a subset of 425 new hires, he claims that the prior pay measure is clearly identified in the data as relating to base pay. The compensation measure for Oracle starting pay is base pay. For the sample of 425, Dr. Neumark states that he has "apples to apples" measures for the pay figures he wishes to compare. For the other 2,358 applicants in his analysis, there is no indication whether the prior pay provided in the data is base pay or total compensation. His Appendix Table C.1 (page A-30 of his report appendix) gives examples of the issues here. One person reported $135K+25K bonus, and presumably the $135K is just base pay. But the person who reported $118,112.48 (inclusive of bonus) cannot be assigned a base pay amount. It is not clear if the person who reported "105000 / 20% bonus" is including the 20% bonus in that dollar amount or if the 20% bonus is an additional amount. Someone who reports just "135,000" may or may not be including bonuses. The person who reported "$190K (base + on-target bonus)" clearly combined the two.

88. Dr. Neumark states, "[...] to create an apples-to-apples comparison with the current measure of starting pay (which uses base pay), I attempt to use prior base pay whenever base pay is explicitly reported (425 employees). However, for most employees (2,358), it is ambiguous whether the salary number given is base pay or total compensation."[72] Yet he never presents any statistical results using just the 425 employees for whom he says he has an "apples to apples" comparison.[73] Instead, the results in his Exhibit 41 are based on all 2,783 observations. His Appendix D, which is intended to test the robustness of his results, only drops observations for which prior pay was in a foreign currency and 62 observations where "prior pay was ambiguous"

[72] Neumark Report, p. 26, paragraph 67.
[73] In other words, even among the 425 "apples to apples," it is not entirely clear whether the dollar amount is just base pay or includes bonuses, stock or any other compensation.

without explaining why they were ambiguous.[74] In other words, in the text of his report, he describes 2,358 prior pay observations as "ambiguous" but only drops 62 as "ambiguous" in his sensitivity test. Below, I test his Exhibit 41 results using the group of 425 employees he claimed would present an "apples to apples" comparison.

89. The first three columns in the table below show Dr. Neumark's results, as reported in his Exhibit 3 on page 34 of his report. The next three columns repeat his analysis, but restrict the population to "apples to apples."[75] The results are quite different when using the "apples to apples" base pay measures. None of the results are statistically significant but the point I want to emphasize is about the relative sizes of the coefficients, which are what Dr. Neumark uses to fashion his argument. The gap in starting pay among the "apples to apples" group is similar to the gap in his full 2,783 population: women's starting pay is 2.16% less than men's, based on his model. However, the gender gap in *prior pay* is much larger: women's prior pay is 4.91% less than men's. The third column shows that the difference between starting pay and prior pay regressed on the same controls shows that women actually do 2.74% better than men upon being hired by Oracle, according to the way Dr. Neumark has set up his analysis.

[74] Neumark Report, page A31.
[75] Three observations drop from the analysis because they are missing information in one or more of the control variables, leaving 422 for analysis.

Dr. Neumark's Prior Pay Results are Driven by Measurement Error

		Neumark Ex. 41 Analysis Sample			Dr. Neumark's "Apples to Apples" Subset		
		(1)	(2)	(3)	(4)	(5)	(6)
Regression Dependent Variable		ln(Starting)	ln(Prior)	ln(Starting)-ln(Prior)	ln(Starting)	ln(Prior)	ln(Starting)-ln(Prior)
Female	Coefficient	-0.0242***	-0.0218**	-0.0025	-0.0216	-0.0491	0.0274
	Std. Error	(0.0057)	(0.0108)	(0.0091)	(0.0221)	(0.0341)	(0.0261)
	T-stat	-4.24	-2.02	-0.27	-0.98	-1.44	1.05
	p-value	0.0000	0.0439	0.7874	0.3280	0.1509	0.2930
Sample Size		2,783	2,783	2,783	422	422	422
Adjusted R-squared		0.8523	0.6702	0.3526	0.8045	0.6853	0.4521

1. All models control for experience, a dummy for whether foreign currencies were converted, and dummies for data ambiguity in the prior pay data (unclear currencies, unclear fractions, unclear hourly, and unclear total compensation). Job controls include controls for job code-job grade interactions, zip code, LOB head and part-time status.

2. Columns 1 to 3 replicate Neumark Ex. 41 Columns 2 to 4.

3. *** denotes a significance level of 1%, ** denotes a significance level of 5%, * denotes a significance level of 10%.

Exhibit 27

90. The results in Dr. Neumark's report appear to be a coincidence due to measurement error, and not evidence that Oracle "mimics" hypothesized gender discrimination in the labor market writ large. The bigger picture, as noted above, is that nothing in his analysis demonstrates that prior pay has a causal effect on starting pay at Oracle.

91. Simply put, 85% of the data Dr. Neumark relied on to draw conclusions about starting pay and prior pay are "ambiguous" in that he states that it cannot be determined if what is being modeled is prior base pay or prior total compensation. However, upon careful inspection, even the 425 observations Dr. Neumark considers "apples to apples" have problems. The table below lists 12 cases where Dr. Neumark's data is compared to the actual variable he uses to measure prior pay. As seen, Dr. Neumark's mismeasurement of prior pay ranges from a $110,000 under-calculation to a $130,000 over-calculation. For example, the first row shows that Dr. Neumark recorded the employee's prior base pay as $290,000. However, in the variable that stores prior pay, it clearly indicates that base pay is $160,000. In the second row, it is not clear where his

$118,000 value comes from. In the third and fourth rows, it appears that a zero was mistakenly dropped in creating a numeric prior pay variable suitable for analyzing quantitatively out of the prior pay text provided in the data.

Inspection of Dr. Neumark's 425 "Apples to Apples" Subset Indicates that Dr. Neumark Incorrectly Identified Base Pay for 12 Members in this Subset

	Person ID	Neumark's Prior Base Pay	Company Data: CANDIDATE_CURRENT_SALARY_ATV	Corrected Prior Base Pay	Extent of Error	Starting Oracle Pay (2017 $)
1	893816829	290,000	290K (160K Base + 28K Bonus + 108K RSU)	160,000	130,000	[redacted]
2	893432257	118,000	228000 with RSUs	228,000	-110,000	[redacted]
3	887184502	12,000	120000 + 10% average bonus	120,000	-108,000	[redacted]
4	893741761	10,000	$100k + OT	100,000	-90,000	[redacted]
5	889969072	180,000	$128K base salary + 3.5%-5% bonus (last year he got 4.5%)	128,000	52,000	[redacted]
6	891674199	186,000	$162k base, $186 total	162,000	24,000	[redacted]
7	894217345	230,000	$210K Salary + $20K Annual Bonus	210,000	20,000	[redacted]
8	891096869	253,000	$235k base + $500K unvested equity	235,000	18,000	[redacted]
9	893720364	240,000	$295K (includes a $45k bonus)	250,000	-10,000	[redacted]
10	891927319	150,000	155k + 75k RSUs + 15% bonus plan	155,000	-5,000	[redacted]
11	890778700	127,000	$127,500 + $5000 bonus	127,500	-500	[redacted]
12	891938057	159,000	159.5K plus bonus of around 12 to 20K	159,500	-500	[redacted]

Note: Neumark's Prior Base Pay comes from "Base_USD"

Exhibit 28

92. Another problem is that Dr. Neumark's prior pay and starting pay models do not control for year, which is especially problematic given how quickly things change in the tech sector.

The hires in his analysis span a period of six years. While inflation (which Dr. Neumark tracks using the CPI) ranged from less than 1% to 2.1% during the 2013-2018 period, median hourly wage growth in the San Francisco / San Mateo / Redwood City region ranged from 3.1% to 5.3%.[76] When I correct the 12 prior pay errors in his "apples to apples" data and add a year control, the results now show that there is no gender gap in prior pay among new hires at Oracle. The female coefficient in the prior pay regression is under 1% and is not even remotely close to being statistically significant. In sum, Dr. Neumark's analysis of prior pay fails to demonstrate a purposeful or causal linkage between prior and starting pay, and fails to support the conclusion he draws that Oracle relies upon prior pay to the detriment of female applicants hired into Oracle.

When the Data Errors in Dr. Neumark's Sample of 425 are Fixed, the Statistical Results Change a lot, and Adding a Control for Year Essentially Eliminates Gender Differences in Base Pay

Description	N	Female Coefficient Estimate	t-value
Dr. Neumark's Full Prior Pay Sample (As Reported in His Exhibit 3)	2,783	-0.0218	-2.02
Limited to "Apples to Apples" Subset with no Correction for Data Errors	422	-0.0491	-1.44
Correcting for Prior Base Pay Data Errors	422	-0.0230	-0.86
Add Year Control	422	-0.0090	-0.34

Exhibit 29

[76] Dr. Neumark adjusts all compensation measures based on the CPI. However, that adjustment underestimates the real salary growth for relevant occupations in the Silicon Valley geographical area. An examination of the Occupational Employment Statistics survey used by the State of California Employment Development Department shows that the median wage growth in Silicon Valley for Computer and Mathematical Occupations (SOC code 15-0000) during the class period far exceeded the CPI. (EDD State of California – OES Employment and Wages, https://www.labormarketinfo.edd.ca.gov/data/oes-employment-and-wages.html#OES).

Relevant experience and starting pay

93. Dr. Neumark's starting pay data includes a variable called APPLICANT_CURRENT_EMPLOYER that indicates the new hire's former company. There are a number of familiar companies in the technology sector (Cisco, Apple, Microsoft, IBM, HP, and so on), as well as less well known technology companies (Violin Systems, Happiest Minds, Hoopla Software, etc.) and some companies that are not in the tech sector (Gap Inc., Bob Evans Farms, Bay Area Video Coalition, among others). When Dr. Neumark refers to "relevant" experience in his analyses and in the information he "scraped" from resumes, what he means is just education level and years of experience since graduating from latest education.[77] He does not use the resumes to identify the quality or the nature of prior experience. A programmer coming from Apple quite likely had a very different experience than a programmer coming from Bob Evans Farms. His analyses do not take these factors into account.

There is wide variation in the relationship between prior and starting pay at Oracle

94. Implicit in Dr. Neumark's analysis is the idea that Oracle sets starting pay by asking about prior pay and then adding some fixed percent (for example, prior pay plus 5% or 10%).[78] If there is such a formula, then when I examine the relationship of starting pay to prior pay in the "apples to apples" population, I would expect to see more or less a narrow horizontal band across

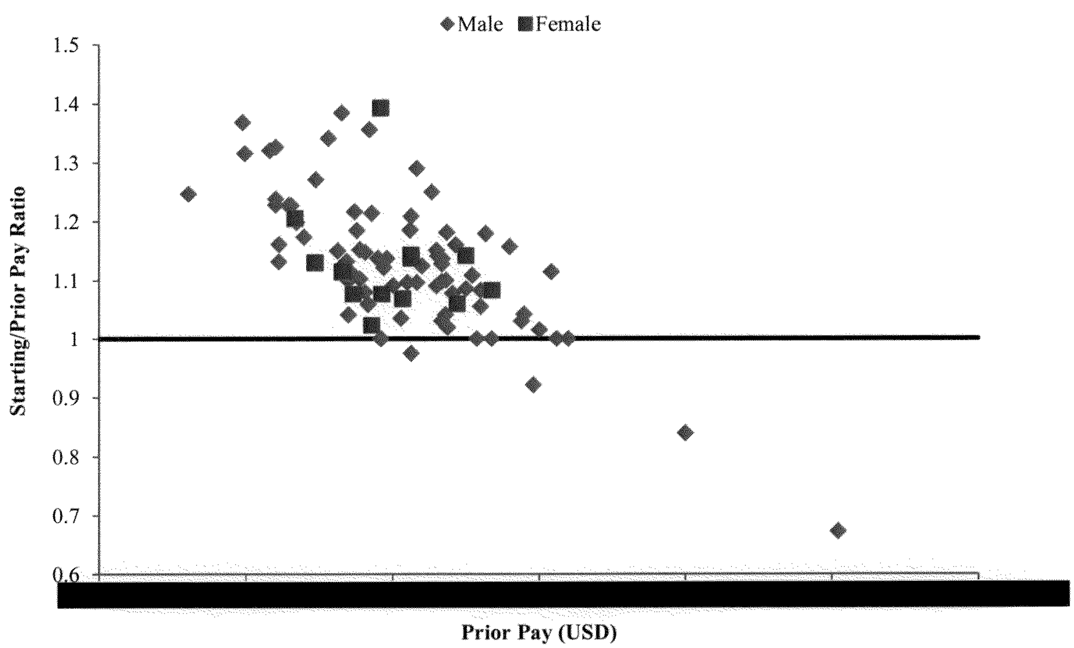
---

[77] In the population of 16,201 employee-years in Dr. Neumark's Exhibit 42, the correlation coefficient between scraped experience and age or experience (that he measured as age minus 22) is 0.84509. The correlation between his scraped experience and his scraped "relevant" experience is 0.9868. In effect, no new information or refinement is being added by his resume processing.

[78] For example, Srividhya Subramanian submitted a declaration that stated, "As part of my job duties as a Software Development Director and Software Development Senior Director, I was involved in hiring new employees and setting the initial pay at Oracle for these new employees. The primary factor I used for setting starting pay for new employees was prior salary. I was trained by my manager, Palanivelu Nagarajam who was an Oracle Vice President, to get starting pay within 10% of the applicant's then current pay." Declaration of Srividhya Subramanian in Support of Representative Plaintiffs' Motion for Class Certification, November 13, 2018.

the graph. The 5 graphs below depict the ratio of starting pay to prior pay for the 5 largest job titles in the starting pay data. The graphs demonstrate that there is no apparent formula: applicants' starting pay outcomes vary a great deal even when hired into the same job code. Also, each job shows new hires with a wide range of prior pay, and it does not appear to be the case that women's prior pay is below that of men. In short, contrary to Plaintiffs' suggestion there does not appear to be any lockstep process in which female new hires earned less at their jobs before coming to Oracle, which then resulted in lower starting pay.

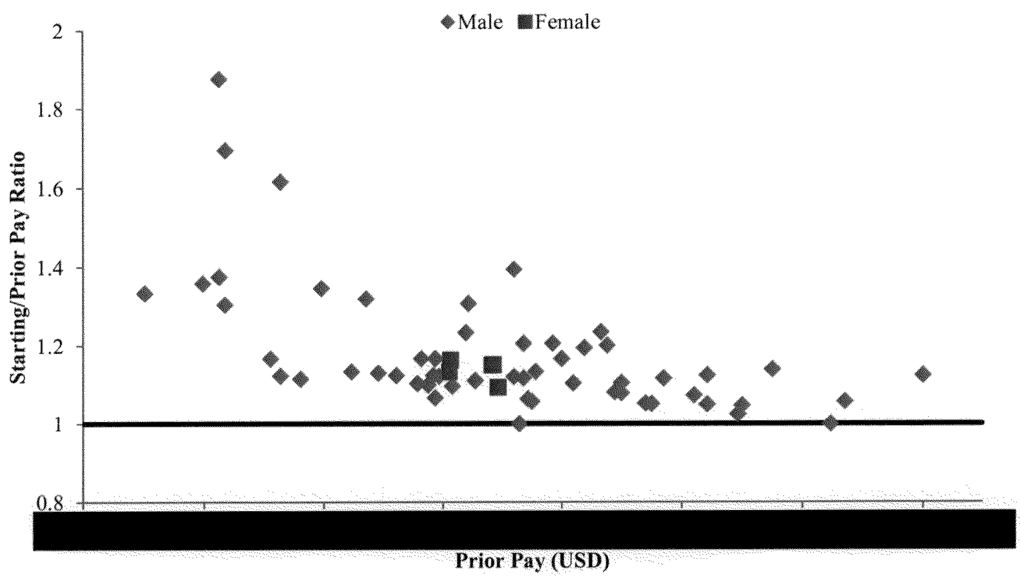
Software Developer 4: Ratio of Starting Pay to Prior Pay



Note: This population is part of Neumark's 425 "apples-to-apples" population

Exhibit 30

Software Developer 5: Ratio of Starting Pay to Prior Pay



Note: This population is part of Neumark's 425 "apples-to-apples" population

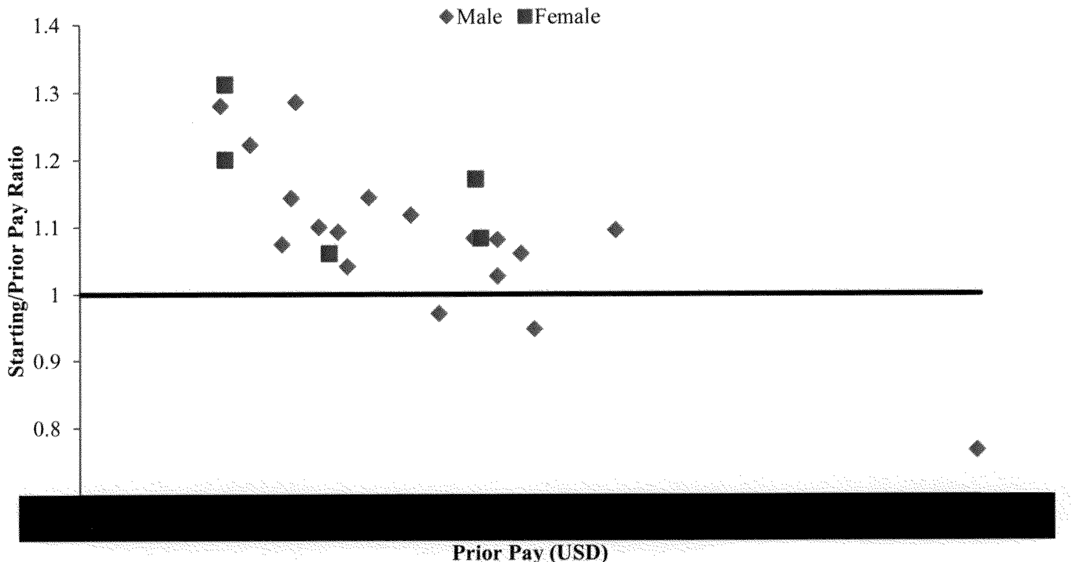
Exhibit 31

Software Developer 3: Ratio of Starting Pay to Prior Pay

Note: This population is part of Neumark's 425 "apples-to-apples" population

Exhibit 32

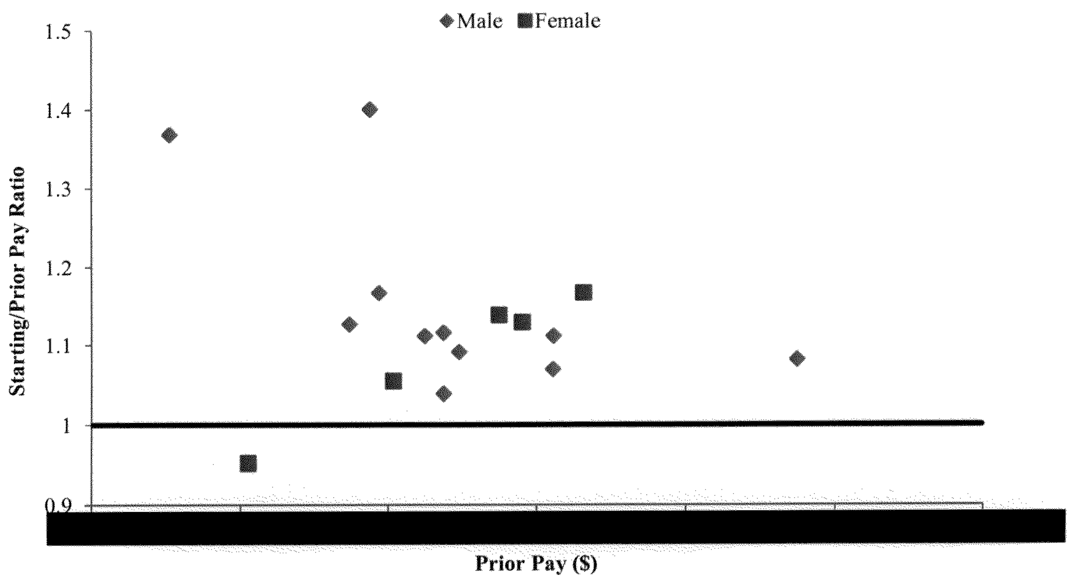
Product Manager/Strategy 5-ProdDev: Ratio of Starting Pay to Prior Pay



Note: This population is part of Neumark's 425 "apples-to-apples" population

Exhibit 33

Product Manager/Strategy 4-ProdDev: Ratio of Starting Pay to Prior Pay



Note: This population is part of Neumark's 425 "apples-to-apples" population.

Exhibit 34

95. Consider, too, that in addition to the prior pay entries for which it is not possible to decipher what is being recorded, the iRecruitment data from 2010 through 2017 (i.e., before the time period during which I understand questions about prior pay were banned by the California Equal Pay Act) includes 845 entries out of a total 7,040 entries (12%) which either record "0" or have no numeric value. These include 198 entries labeled not applicable and 6 entries indicating that the candidate did not provide the information.

NAMED PLAINTIFFS

96. Dr. Neumark performs a regression analysis in which he estimates pay shortfalls for the three Named Plaintiffs for five distinct measures of pay: Base Pay, Medicare Wages, Bonuses, Stock Grants, and Total Compensation.[79] His regression controls in these models are the same, but the variable for female codes all female employees other than the named plaintiffs, and he then adds indicator variables for each of the named plaintiffs. The results show that the coefficient for each named plaintiff is negative and statistically significant.[80] Dr. Neumark remarks that the results of these regressions indicate that the three proposed class representatives—Marilyn Clark, Manjari Kant, and Elizabeth Sue Petersen—received significantly lower compensation than comparable male employees (based on his definition of employees doing substantially similar work). [81]

[79] See Exhibit 4 of the Neumark Report.
[80] One exception is the model for Stock Grants, where Named Plaintiff Petersen's results show a surplus of stock grants given to her when compared to male employees, though this surplus is not statistically significant.
[81] Neumark Report, p. 6, paragraph 8e.

97. However, these results have no bearing on the class as a whole. To illustrate that point, I selected three additional women in the data and re-estimate the regression added an indicator variable for each, just as Dr. Neumark did for the named plaintiffs. The three women I selected are [redacted], and [redacted][82] The results show that the regression coefficients for the three women are positive and significant, meaning that [redacted], and [redacted] were all paid significantly *more* than men doing substantially similar work according to Dr. Neumark's model. Their estimated compensation premiums range from a 23% to 49.7% premium in base pay, and a 26% to 76.1% premium in Medicare wages. All of these premiums are highly statistically significant for these three women. Simply because Dr. Neumark estimates coefficients that indicates that the named plaintiffs were paid less than predicted by his model does not automatically make them representative of outcomes for all women, and his analysis certainly does not support the conclusion that all women in the class are underpaid when compared to men who he defines as performing substantially similar work.

[82] Their IDs are 889997603, 5018, and 888511678, respectively.

Adding Three Additional Female Employees to Dr. Neumark's Exhibit 4 Regarding Named Plaintiffs Indicates That There are Also Women Who are Paid More Than Dr. Neumark's Model Predicts

	Base Pay	Medicare Wages	Bonuses	Stock Grants	Total Compensation
	(1)	(2)	(3)	(4)	(5)
[redacted]	0.4974***	0.2940***	1.4046***	1.2709	0.3718***
	(0.0486)	(0.0742)	(0.2501)	(1.3239)	(0.0661)
Standard deviations	10.24	3.96	5.62	0.96	5.62
[redacted]	0.2388***	0.7611***	1.4883***	6.5968***	0.7699***
	(0.0193)	(0.0285)	(0.1479)	(0.3095)	(0.0268)
Standard deviations	12.35	26.73	10.06	21.31	28.67
[redacted]	0.3523***	0.2604***	0.3925**	8.0301***	0.1435***
	(0.0353)	(0.0448)	(0.158)	(0.3934)	(0.0433)
Standard deviations	9.98	5.81	2.48	20.41	3.31
Clark, Marilyn J.	-0.1317***	-0.2194***	2.2507***	-3.4035***	-0.1400***
	(0.0131)	(0.0189)	(0.1372)	(0.3292)	(0.0181)
Standard deviations	-10.08	-11.62	16.41	-10.34	-7.73
Kant, Manjari	-0.1458***	-0.2854***	-1.6584***	-3.9691***	-0.2442***
	(0.0155)	(0.033)	(0.1461)	(0.4176)	(0.0311)
Standard deviations	-9.42	-8.64	-11.35	-9.50	-7.84
Petersen, Elizabeth Sue	-0.2474***	-0.3552***	0.5364***	0.2020	-0.3292***
	(0.0082)	(0.0139)	(0.0785)	(0.1599)	(0.0124)
Standard deviations	-30.01	-25.49	6.83	1.26	-26.46
Observations	66,928	57,066	58,256	58,256	58,256

Note: This exhibit corresponds to Dr. Neumark's Exhibit 4.

Exhibit 35

<u>Named Plaintiffs' Comparators</u>

98. Dr. Neumark's regression model compares the Named Plaintiffs to employees in the same job code/grade, line of business head, and zip code, with similar tenure profiles. To look more closely at what this means, I flagged an individual in the data as a "comparator" if he or she his similar values to the Named Plaintiffs for the following characteristics from Dr. Neumark's

model in a particular year: 1) same part-time status, 2) same hourly status, 3) same line of business head, 4) same zip code, 5) same job code and job grade, and 6) three different measures of tenure within two years of the Named Plaintiffs' corresponding measures of tenure.[83]

99. The following charts show salaries for Ms. Clark's, Ms. Kant's and Ms. Petersen's comparators, respectively, for each year the individual Named Plaintiff is present in the data. In the following graphs, the individual Named Plaintiff is denoted by a red square, female comparators are denoted by a blue dot, and male comparators are denoted by a blue X.

Ms. Clark's Comparators:

100. Ms. Clark was employed during three years of the relevant period as a Database Administrator 4-IT. The charts below show her base pay, as well as that of her comparators for each year from 2013 through 2015.

[83] The three measures of tenure used are: tenure in job, which is the length of time spent in the current job code; Oracle tenure, which is the length of time spent at Oracle; and overall tenure, which is the length of time spent at Oracle and any company acquired by Oracle. Note that I did not also subset comparators within two years of his general experience variable (age minus 22). This is because Elizabeth Petersen would then have no comparators in any year, Marilyn Clark would have just one comparator in each year for 2013-2015, and Manjari Kant would have seven comparators in 2013, just one comparator in 2014 and 2016, two comparators in 2015, and none in 2017. Of the tenure measures adopted by Dr. Neumark, age is the least relevant, and thus that is the one I dropped. However, when I do keep the age measure to define comparators to the Named Plaintiffs, there is no pattern that emerges among the handful of comparators that remain.

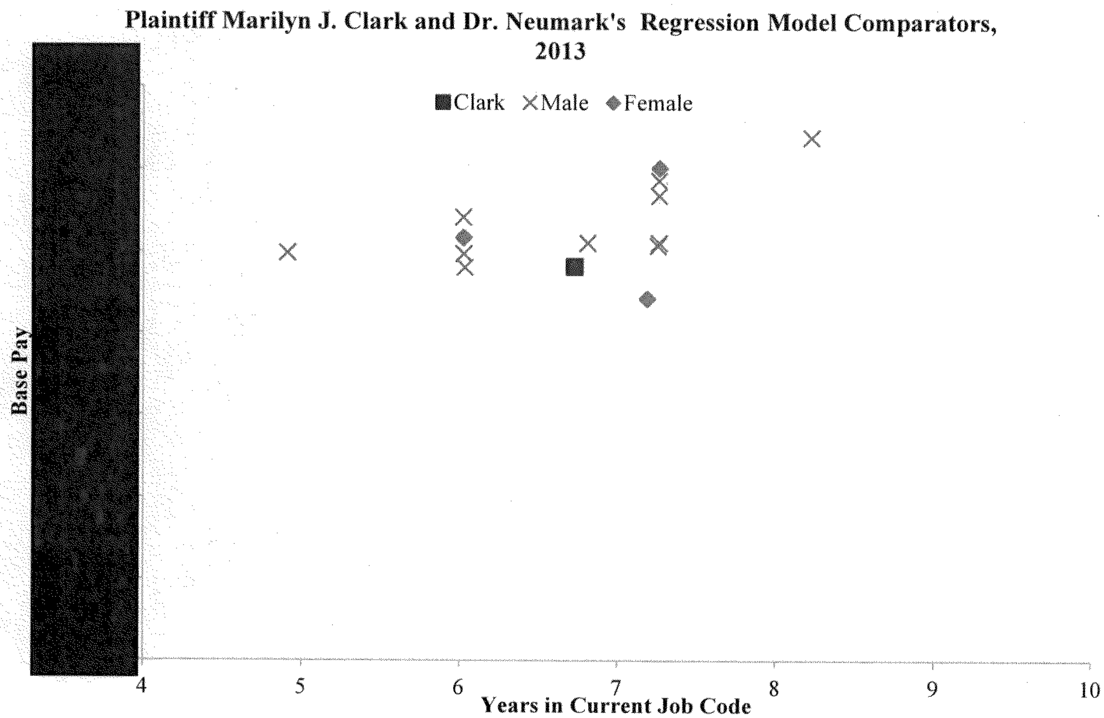


Exhibit 36

101. In 2013, Ms. Clark had 13 comparators selected based on the criteria described above. Ms. Clark's salary was $96,315 in that year; her comparators were paid salaries between [redacted] and [redacted] That is a salary range of almost [redacted], which is very wide especially when considering that comparators have the nearly all same job attributes as those controlled for by Dr. Neumark in his regressions. As can be seen in the chart, there were comparators in 2013, both m ale and female, that were paid less than Ms. Clark, as well as male and female comparators that w ere paid more than Ms. Clark. The lowest and highest salaries for a female comparator in 2013 were [redacted], respectively.

Plaintiff Marilyn J. Clark and Dr. Neumark's Regression Model Comparators, 2014

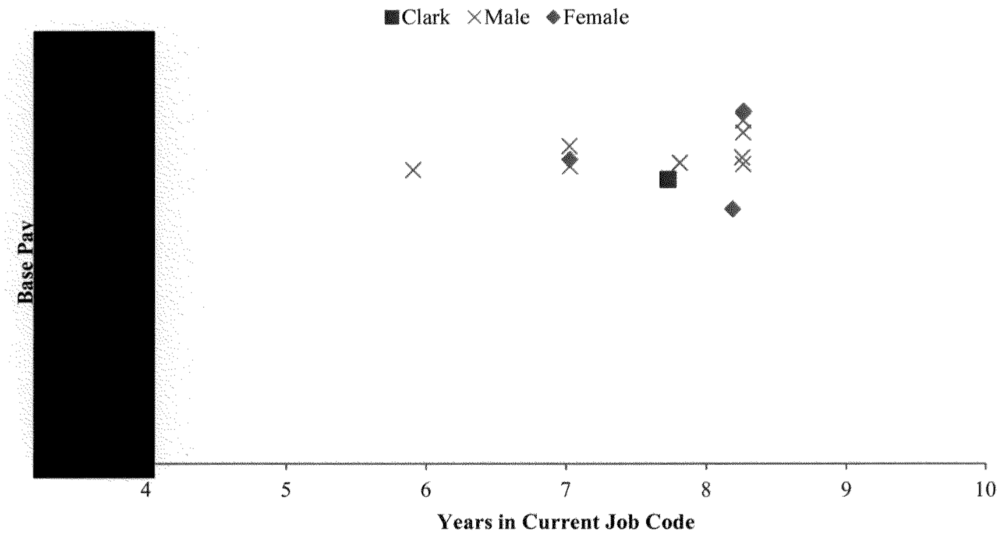


Exhibit 37

Plaintiff Marilyn J. Clark and Dr. Neumark's Regression Model Comparators, 2015

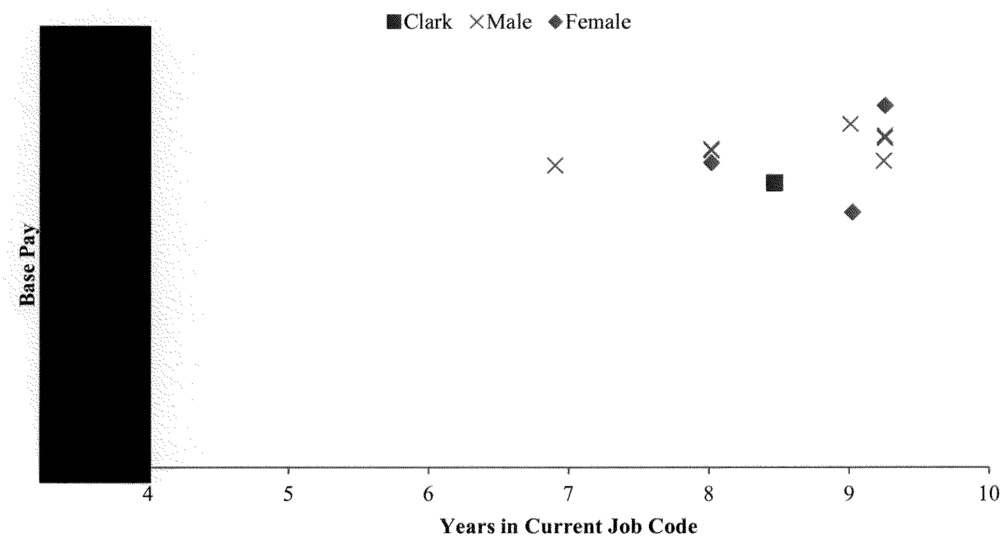


Exhibit 38

102. In 2014 and 2015 the range of salaries for Ms. Clark's comparators is also very wide, approximately [redacted] in 2014 (with the lowest salary at [redacted] and the highest at [redacted] and [redacted] in 2015 (with the lowest salary at [redacted] and the highest at [redacted]). Note that in both 2014 and 2015, the highest salary paid to Ms. Clark's comparator group was paid to a female comparator.

103. Ms. Clark identified four male colleagues whom she considered to be her comparators. Three of the four identified comparators are present in the comparator group selected based on Dr. Neumark's regression criteria.[84] All three are paid more than she was.

Ms. Kant's Comparators:

104. Ms. Kant was employed during five years of the relevant period. The next set of charts show her salary, as well as that of her comparators for each year from 2013 through 2017. In 2013, Ms. Kant held a QA Analyst 4-ProdDev job title. In that year, using the criteria described above, 18 comparators were selected for Ms. Kant. Ms. Kant was paid $89,896 in base salary in 2013; her comparators were paid salaries between [redacted]. The highest salary is over [redacted] higher than the lowest salary. There were comparators, both male and female, that were paid more than Ms. Kant in 2013, and the highest salary for a female comparator was [redacted]

[84] The three comparators are: 1) Michael Burrows in 2013 and 2014; 2) Tuan Karsevar in 2013, 2014, and 2015; and 3) Alejandro Espinosa in 2015.

Plaintiff Manjari Kant and Dr. Neumark's Regression Model Comparators, 2013

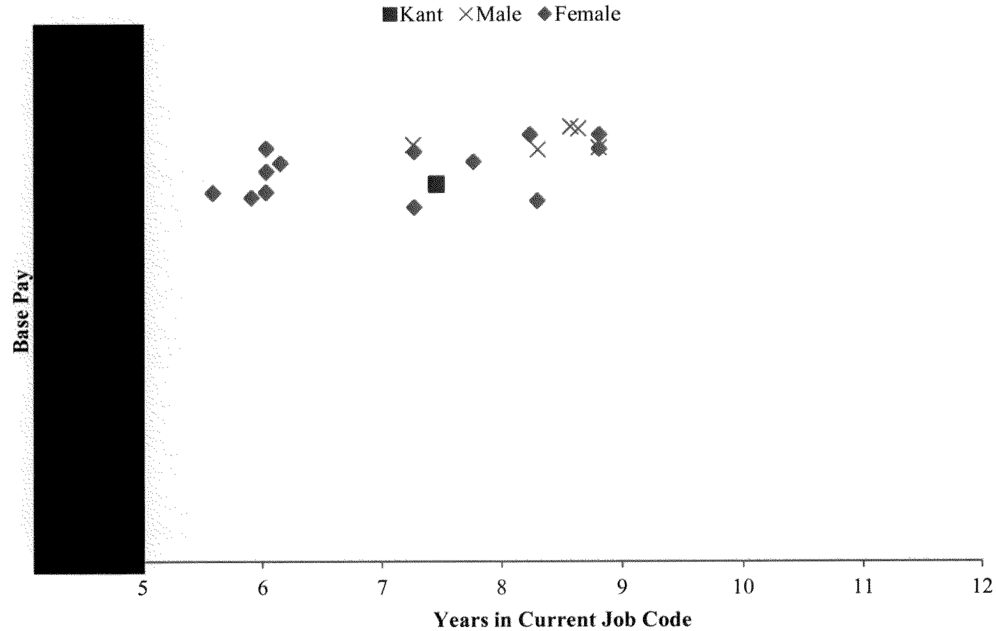


Exhibit 39

105. In 2014, Ms. Kant became a QA Analyst 5-ProdDev, and as a result the number of her comparators selected based on the criteria discussed dropped. In 2014, there were two other comparators to Ms. Kant, and both of them were female. While Ms. Kant's salary was $94,396 in that year, one of her comparators had a salary lower than hers ([redacted]) and one had a salary that was higher than hers ([redacted]). Similarly, in 2015, 2016, and 2017, Ms. Kant had a smaller number of comparators, all female, with salaries sometimes lower and sometimes higher than her own.

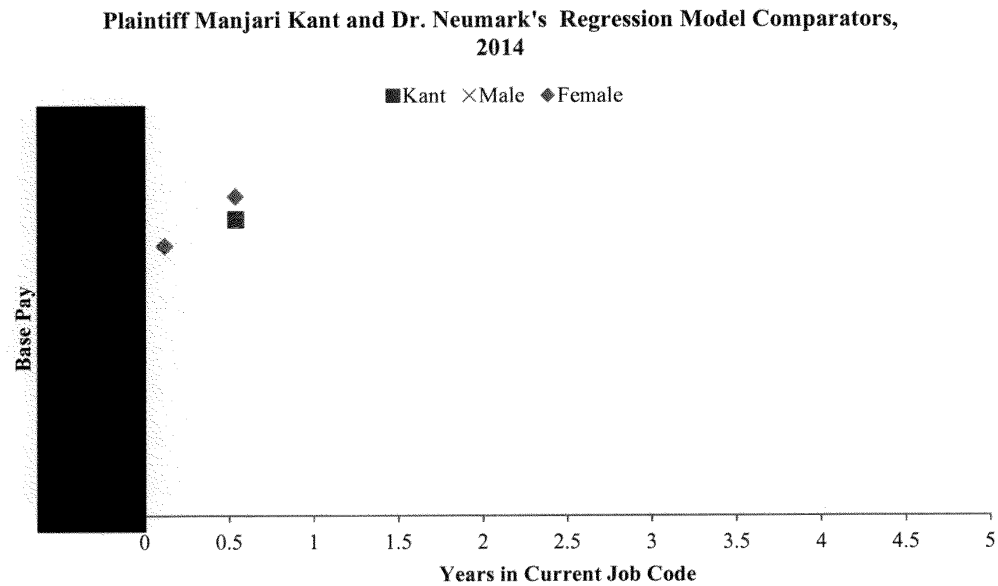


Exhibit 40

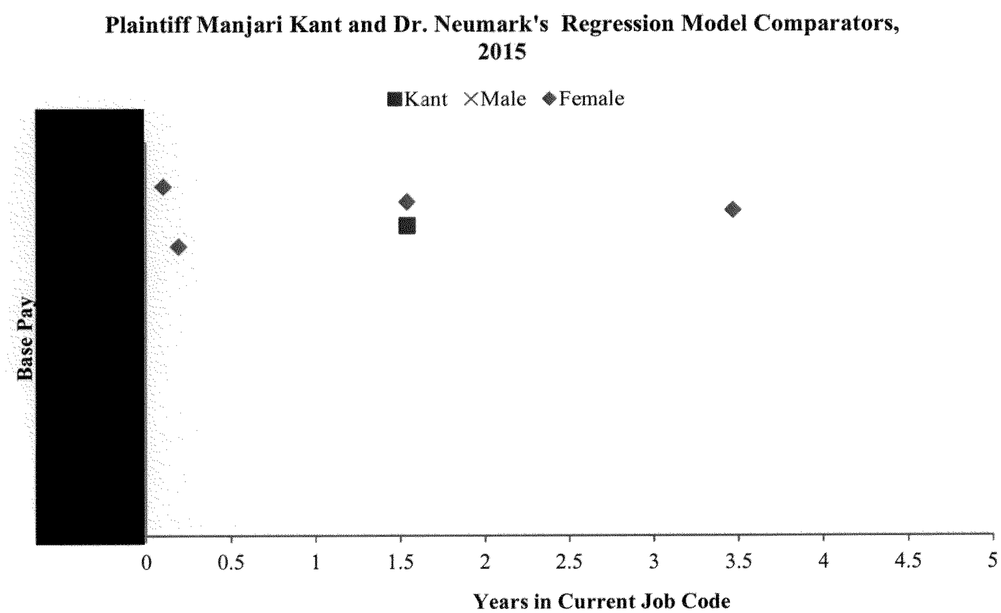


Exhibit 41

Plaintiff Manjari Kant and Dr. Neumark's Regression Model Comparators, 2016

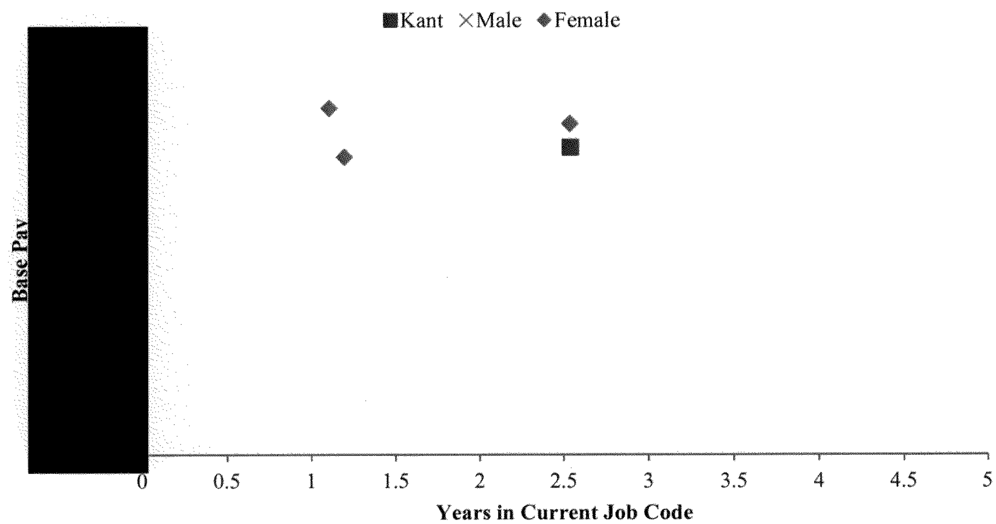


Exhibit 42

Plaintiff Manjari Kant and Dr. Neumark's Regression Model Comparators, 2017

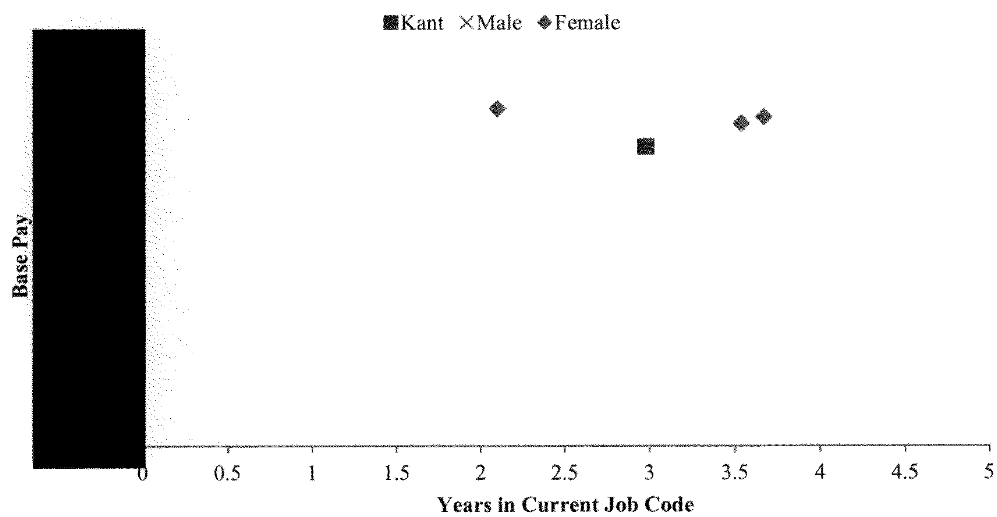


Exhibit 43

106. Ms. Kant identified three male colleagues whom she considered to be her comparators.[85] None of them appear in the group of comparators selected based on Dr. Neumark's regression criteria.

Ms. Petersen's Comparators:

107. Ms. Petersen was employed during six years of the relevant period. The next charts show her salary, as well as that of her comparators for each year from 2013 through 2018. Evaluated against her comparators, Ms. Petersen was paid a lower salary than all of them in every year ($78,000 in 2013, $80,000 in 2014 and 2015, and $81,600 in 2016, 2017, and 2018). She had three comparators in 2013, two women and one man, who were all paid more than she was, with the male comparator being paid the salary closest to hers [redacted] and the two female comparators paid approximately [redacted] than she was paid (both female comparators were paid [redacted]).

[85] Ms. Kant's three comparators are: 1) "Ed," 2) "Ram," 3) Raymond Winther.

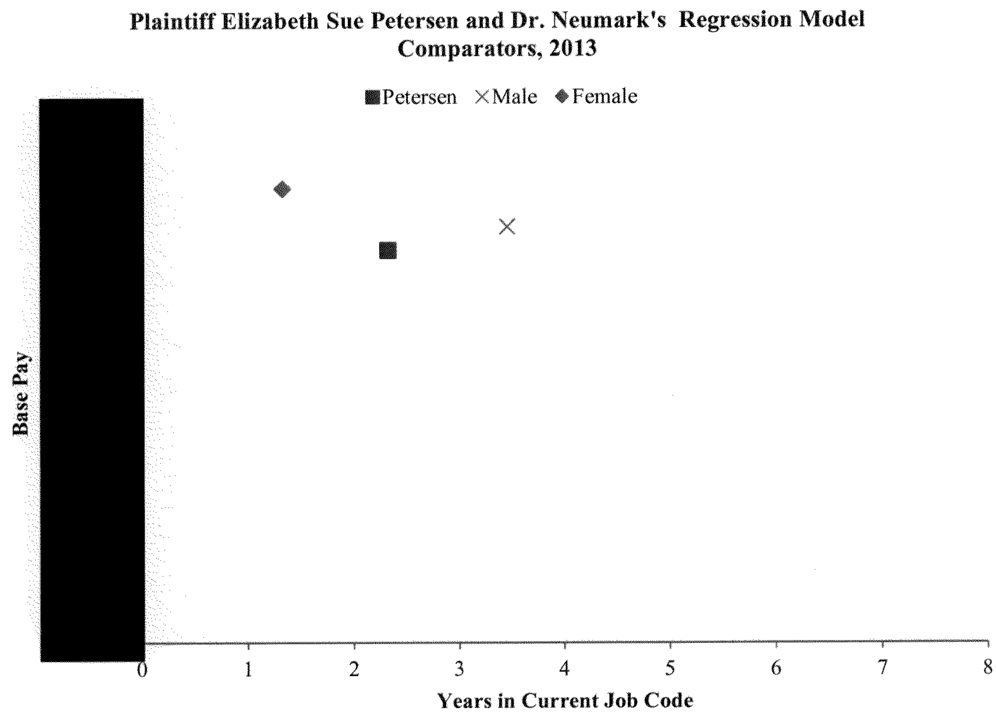


Exhibit 44

108. From 2014 through 2018, only one person was a comparator to Ms. Petersen, based on the criteria discussed above. This comparator (who was a woman) was paid considerably more than Ms. Petersen was paid in all of the years.

Plaintiff Elizabeth Sue Petersen and Dr. Neumark's Regression Model Comparators, 2014

Exhibit 45

Plaintiff Elizabeth Sue Petersen and Dr. Neumark's Regression Model Comparators, 2015

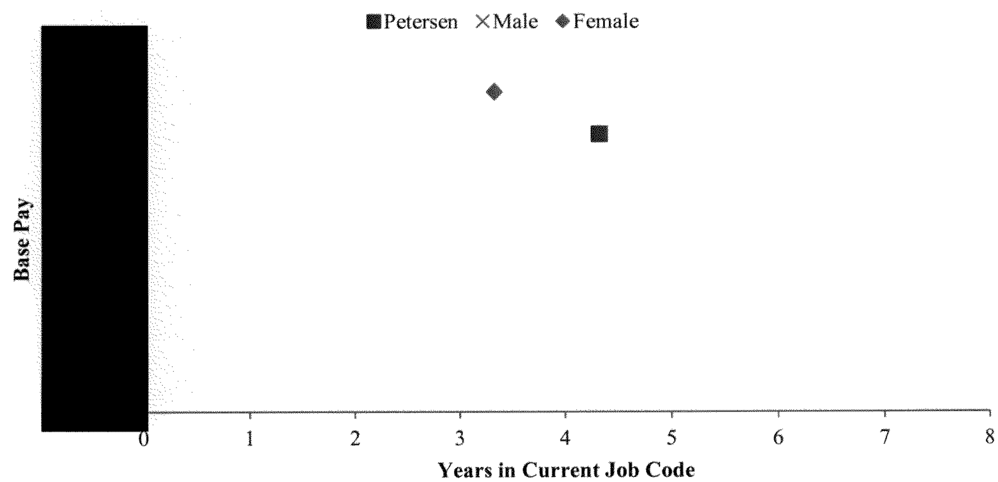


Exhibit 46

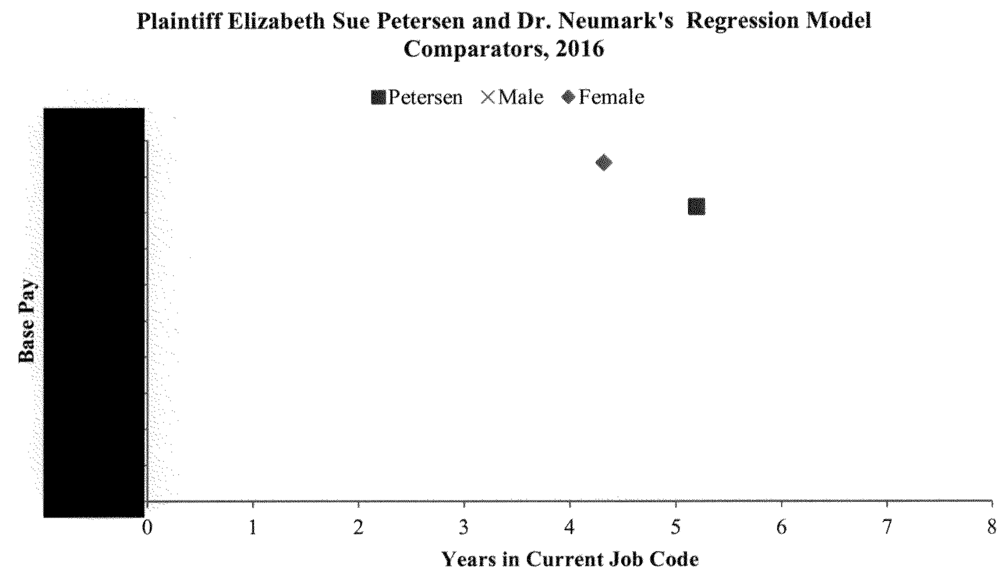


Exhibit 47

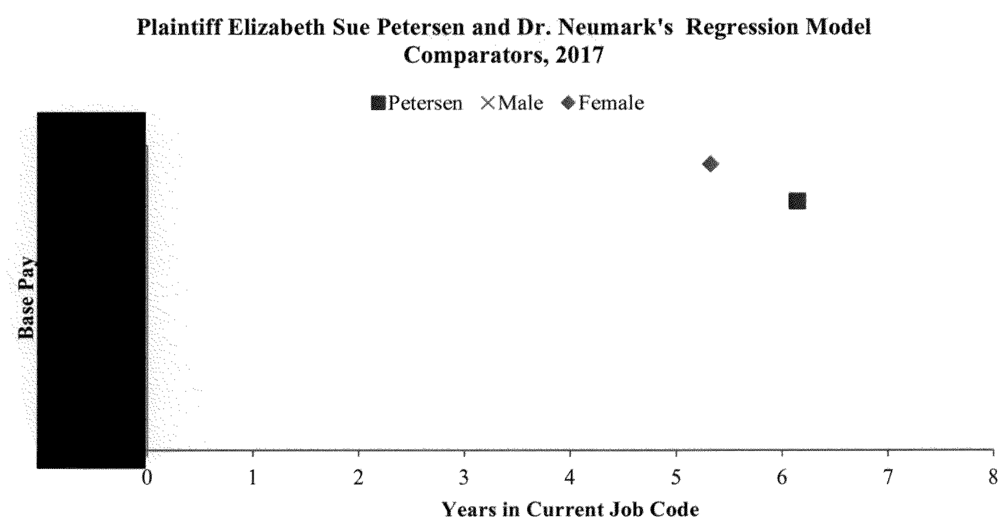


Exhibit 48

Plaintiff Elizabeth Sue Petersen and Dr. Neumark's Regression Model Comparators, 2018

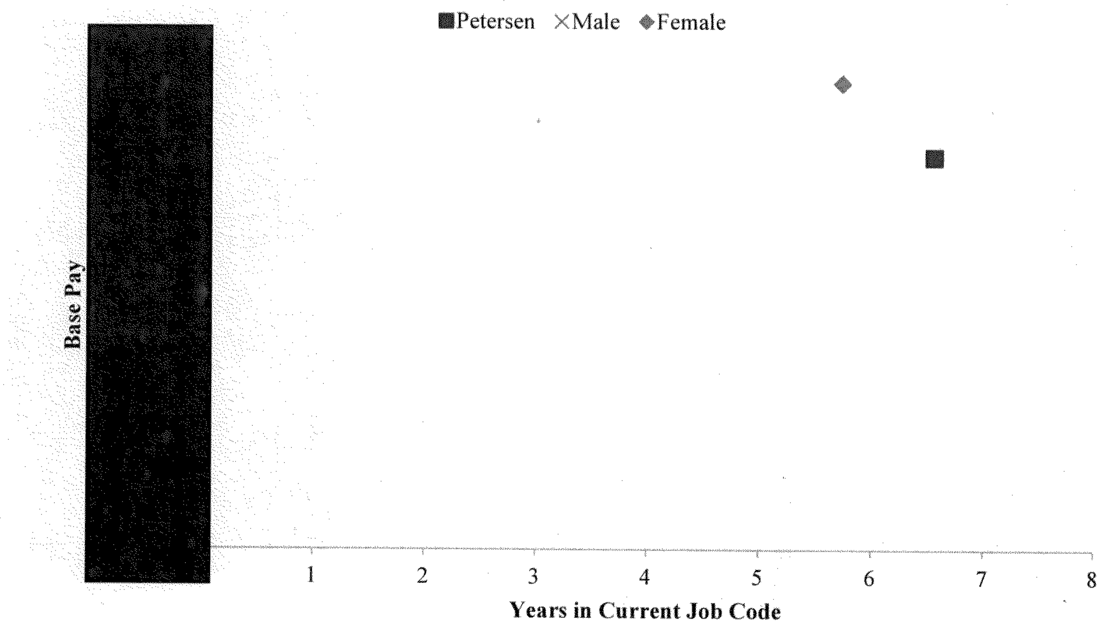


Exhibit 49

109. Ms. Petersen identified three male colleagues whom she considered to be her comparators.[86] None of them appear in the group of comparators selected based on Dr. Neumark's regression criteria.

[86] The comparators she selected are David Manes, Owen Richards, and Victor Cecena.

ADDITIONAL CRITIQUES OF DR. NEUMARK'S PAY ANALYSES

110. I have reviewed Dr. Neumark's computer programs and I have many concerns about how he constructed his dataset, as well as concerns about the variables he selected for his analysis.

Dr. Neumark does not properly define the relevant population

111. Dr. Neumark includes 1,979 non-California observations – that is, years when the employee at issue did not work in California – in his analyses and in his damages estimates. My understanding is that Oracle produced complete records for all employees who were ever employed in Product Development, Information Technology, and Support in California during the class period. This means that if, for example, in one year the individual was in California but in the next year they moved to Washington, their employment history for the time that they were in Washington was included in Dr. Neumark's analysis. Dr. Neumark does not limit his analysis file to employment spells in California, as he should have done.

112. The second problem with Dr. Neumark's analysis is that he does not analyze the correct population. His programs incorrectly identify college hires. He excluded 777 college hires from 2010 through the end of the data based on a variable called "Change_Reason" reflecting a value of "Campus Hire."[87] In his report, he notes that CR stands for College Recruiting,[88] but he did not exclude observations for which the "Change_Reason" variable was coded as "CR- Dev Hire." In the notes to Exhibit 4, he states that Xian Wang is not coded as a campus hire. Had he examined the variable "Justif_for_this_hire," which is in the same data file as "Change_Reason," he would have seen that her data shows "Justif_for_this_hire" reflecting a value of "College Hire." In another data file, the variable HIRE_TYPE also includes the codes "CR – Dev Hire"

[87] This variable appears in ORACLE_JEWETT_00001180.
[88] Neumark Report, p. 26, paragraph 65: "Other types, such as CR (college recruiting) do not have any information about prior pay [...]."

and "Campus Hire."[89] Had Dr. Neumark also used those variables to identify college hires, he would have excluded another 1,313 observations from his data. Moreover, it is not possible to identify college hires hired prior to 2010, meaning that they remain in the data even though they are not part of the proposed class as defined. The erroneous inclusion of these college hires means that his results do not provide reliable estimates for the proposed class as defined by Plaintiffs.

<u>Dr. Neumark's failure to control for all leaves of absence taken by various Oracle employees overestimates tenure, especially for women</u>

113. A third problem with Dr. Neumark's data construction is that he does not account for leaves of absence that occur prior to the current position, which he acknowledges in his testimony should be done.[90] This is surprising because incorrect tenure measures for women are a well-documented problem in the economics literature, as he testified and as he noted he has done in his other work[91] and it is a significant data construction problem.[92]

114. Dr. Neumark calculates tenure as the difference between the Oracle hire date and the date of the salary observation. He experiments with using "continuous service" hire date, which includes any time spent in other Oracle entities (like Oracle India) or at a firm that was acquired. He does not, however, subtract time on leave from these tenure variables. The only way he accounts for time not worked is in a person's current job code – if they happened to take their

---

[89] ORACLE_JEWETT_00007304

[90] Neumark Deposition, 260:1-262-1.

[91] Neumark Deposition, 260:17-20. "If it's true that there's actually a meaningful positive effect of a company tenure question that doesn't account for leave, then yes, I would expect the gap to be reduced if I – if I measure it." Neumark Deposition, 261:15-19.

[92] Mincer, Jacob, and Solomon Polachek (1978). An Exchange: The Theory of Human Capital and the Earnings of Women: Women's Earnings Reexamined. *The Journal of Human Resources* 13(1), pp.118-134. Bertrand, Marianne, Claudia Goldin, and Lawrence F. Katz (2010). Dynamics of the Gender Gap for Young Professionals in the Financial and Corporate Sectors. *American Economic Journal: Applied Economics,* 2(3), pp. 228-55.

leave of absence while in their current job code, he subtracts those days from their tenure. If they took leave at any prior time, he does not. The reason mishandling leaves of absence is especially problematic when estimating gender gaps in pay is because, on average, women take more leave than men at Oracle, as the data that was available to Dr. Neumark shows. By not accounting for tenure properly, his estimated gender coefficient is thereby biased.

115. The graph below shows average leave taken by gender, depending on whether someone was hired directly by Oracle or came through one of the 72 acquisitions reflected in the data. It shows that women hired directly into Oracle averaged 79.2 days of leave at Oracle and women who came through acquisitions averaged 56.9 days of leave at Oracle; for men, these numbers were 7.8 days of leave and 8.7 days of leave, respectively. For acquisitions, the available data indicates their original hire date at the acquired firm but does not record leave histories; thus, the record is incomplete for both men and women who were at acquired firms. That Dr. Neumark's model does not take cumulative leave into account is clearly problematic, since leave is not evenly distributed between men and women in this population of Oracle employees and thus failing to control for leave biases the results he obtains for gender.

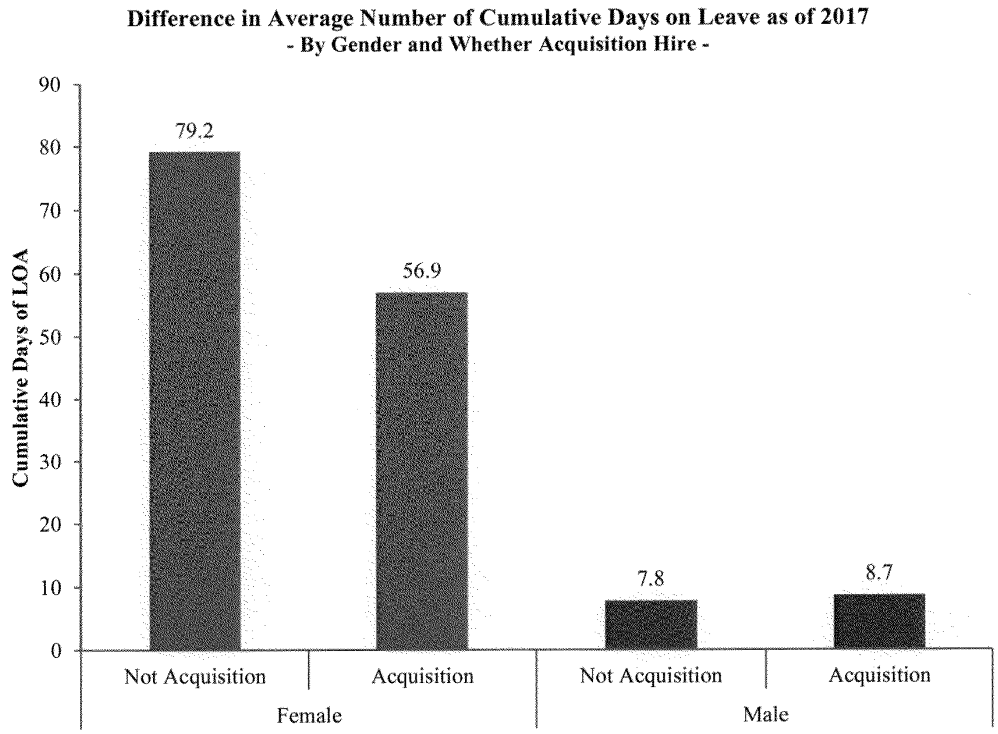


Exhibit 50

Dr. Neumark uses zip code to group employees rather than establishment

116. A fourth data construction issue is that Dr. Neumark uses zip code as his location variable rather than the establishment. My understanding is that prior to 2016, California Equal Pay Act comparisons were to be made within an "establishment." In any given year, about 6% of employees work from home in California, and in Dr. Neumark's analysis, their home zip codes will be used to identify the work location. They nonetheless work on teams based at Oracle locations that may or may not be located anywhere near this zip code. Comparing these work from home employees to other individuals who also happen to work from home and live in the same zip code as them, or who work at physical establishments that happen to be nearby – regardless of the managers to whom they report or teams they work on – leads to inappropriate

comparisons. Dr. Neumark thus does not properly account for location, and its potential impact of pay scales and compensation rates, in his models.

Dr. Neumark overestimates total compensation for part year employees

117. Dr. Neumark builds his total compensation measure by adding together regular earnings, stock awards and bonus amounts.[93] He then "grosses up" that total for employees who only worked part of the year by dividing total compensation by the proportion of the year that they worked. For example, for someone who only worked half the year, Dr. Neumark would double their total compensation. This method assumes that not only were regular earnings half as much as they would have been had the individual worked for the entire year, but so also were stock awards and bonuses. This assumption is problematic.

118. To see why, imagine that Oracle acquires a company in the month of July, and brings on the founder of that company with a base salary plus $2 million dollars of stock. Dr. Neumark's extrapolation method would assume that the founder would have received $4 million in stock had the company been acquired in January instead of July. Similarly, if Oracle hires someone and needs to grant them stock in order to "make them whole" (for example, because they walked away from an impending stock award with their previous employer), that one-time payment would also be scaled up using Dr. Neumark's method. This is true as employees exit the firm as well. [redacted] (PERSON ID 85058), for example, worked only 3% of 2016 because he left Oracle on January 11, after receiving what was (for him) an unusually large stock award. Dr.

[93] He uses the file ORACLE_JEWETT_00001167 for "Regular Earnings." From the same file, he estimates total stock amount using the variable balance_name to identify "NonQual Stock Opt," and "Restricted Stock Units." He estimates bonuses by cumulating anything that has the word "bonus" in it from that file and combines it with the amounts in the bonus tab in ORACLE_JEWETT_00030955_Jewett_gsi_comp_history_native (excluding severance pay). After he adds regular earnings, stock and bonus amounts to calculate total compensation, he then scales up the total compensation for employees who worked part year.

Neumark takes his total compensation and scales it up based on [redacted] working 3% of the year. As a result, Dr. Neumark's data codes [redacted] as earning over [redacted] in 2016.

Dr. Neumark's Adjustment for Part Year Work Creates Serious Measurement Error Problems in Total Compensation

- Person ID: 85068, Software Development VP -

Year	Base Salary at End of Year	Regular Earnings	Total Stock Amount	Total Bonus Amount	Total Compensation	Percentage of Year Worked	Total Compensation Adjusted for % of Year Worked	CPI Adjustment to 2017 Dollars	Total Compensation Adjusted to 2017 Dollars
2013	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	100.00%	[redacted]	1.0578	[redacted]
2014	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	100.00%	[redacted]	1.0499	[redacted]
2015	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	100.00%	[redacted]	1.0423	[redacted]
2016	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	3.01%	[redacted]	1.0211	[redacted]

Exhibit 51

119. This decision to extrapolate compensation for part year workers is also a problem for defining annual equity compensation. Restricted stock units ("RSUs") vest once a year, typically at the same time each year. Stock options would unlikely be exercised more than once per year. Consider an employee who exercised stock options on January 2 of a given year. Whether the employee leaves the company on January 3 or on December 1, it is inappropriate to extrapolate equity income events in the same manner that one would extrapolate base salary, as equity income occurs as infrequent, discrete events, whereas base salary is typically earned in equal increments throughout the year. In other words, if an employee worked at Oracle for half a year, and exercised $10,000 in taxable stock options that year, Dr. Neumark would gross up the equity income to $20,000. In this example, the extrapolation of the same stock award would look very different if the employee left in January as opposed to in December.

Dr. Neumark incorrectly values stock awards

120. More generally, Dr. Neumark does not calculate equity compensation correctly because he inappropriately combines taxable equity income with actual base and bonus earnings within the same calendar year. Exhibit 10 of his report is titled "Adjusted Total Stock Grants (RSUs + Options) (2013-2017)."[94] However, this exhibit represents the actual taxable equity income in each calendar year, which is not the same as the year in which the equity awards were granted to (and thus earned by) the employee. This distinction is important for several reasons.

121. Oracle grants two different types of equity awards: stock options ("SOs") and restricted stock units. It is my understanding that employees have at times had the choice of receiving SOs or RSUs.[95] There are basic differences between these two types of awards. A stock option gives an employee the right to purchase a number of company shares by a future date (the "expiration date," typically ten years) for a specific amount per share (the "strike price"). Typically, there is also a vesting schedule, prior to which an SO cannot be exercised.[96] If the price of the company's stock is higher than the strike price at a future date, the employee can exercise his SOs and profit from the difference. This profit would be reported as taxable income in the year in which the options are exercised. If the price of the company's stock falls and the SOs are about to expire, the SOs become worthless, and the employee would not exercise them. Alternatively, an employee might not exercise his options for other reasons, such as when he or she leaves the company before their SOs vest, or if he or she chose simply not to pay the strike price required to exercise the SOs.

122. An RSU is different in that at the time of the grant, the company awards a specific number of company shares that will be distributed to the employee in the future, pending certain

[94] Neumark Report, Exhibit 10.
[95] See, for example, Equity Choice FAQ (WANG_00001).
[96] Ibid.

vesting requirements. That is, the employee only receives the awarded shares once they are fully vested. At Oracle, the typical vesting schedule is four years, with one-fourth of the RSUs being vested each year.[97] Thus, even if the share price falls, the vested RSUs still have value. Again, only the fraction of RSUs that vest in a given year are typically reported as taxable income in that same year.

123. By using "Medicare wages" and "total compensation" in his analyses, Dr. Neumark combines the actual base and bonus pay an employee *earned* in a given year with *exercised* SOs and *vested* RSUs in the same calendar year. He is therefore combining earnings that were earned over different time periods within the same annual measure of income.

124. Consider this example: Employee A is granted stock options in 2004, which he exercises ten years later in 2014 and which are worth $10,000 on the exercise date. Employee B is awarded an RSU grant in 2013, of which one-fourth vests in 2014, with the vested amount also worth $10,000. The "Medicare wages" and "total compensation" analysis would show this $10,000 income as having been earned in the same year by each employee, whereas in reality, it was *earned* nine years apart (2004 versus 2013). What Dr. Neumark calls "stock grants," in reality is "taxable equity income."

125. Consider another example: Employees X and Y each receive an award of 100 stock options in 2007. Employee X exercises his options in 2013 and earns $5,000 of taxable equity income in that year. Employee Y does not exercise his options and they expire. Dr. Neumark's analyses would treat Employee X as having been granted (or earned) $5,000 worth of stock options in 2013, and Employee Y was granted (or earned) zero stock options in any year (stemming from the original grant). The reality is, however, that both Employee X and Y *earned* an equal amount in equity compensation in 2007. This would not be reflected at all in Dr.

[97] Ibid.

Neumark's analyses. Further, by using data that combines equity income earned over different years with actual base and bonus earnings for possibly a different year, Dr. Neumark's analyses using Medicare wages and total compensation cannot possibly provide an actual portrayal of total *earnings* in any given year.

126. It is clear that Dr. Neumark does not fully appreciate how equity compensation works at Oracle. When asked at his deposition, "I believe, you evaluated two different kinds of equity compensation at Oracle, correct?" he responds, "Everything I do, I believe, aggregates them. What I was saying was that I'm somewhat less sure about how the—the non-RSUs should be reflected in pay, which I think is – that's one of the reasons I'm – I – makes [sic] a lot of tables, but I keep doing multiple pay measures, including pay measures that don't include stocks, and including pay measures that don't include stocks for the people who never get stocks, for whom this simply can't be an issue."[98] Further, when asked, "Do you know whether employees at Oracle, who were awarded equity, could choose whether to receive that equity as stock options versus RSUs?" Dr. Neumark responds, "I don't know that."[99] Interestingly, when given a hypothetical example asking him how he would treat a grant in 2014 that would vest over four years, he responds, "So the RSUs, I treat as the – as exact – as – as the vesting. And there's not a – I don't think there's a good answer to this. I would say two things: One is, from the point of what the government considers income, that is – that is income, so that's why – that's why it's – that's one reason to do it that way. The second thing I would say is, you know, it's not at all obvious to me why there would be a gender difference in any of this stuff, but who knows."[100] He also agrees that the year in which a stock grant is considered a taxable Medicare wage may

[98] Neumark Deposition, 310:3-12.
[99] Ibid, 310:13-16.
[100] Ibid, 315:16-25.

not be the same as the year in which it was earned.[101] Dr. Neumark even concedes that "the taxable events from the non-RSUs may depend, in part, on decisions employees make. And in that sense, those – that variation should not be viewed directly as a measure of compensation."[102] It is clear from his testimony that Dr. Neumark is aware of some of the potential issues in the way in which he includes equity income in his analyses, yet he nevertheless ignores them.

Dr. Neumark's line of business head variable is too high in the company to demarcate groups of employees doing substantially similar work

127. [redacted] of the 68,510 observations in Dr. Neumark's data relate to work in a single line of business that was led by Thomas Kurian. As the earlier analysis showed, even when restricted to a single job title, employees in Mr. Kurian's line of business work on very different products and accordingly, their starting pay exhibits variation.[103]

128. For example, below I compare two employees hired in 2016. Both were hired as Product Managers/Strategy in IC2 in Mr. Kurian's line of business. They have similar tenures and similar ages. This is what Dr. Neumark's model would control for. Discretionary job titles are not fully populated in the data, giving them limited use for comparing all employees, but both employees discussed here have a discretionary title that provides more detail on the kind of work they do. Employee 1 is in business operations while Employee 2 has a creative function,

[101] Ibid, 313:19-25.
[102] Ibid, 309:12-16.
[103] It also does not appear to be the case that the heads of lines of business are making day to day pay decisions even if their names appear in the list of potential approvers, as Ms. Waggoner testified. (See Waggoner Deposition, 191:1-6: "Q. And it has to get approved all the way back up to the CEO level? A. Individual recommendations aren't really reviewed and approved. Again, it's about, did they stay in the budget.") I examined the 2013-2017 audit data for focal pay decisions. Of the 7 heads of lines of business Dr. Neumark controls for in his model, 6 do not appear as ever having changed a pay decision. The exception is Mr. Screven, who is recorded as having made pay decisions for 9 high level employees in the relevant population, all M6 or higher, in his chain of command. This is not relevant to lower level employees like Software Developer 3s or 4s, among others.

according to the discretionary titles. Their educational backgrounds differ not so much in quantity (both have a B.A.) but in the details – Employee 1 has a degree in Government and Economics from Harvard and Employee 2 has a more creative background, with a degree in Broadcast and Electronic Communication Arts from San Francisco State University and prior experience as a video producer. Yet Dr. Neumark treats them as equivalent in his model.

	Employee 1 **Person ID: 894031109**	**Employee 2** **Person ID: 894048892**
Hire Date	03/14/2016	02/29/2016
Age	24.8	25.3
Line of Business Head	Kurian	Kurian
Job Title	Product Manager/Strategy 2-ProdDev	Product Manager/Strategy 2-ProdDev
Discretionary Job Title	Senior Associate, Business Operations & Strategy	Product Manager, Studio Specialist
Starting Salary	[redacted]	
Full-Time Status	Full Time – Regular	Full Time – Regular
Organization	BG16 - Public Cloud Platform Development - ORCL USA	PL07 - Fusion Development Management - ORCL USA
Education	Bachelor of Arts, Government & Economics from Harvard University - HBX Credential of Readiness from Harvard Business School	Bachelor of Arts, Broadcast and Electronic Communication Arts from San Francisco State University
Previous Experience	Participated in a rotational program at LinkedIn as Business Leadership Program Associate after graduation - Internship at a start-up company generating clients and creating business plans	1+ years of other company experience at Tribune Media - KTXL FOX40 as a Creative Services Producer/Editor - 2+ years of experience as freelancer/video producer

Exhibit 52

Dr. Neumark's results show no gender difference in performance ratings

129. Dr. Neumark claims that performance ratings are a "directly observable measure of productivity."[104] But he concedes that this is not based on any objective evidence, and is only his "expert, as well as common sense understanding, of what a performance review is supposed to be." He continues, "I've had performance reviews myself, and they're meant to be rel – tell you something about my productivity."[105] With regard to including performance scores in his analysis, Dr. Neumark acknowledges that he has "some reservations about doing this"[106] and lists numerous reasons why he believes that Oracle performance data are inadequate. In particular, he expresses the hypothetical concern that performance scores can, in some instances, be susceptible to bias. Yet in his models of Oracle, where ratings at Oracle are taken into account, the gender coefficient on pay does not change. In other words, there is no meaningful gender gap in performance measures because if there were, it would have affected the gender gap in pay in his models. Moreover, if there are different kinds of work being done within a job code, then to the extent that men and women perform different kinds of work, even if they will on average receive the same performance ratings they may earn differential pay because the work they are doing is different. As he noted in his deposition, a distinguished professor in the English department does not earn the same as a distinguished professor in his department (Economics), because the pay is not set by the work that makes them "distinguished" alone but as much if not more by their department.[107]

---

[104] Neumark Deposition, 357:23-25.
[105] Neumark Deposition, 358:22-359:1.
[106] Neumark Report, p. 15, paragraph 31. He also notes that "there is not a regular performance appraisal process at Oracle; managers are not required to give formal performance appraisals, and frequently do not do so."
[107] Neumark Deposition, 107:11-21: "Pay at the university is not set across departments, right. Pay is set by department. So just -- I -- I may have a very different salary from a distinguished

"Deciding whether residual wage estimates capture discrimination may be more an act of faith than an act of science": Fundamentally, Dr. Neumark's regression model does not compare pay outcomes among employees doing substantially similar work

130. Discrimination itself can never be directly observed in *any* data and simply included in pay models as a control variable. Instead, researchers seek to use control variables to compare employees that are as similar as the data allows in terms of their skill, effort and level of responsibility, and working on substantially similar work tasks, and may if appropriate then they infer that any remaining and thus unexplained variation in pay that is correlated with gender may be due to discrimination. These analyses are always, however, subject to the criticism that the model is not in fact comparing employees who are economically similar and thus that no such inference of discrimination can be made. As Dr. Neumark has written, "Deciding whether residual wage estimates capture discrimination may be more an act of faith than an act of science."[108]

131. As outlined above, statisticians and econometricians call the failure to include variables that are correlated with variables in the regression model that are of interest (e.g., gender) "omitted variable bias." In the human capital literature, a considerable amount of research has focused on the biases caused by omitting important statistical determinants of earnings when conducting studies designed to measure the impact on earnings of other specific variables of interest.[109] For example, some of the early studies in the human capital literature sought to

professor in another department. There's no -- there's no -- there's no pretense that there's this, you know, assistant professor, associate professor. These other ranks is the way that -- it's -- it's a ladder within a field, but it doesn't determine your pay."

[108] Hellerstein, Judith K., and David Neumark (2006). Using Matched Employer-Employee Data to Study Labor Market Discrimination. In *Handbook on the Economics of Discrimination*, ed. William Rodgers (pp. 29–60). Cheltenham, UK: Edgar Elgar.

[109] Mincer, Jacob, *Schooling, Experience and Earnings*, National Bureau of Economic Research, 1974, pp.83-96. Griliches, Zvi, (1977) Estimating the Returns to Schooling: Some Econometric Problems. *Econometrica*, 45(1), pp.1-22. Willis, Robert, (1986) Wage Determinants: A Survey and Reinterpretation of Human Capital Earnings Functions, Chapter 10, *Handbook of Labor*

estimate the financial return to schooling. The interest in this area was partly to understand if schooling, and in particular higher education, was a good investment, not just for an individual, but for society. Initial estimates were that there was a strong and highly positive impact of college education on earnings. The policy implication was that significant public funds should be devoted to increasing the numbers of persons attending college. Subsequently, labor economists determined that there was an important omitted variable that biased upwards the estimated rate of return to college education. According to the economic theory, those who can earn most from investing in a college education are the same ones who would be the most likely to do so. And in fact, it was clear that they did so. These individuals are likely to have higher intelligence, or ability. In other words, in this case the omitted variable in estimating the true return to schooling was ability – when measures of ability are inserted into earnings regressions that seek to estimate the return to schooling, the regression coefficient on schooling falls considerably. While in retrospect this is an obvious example of omitted variable bias, the principle applies equally to analysis in a discrimination context. Indeed, one of Dr. Neumark's contributions to the discrimination literature has been to suggest researchers move away from sole reliance upon pay regression models of exactly the type he presents in his report and toward other methods that are not subject to the same concerns about omitted variable bias.

132. When confronted with his own writings in this area, Dr. Neumark claimed there is no contradiction between the opinions expressed in his publications and what he does in this case, because he claims that the Oracle data are "even more detailed" than the usual job-level data available to labor economists[110] and that "a researcher would die to have this much detail… on

Economics, Volume 1, Edited by Orley Ashenfelter and Richard Layard. Elsevier Science Publishers BV.

[110] Neumark Deposition, 181:4-5.

workers and jobs at the same company."[111] The implication appears to be that company data are much more detailed than what is usually available to labor economists, and therefore Oracle's company data must be sufficiently detailed to understand which employees are doing substantially similar work. As I have shown above, Dr. Neumark may have indeed had a lot of information with which to work in this case when compared with having to rely upon census or other public data, but he failed to utilize much of it. In addition, he utilized variables that simply do not stand up to scrutiny as variables that hold the nature and circumstances of work constant such that reliable inferences regarding the meaning of his gender coefficients. Just because the data he used here comes from one company does not mean that all relevant variables have been accounted for.

133. Consider Exhibit 13 in Dr. Neumark's report. The first column shows the results of a regression model with no other control than gender. The table reports that the probability of finding a coefficient on female of -0.1469 under a null hypothesis of no gender gap is less than 1 in 1 billion. Yet not even Dr. Neumark would seriously argue that this is evidence of discrimination, because discrimination is left as something to be inferred <u>only after accounting for the relevant variables.</u> Dr. Neumark does not claim that this "raw" result means anything, but importantly, he reports it and fails to state that it means essentially nothing. Every variable is omitted except female in this "model," and the "1 in a billion" statistic is therefore meaningless.

134. However, in this matter, Dr. Neumark appears to simply "assume away" any issues with regard to whether jobs properly group employees in terms of skill, effort, responsibility, and working conditions. He claims that "we adjust the pay gap for differences in the jobs employees hold, and the skills they have, so that we are comparing pay between women and men in similar

[111] Ibid, 165:15-18.

jobs with similar skills."[112] In other words, Dr. Neumark simply assumes that software engineers in the same job code have the same skills and, more importantly, are performing similar work. But he has provided no test of that assumption, and therefore he has no basis for concluding that his job code, grade and line of business head controls are sufficient to group employees doing substantially similar work. In the work described above I show, for example, an enormous range in compensation between people doing what Dr. Neumark describes as substantially similar work, simply because their job code/grade and line of business head are held constant. Elsewhere in his academic research, he has described this as a "fundamental" problem." He writes "[…] perhaps the most fundamental problem is that the control variables that are included in X may not fully capture marginal productivity differences. […] The bottom line, in our view, is that because one can always tell a story about an unobservable that is related to productivity […], deciding whether residual wage estimates capture discrimination may be more an act of faith than an act of science." [113]

135. Dr. Neumark did not perform any investigation of the job codes that he believes are sufficiently detailed to account for "differences in the jobs employees hold."[114] He states that "Job code fully encompasses title, function, specialty area, and global career level"[115] and includes dummy variables for other job-related factors not included on that list, concluding that "Including this highly-detailed set of controls in my regression model allows me to compare women's and men's pay within very narrowly defined jobs."[116] But simply stating that the job

---

[112] Neumark report, paragraph 11.
[113] Hellerstein, Judith K., and David Neumark (2006). Using Matched Employer-Employee Data to Study Labor Market Discrimination. In *Handbook on the Economics of Discrimination*, ed. William Rodgers (pp. 29–60). Cheltenham, UK: Edgar Elgar.
[114] Neumark Report, p. 7, paragraph 11.
[115] Ibid, p. 13, paragraph 27.
[116] Ibid, p. 13-14, paragraph 27.

codes reflect "narrowly defined jobs" does not make it so. Dr. Neumark provides no qualitative or quantitative analysis to back up this statement.

136. As shown elsewhere in this report, there is a wide range of salary outcomes within the Oracle job codes. If, as Dr. Neumark claims, all workers within a job code are "performing substantially equal work in jobs the performance of which required substantially equal skill, effort, and responsibility, performed under similar working conditions,"[117] then one would not expect to see such variation in the pay outcomes for those ostensibly similar workers. He claims that he "treated persons in the same job code and grade as performing substantially the same or similar work, which is how Oracle treats such persons,"[118] but provides no detail on what he means by "treats." He certainly cannot mean that Oracle pays such persons the same, as we observe wide variation in the pay of individuals in the same job code and grade. It appears that he simply assumes that Oracle must agree that they are "performing substantially the same or similar work" without any basis for that assumption.

137. My review of the data in this case instead suggests that Oracle, like many large employers, utilizes a job code nomenclature in order to organize their workforce into buckets for various purposes. However, one cannot simply assume these categorizations are for the purpose of allowing a labor economist to identify employees who are in fact performing work similar enough to be deemed comparators in an equal pay or pay discrimination context without looking further at the actual content of the jobs themselves..

138. Dr. Neumark, however, takes all of these jobs and aggregates them into a single model. As I showed earlier, different jobs face different pay structures. For example, employees in the Manager Career Levels have a [redacted]

---

[117] Ibid, p. 4, paragraph 8.
[118] Ibid, p. 4, paragraph 8.

than employees in the IC Career Levels, and as the exhibits showed, this varies dramatically between lower level managers and higher level managers. It is not necessarily only the composition of pay that differs. If some jobs, for example, prize cutting edge programming and product design skills that are rare, then tenure might be less important in explaining pay for those jobs. Certainly there do appear to be varying returns to tenure in the data. The graph below shows the regression coefficient on tenure after estimating Dr. Neumark's model separately by system job title.[119]

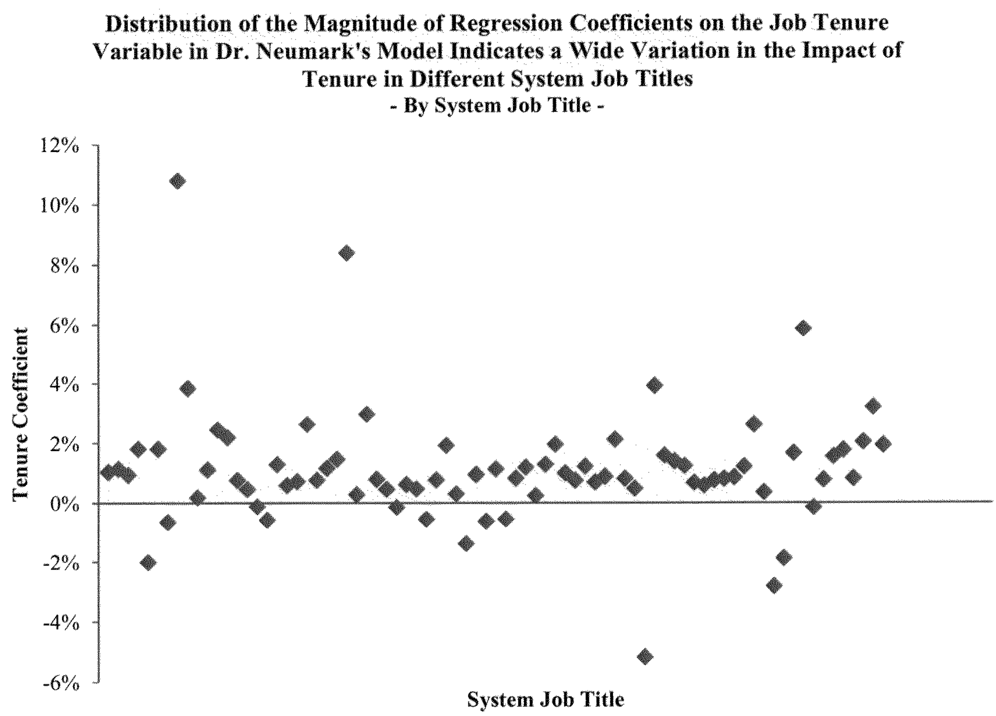


Exhibit 53

[119] These system job titles have at least 100 observations in the incumbent base pay regression population.

139. A statistical test, the F-test, can be used to test whether or not the magnitude of the regression coefficients differ in a statistically significantly manner between any two subgroups. This makes it possible to compare the coefficient on tenure in one system job title to each of the tenure coefficients in other system job titles.[120] An F-test indicates whether the pair of coefficients is statistically significantly different. This test is repeated for all pairs, and the results are shown in the chart below.

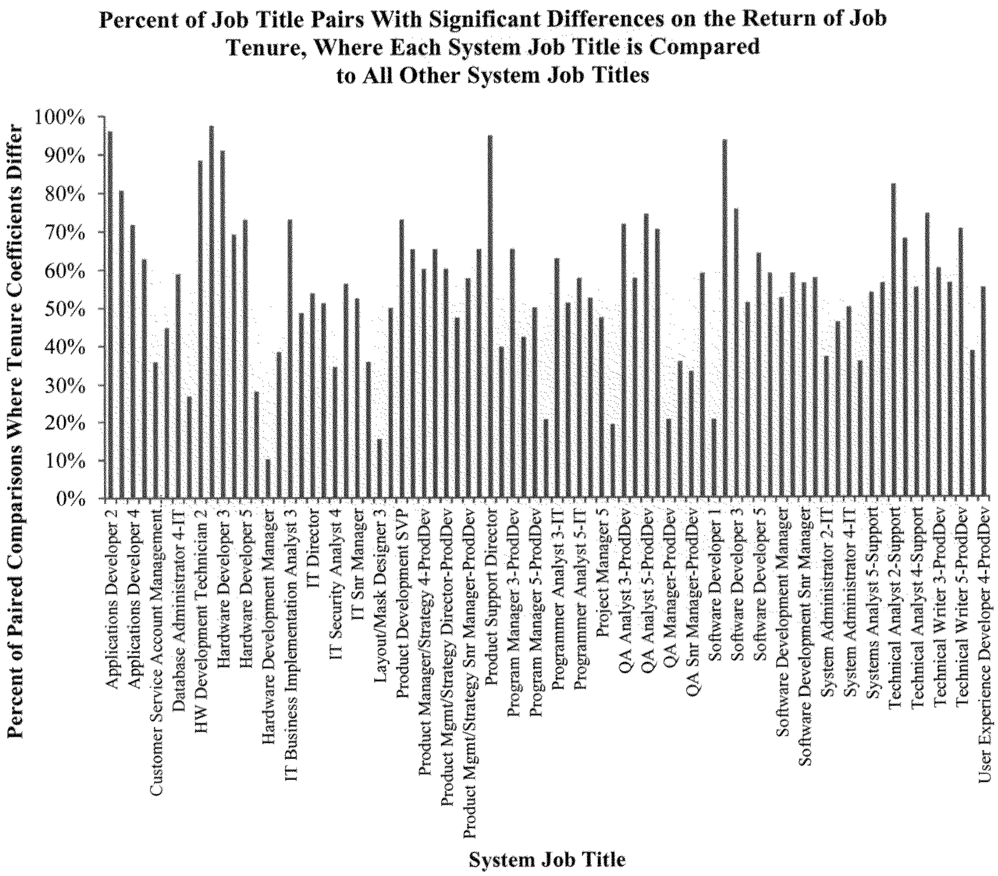


Exhibit 54

---

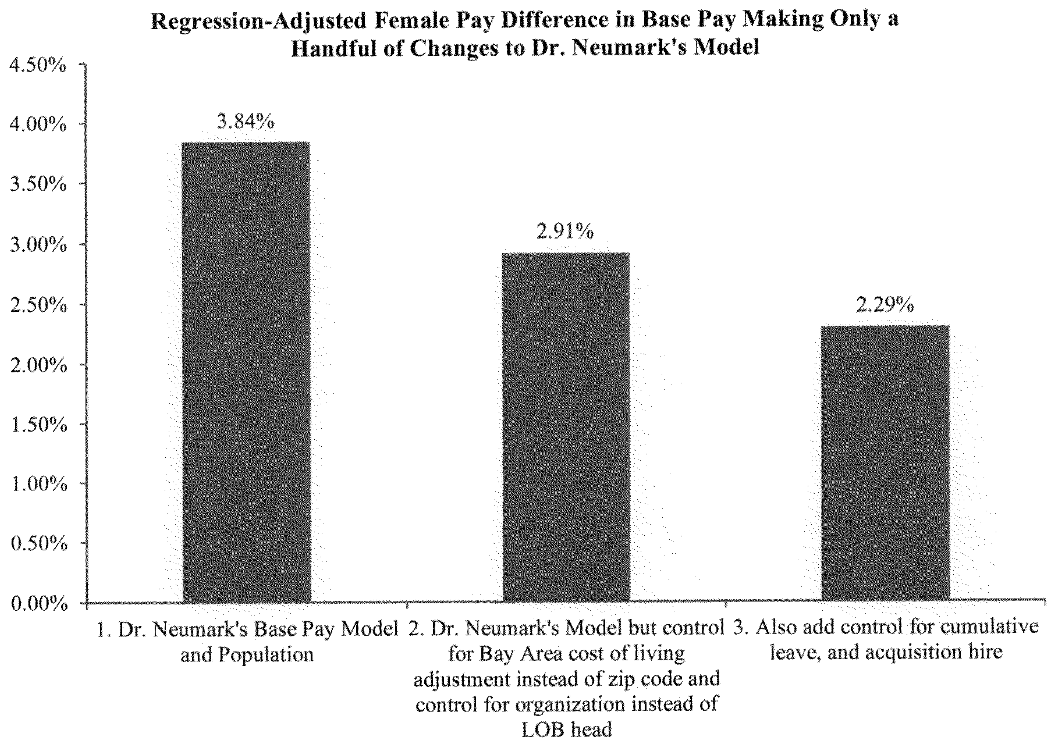
[120] This analysis is based on the regressions by job codes from the previous exhibit.

140. The results of this analysis show that the sizes of the coefficients vary widely across different subsets of employees. It is entirely possible that this is a result of different parts of the company rely on different factors in setting pay. And yet Dr. Neumark's one-size-fits-all model fails to allow for or appropriately reflect this variability.

Dr. Neumark's results are cut almost in half simply by adding a handful of readily-available variables

141. In my variability analysis, I took Dr. Neumark's data and regression model as a given. However, that should not be read as an endorsement of his method or his model. Even with the data at hand, I am able to reduce his estimated gender pay gap by 43% simply by including readily available variables that more closely group employees by what they do, and correcting some of the other mistakes he made. The first column of the graph below indicates Dr. Neumark's estimated gender pay gap for base pay from his Exhibit 2. After I replace his zip codes with a variable for working in the Bay Area (because Bay Area employees receive a salary bump relative to those elsewhere in California),[121] and replace line of business head with organization (which is more closely correlated with specific products and services than LOB head), the gender gap in pay is reduced down to 2.91%, with no other changes whatsoever to the structure or content of his model. This is depicted by the middle blue bar in the graph. The third bar shows the estimated gender pay gap (2.29%) after I add variable that account for all leaves of absence and a squared term, and a flag for acquisition hires (because their leave information is not carried over during the acquisition).

[121] Waggoner Deposition, 174:12-175:20.

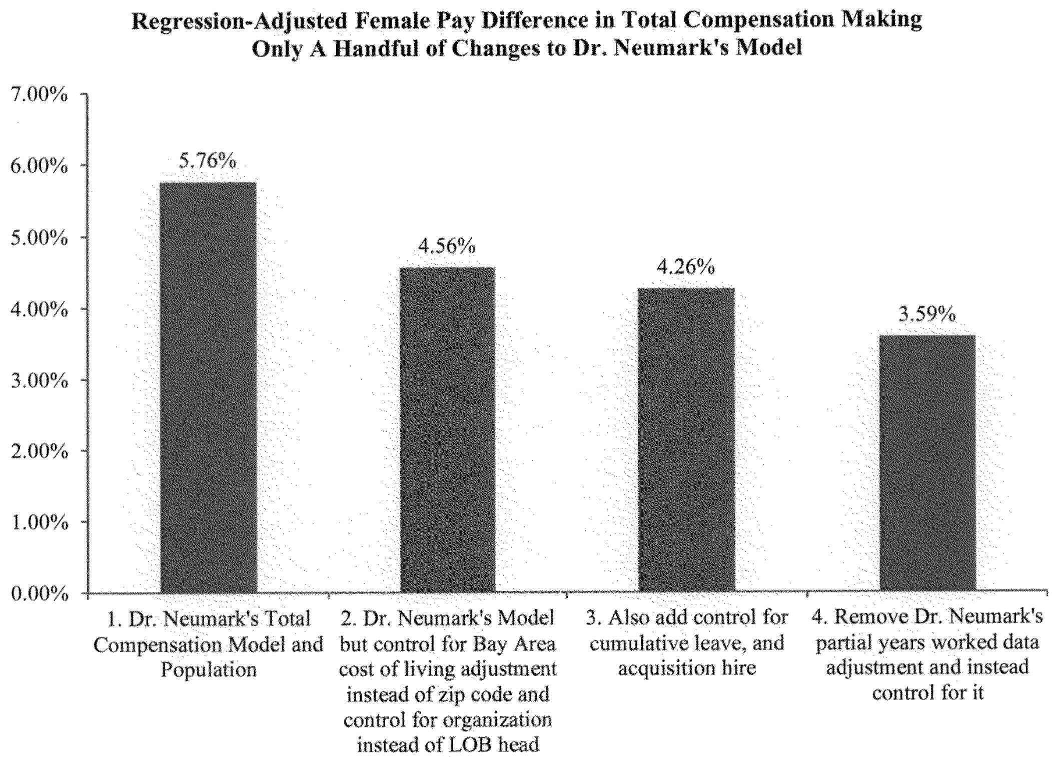


Note: Dr. Neumark's base pay model controls for female, hourly, part-time, Oracle US tenure, Overall Tenure, Tenure in Job, experience, zip code, job code and grade interaction, and LOB head.

Exhibit 55

142. A similar exercise can be performed on Dr. Neumark's total compensation regression model (keeping in mind that Dr. Neumark's treatment of stock awards makes his measure of total compensation highly dubious). The first bar on the chart below shows the gender gap according to Dr. Neumark's model. Simply controlling for organization and whether the employee is in the Bay Area rather than for line of business head and home zip code reduces the gender gap by 21%. Adding cumulative leave and acquisition status to the model reduces the gap by an additional 6%. The last bar on the right shows the effect of undoing Dr. Neumark's flawed annualization of total compensation. If instead of inflating everything for part-year

employees, I simply use their actual total compensation as the dependent variable to be explained by the regression and include a set of indicator variables on the right hand side controlling for what portion of the year they worked, the estimated gender gap is reduced by an additional 15%.



Note: Dr. Neumark's total compensation model controls for female, part-time, Oracle US tenure, Overall Tenure, Tenure in Job, experience, zip code, job code and grade interaction, and LOB head.

Exhibit 56

143. To be clear, I do not believe that organizations, days not worked, and fraction of the year worked alone are sufficient to fully identify the productivity of individual employees or to group them according to the skills or responsibilities their jobs require, for the reasons I detail above.

Nor do I trust Dr. Neumark's definition of total compensation given the issues described at length above with how he valued them. These "tweaks" to his model are presented simply to show, using a handful of readily available data already at hand, that Dr. Neumark's models are flawed and greatly overstate the extent of any difference in pay between men and women, and for that further reason do not support an inference that women are discriminated against by or paid less than similar men at Oracle even if one aggregates all of the data as he has done.

CONCLUSION

144. In this final section of my report, I summarize some of the key conclusions I have reached after considering Dr. Neumark's report against the data, documents, and other information available about work at Oracle. Dr. Neumark's analysis of compensation is flawed in a number of ways, and does not support the inference that all women in the proposed class are paid less than men "performing substantially equal work in jobs the performance of which required substantially equal skill, effort, and responsibility, performed under similar working conditions."[122] Applying Dr. Neumark's aggregated company-wide common statistical model to individuals in the putative class produces wide variations in pay outcomes. For example, his model shows that almost half the women analyzed were paid more than other employees considered under *his* model to be performing substantially similar work. This wide variation in outcomes is inconsistent with the notion that a common model can be used to explain and meaningfully analyze the claims of all of the individual women that the putative class in this case would encompass.

145. Dr. Neumark does not include sufficient factors in his multiple regression models to allow him to compare employees who are performing "substantially similar" work. His use of

[122] Neumark Report, p. 4, paragraph 8.b.

job code/job grade together with line of business head does not, contrary to his claim, "very narrowly define" the work that Oracle employees are doing. This is evident from the wide spread in compensation within each of the job code/job grade buckets within each of the line of business heads he analyzes. It is very unlikely from a labor economics perspective that two individuals sharing the identical job code/grade, with one paid double what the other earns, are doing substantially similar work. Yet for virtually all job codes in Dr. Neumark's data, one sees wide ranges in pay. The Oracle job codes are apparently not designed to "narrowly define" the nature of work from the perspective of a labor economist. That Oracle organizes its workforce with a particular hierarchical and task-type structure does not mean that this structure alone is sufficient for a labor economist asked to analyze this data in an equal pay context. These are crude measures from an analytical perspective.

146. Reviewing the thousands of detailed descriptions of the work tasks associated with job requisitions for the same job code reveals wide differences in the nature of these tasks. Dr. Neumark did not review any of these materials. Paralleling the wide variation in the tasks and duties called for by different requisitions, those hired into the company into positions with identical job codes earn widely varying amounts. This variation has little to do with years of labor market experience or age; instead, it appears that if a successful candidate has the requisite specific skills, they can be hired and paid commensurate with the skills and responsibilities that the position requires. For Software Developer 4 jobs, the largest single job code in the data, the range of ages hired at the same pay level spans from 25 to 60, and at any given age, the range of pay is almost a two to one ratio. Job title/code alone tells you almost nothing here. Dr. Neumark's additional job-related variables, like line of business head and job grade, also do not adequately distinguish the type of work being done.

147. What I take away from this is that there are very likely to be systematic differences within the job codes that Dr. Neumark assumes are "very narrowly defined." Dr. Neumark testified that he did not review the thousands of pages of additional, relevant material that was available. As a result, Dr. Neumark's regression model is not correctly specified, and does not compare employees performing substantially similar work.

148. Dr. Neumark also analyzed the relationship between prior pay and starting pay for those employees hired into Oracle from other employers. He concludes that his analysis is consistent with Oracle relying upon prior pay in setting starting pay, and claims that his results are evidence consistent with the notion that women hired by Oracle suffered discrimination in the wider labor market, and further that Oracle's presumed practice of relying upon prior pay imported this labor market discrimination into Oracle. His analysis supports none of these conclusions, due to technical flaws, that once corrected, totally change the statistical results.

149. Dr. Neumark starts his analysis of prior pay by observing that there is a high correlation between prior and starting pay for Oracle. While careful not to directly infer causality from this observation, he nevertheless states in a Table that this correlation is so high as to have a less than one in a billion probability of arising by chance. This statement is extremely misleading. It implies there is something worth taking note of. Tests of statistical significance must use a benchmark against which the observed outcomes are compared. His "one in a billion" tests the hypothesis of no relationship whatsoever between prior and starting pay. This is an impossibility; it implies that there is no expected connection whatsoever between workers' prior and new pay when they change jobs. This is of course untrue, as Dr. Neumark acknowledged in his deposition testimony. An appropriate hypothesis would be to test the Oracle correlation against a benchmark like the observed correlation at other firms, or the observed correlation in

the economy as a whole. Such a comparison would still not indicate anything about causality, but it would be at least more appropriate to test whether the correlation at Oracle is different than one would expect in any company. I looked at the correlation between pay measures in adjacent employee jobs using a nationwide longitudinal data sample, and found results similar to those for Oracle. Dr. Neumark's implication that a correlation between prior and starting pay means anything regarding Oracle's use of prior pay should be entirely discounted.

150. Once I correct the data used to support his prior pay analysis and add a year control variable to account for the almost six years covered by the data, the results completely change and no longer support Dr. Neumark's conclusion that Oracle "imports" gender discrimination from the external labor market into the company.

151. Finally, there are a number of data errors and other technical issues in Dr. Neumark's data. These include incorrect measures of tenure that do not account for leaves of absence correctly (as he testified he does in his academic work). Another deeply problematic error is Dr. Neumark's definition of total compensation. He valued stock on an annual basis (essentially attributing it to when it was taxed and not when it was earned) which causes him to incorrectly assign stock earnings to years other than the year they were earned. Dr. Neumark's total compensation results are uninterpretable, because stock awards make up sizeable percentages of compensation, especially in the higher career levels and in management in particular.

152. For the reasons enumerated herein, it is my opinion that Dr. Neumark's analyses are mis-specified, suffer from omitted variable bias, and have a number of important methodological flaws. As a result, my opinion is that Dr. Neumark's analyses do not permit the inferences he seeks to make.

Attachment A: Saad CV and Testimony

resolution economics LLC

1925 Century Park East 15th Floor
Los Angeles, CA, 90067
Office: 310.275.9137
Fax: 310.943.3529

ALI SAAD, Ph.D., MANAGING PARTNER

Dr. Saad is the Managing Partner of Resolution Economics LLC. He has a Ph.D. in Economics from the University of Chicago. Prior to Resolution Economics, Dr. Saad was a partner at Deloitte & Touche LLP and at Altschuler, Melvoin and Glasser LLP. Before that he was in the disputes consulting group at Price Waterhouse, first in New York, and then in Los Angeles. Prior to his consulting career, Dr. Saad served as an Assistant Professor of Economics at Baruch College of the City University of New York (CUNY).

Professional Experience

Dr. Saad's experience is extensive in the area of statistical and economic analysis of liability and damages related to employment litigation matters. His experience is extensive in the application of economics and statistical methods to class action employment discrimination matters. He is also experienced in designing, implementing, and analyzing surveys and observation studies as well as conducting empirical analyses related to exempt/non-exempt status, hours worked, uncompensated time, meal and rest breaks, rounding, and other wage and hour issues. He has also performed statistical and damages analyses for a broad range of commercial litigation matters including breach of contract, insurance coverage, environmental claims, patent infringement, antitrust and real estate financing. Dr. Saad has testified a number of times at deposition and trial. Dr. Saad also regularly consults to clients regarding business issues related to employment practices.

Employment Matters

Dr. Saad provides a variety of services related to employment litigation. His experience is extensive in conducting statistical and economic analysis related to issues of liability for employment discrimination matters. He also has designed and conducted many surveys and observational studies related to wage and hour issues. Dr. Saad has also performed analyses of economic damages in both class action and single plaintiff matters.

Statistical and Economic Analysis in Discrimination Matters

Assignments representative of Dr. Saad's experience in performing analyses in connection with employment discrimination matters include the following:

- Consulting and expert witness services in national class action race discrimination matter involving issues of pay, promotion, work assignment, and a variety of other challenged employment practices. Services included creating databases from diverse and voluminous source materials, and conducting extensive statistical analyses.

- Consulting and expert witness services in national class action gender discrimination matter involving issues of job assignment and promotion. Services included creating databases from diverse and voluminous source materials, and conducting extensive statistical analyses.

- Consulting and expert witness services in a class action case alleging that contracts were misleading. Services included processing and analyzing large quantities of data, and performing statistical analysis of the criteria determining class membership.

- Consulting and expert witness services in connection with a major class action alleging gender discrimination in pay and promotion at a large high-tech employer. Services included creating analytical databases, and developing economic and statistical arguments concerning the relationship between productivity-related variables, pay/promotion, and gender.

- Consulting and expert witness services in an antitrust and discrimination matter in which a group of businesses alleged violations of antitrust and discrimination laws by another group of businesses. Services included data construction, and statistical analysis related to issues of liability.

- Consulting and expert witness services on behalf of plaintiffs' counsel in a series of cases alleging race discrimination in hiring. Services included creating analytical databases, studying the relationship between race and hiring, and examining the features of the external labor market.

- Consulting and expert witness services in connection with a class action claim of discrimination based on age in connection with a series of layoffs resulting from the combination of two large retail chains. Services included creating analytical databases, studying the relationship between layoff and age, and examining the relationship between age and workforce composition over.

- Consulting and expert witness services in connection with EEOC allegations of race discrimination in recruiting, hiring, and initial placement at a large service providing company. Services included developing databases from diverse paper and electronic sources, and providing statistical arguments concerning the relationship between race and various other factors.

- Consulting and expert witness services to defendant's counsel in connection with a major class action alleging gender discrimination in multiple employment practices at a national retail chain. Services included developing a database from voluminous paper documents, and conducting analysis related to hiring, initial placement, and initial pay.

- Consulting and expert witness services to defendant's counsel in connection with an EEOC investigation of racial discrimination in hiring by a major service providing organization. Services included developing a database, and conducting statistical analysis related to hiring.

- Consulting services to defendant's counsel in connection with a U.S. Department of Labor OFCCP investigation of pay equity at a high-tech company. Services included design and oversight of a statistical analysis of pay equity, assessment of the OFCCP methodology, and participation in conciliation discussions between the company and the OFCCP.

- Consulting and expert witness services to defendant's counsel in connection with an allegation of age discrimination in terminations resulting from a series of mass layoffs. Services provided included developing statistical arguments concerning the relationship between age and termination.

- Consulting services to defendant's counsel in connection with a Department of Justice investigation regarding allegations of racial profiling by a large city police department. Analyzed departmental data related to over 130,000 traffic stops, pedestrian stops, and other types of police contacts that occurred in four selected weeks in 1997 and four selected weeks in 1999. Cross-referenced traffic stops data with other information sources including human resources data, precinct level paper records, and the officer discipline system to test various hypotheses.

- Consulting services and expert testimony to defendant's counsel in connection with a multi-plaintiff matter alleging race and gender discrimination in promotion and placement into coveted positions by a large city police department. Performed statistical analysis of promotion and placement into coveted positions. Quantified economic damages for several plaintiffs under failure to promote and wrongful termination theories.

- Consulting services in a case against a city government alleging discrimination in recruiting and hiring of police and firefighters. Services included using Census and other large-scale data sources to assess labor market characteristics by detailed geographic location, and conducting extensive analysis of the impact of employment tests on hiring.

- Consulting and expert witness services to defendant's counsel in a matter where plaintiff alleged that defendant's hiring practices discriminated against women. Services included converting diverse paper source materials into a usable database, and developing statistical evidence concerning plaintiff's allegation.

- Consulting services in several class action recruiting and hiring matters. Services included use of detailed census and other data to estimate labor market availabilities by geographic location, and analyzing employment practices in light of these availability findings.

- Consulting services to a major bank involved in an analysis of its fair lending practices. Services included using bank data on applicants for mortgages and other loans, and adding various demographic and geographic information to assess if the bank made loans on the basis of race, or controlling for other, observable factors could explain patterns in loan making.

- Consulting services on behalf of defendant's counsel in a major class action matter involving allegations of gender discrimination in promotion. Services included building analytical database from many sources, using the database to conduct extensive statistical analysis of plaintiffs' allegations, and estimating damages resulting from non-promotion for approximately 3,000 women occupying different jobs over a ten-year period.

- Consulting and expert witness services on behalf of defendant's counsel in two related cases alleging age discrimination in termination. Prior to plaintiffs' vesting for certain long term benefits. Services included using defendant's human resource data to test plaintiffs' specific allegations, developing statistical arguments concerning the relationship between age and termination, and performing analyses of plaintiff's damages in each case.

- Consulting services on behalf of plaintiff's counsel in distribution of award in an age discrimination matter with 75 plaintiffs. Services included developing a method to efficiently compute damages for all plaintiffs, and working with counsel, an arbitrator, and plaintiffs' committee to explain the process to plaintiffs' group.

Wage and Hour Matters

Assignments representative of Dr. Saad's experience in wage and hours matters include:

- Consulting and expert witness services to defense counsel in a national class-action wage and hour matter alleging that several thousand loan originators at a large financial institution were misclassified under FLSA. Conducted statistical analyses of hours worked records, compensation data, plaintiffs' declarations, and other data to determine if select groups of plaintiffs would be representative of the class.

- Consulting and expert witness services to defense counsel in a wage and hour matter alleging that several thousand General Managers and Assistant Managers at a large office supply retailer were misclassified as exempt employees. Services included designing and conducting a survey to examine whether class members were appropriately classified, analyzing the company's labor model and human resources data, and conducting statistical analyses related to a variety of class certification issues.

- Consulting and expert witness services to defense counsel in a wage and hour matter alleging that several thousand Assistant Managers at a large general merchandise retailer were misclassified as exempt employees. Services included designing and conducting both a survey and an observational study, to examine whether or not class members were appropriately classified. Services also included conducting extensive statistical analyses of the data collected by the survey and the observational study, and preparing materials for use in class certification proceedings.

- Consulting services to defense counsel in a class action matter alleging failure to pay overtime wages to independent sales and service representatives for a large national tool franchiser. Services included designing and implementing an hours survey to determine whether the additional hours worked claimed by some plaintiffs was representative of the additional hours worked by the class as a whole. Determined that the problem was isolated to certain geographic areas rather than nationwide.

- Consulting and expert witness services to defense counsel in a wage and hour matter alleging that several hundred store managers and assistant store managers at a chain of retail discount stores were misclassified. Services included creating and implementing a survey to examine whether class members were classified appropriately and conducting statistical analyses related to commonality of class-members and other class certification issues.

- Consulting services to defense counsel in a multi-plaintiff wage and hour matter alleging that the defendant employer failed to compensate security guards for uniform changing time and other claims of off-the-clock work. Services included designing and conducting an observation study to measure time associated with various activities.

- Consulting services to defense counsel in wage and hour matter alleging that store managers at a chain of convenience store/ gas station operations were misclassified as exempt workers. Services included designing and conducting a random sampling scheme and observational study to evaluate the amount of time that class members spent on exempt and non-exempt duties.

- Consulting services to defense counsel in a class-action wage and hour matter alleging uncompensated meal periods and breaks, unpaid overtime wages, and minimum wage violations at a field maintenance company.

Services included creating a database of hours worked from paper and electronic records, and then providing damages estimates based on a variety of assumptions and legal theories.

- Consulting services to defense counsel in a class action matter alleging a variety of wage and hour violations for hourly workers at a chain of warehouse stores. Services included analyzing data to test allegations of improper time adjustments, missed meal and rest periods, uncompensated split shifts, reporting time violations, overtime and regular rate issues, and off-the-clock work.

Employment Damages

Assignments representative of Dr. Saad's experience estimating economic damages include the following:

- Consulting services to plaintiff's counsel in a case involving a breach of employment contract allegation by a high-level executive in the emerging communications industry. Services included damages analysis based on valuation of stock options and estimation of future earnings.

- Consulting services to defendant's counsel in a case involving a wrongful termination allegation by a high-level executive in the telecommunication industry. Services included damages analysis based on valuation of stock options using the Black-Scholes Option Pricing Framework and a Monte Carlo Simulation Model.

- Consulting and expert witness services on behalf of defendant's counsel in a matter brought by a former executive who alleged wrongful termination and age discrimination against a major defense contractor following a reduction in force. Critiqued work product of the opposing expert, evaluated mitigation issues, calculated loss of earnings damages and valued losses related to stock options.

- Consulting and expert witness services on behalf of defendant's counsel in a medical malpractice action where the underlying damages issue was valuing an income stream from a closely held cash business. Performed accounting of plaintiff's financial records to determine the existence and the extent of fraud. Created financial models to calculate damages under a variety of scenarios.

- Consulting and expert witness services to defendant's counsel in a wrongful termination matter brought by senior executive of a high-tech company who alleged age discrimination. Performed analysis of mitigation factors, calculated loss of earnings, and valued future stock options.

Commercial Litigation

Dr. Saad has assisted clients in a variety of commercial litigation matters, including patent infringement, insurance coverage, antitrust, breach of contract, and real estate financing. Assignments representative of Dr. Saad's experience in these areas include the following:

- Consulting and expert witness services in a series of cases involving the real property title insurance industry. Services included performing extensive statistical analyses in connection with both liability and damages issues.

- Consulting and expert witness services in a case alleging breach of loan commitment to a commercial real estate concern. Services included constructing financial models, developing economic arguments relating to fixed versus variable rate loans, and assisting counsel in deposing the opposing expert.

- Consulting and expert witness services in a case involving a breach of contract allegation in the computer hardware industry. Services consisted of performing a damages calculation, and rebutting the opposing expert's analysis.

- Consulting and expert witness services in a case alleging that one entity caused another entity's property to be misused. Services included database creation, and statistical analysis related to issues of causation. Results indicated that there was a statistically significant relationship between defendant's actions and plaintiff's economic condition.

- Consulting services on behalf of defendant's counsel in a breach of contract matter in the context of natural resource raw materials shipping. Services included developing economic arguments regarding the but-for pricing of both the shipping service as well as the material being shipped.

- Consulting and expert witness services on behalf of defendant's counsel in a major insurance coverage case, in which the underlying claims resulted from tens of thousands of asbestos claims. Services included developing strategy for dealing with large amounts of paper information, creating a database for analysis, and performing a variety of statistical analyses.

- Consulting services on behalf of plaintiff's counsel in an antitrust matter in the consumer electronics product market. The antitrust practice alleged was predatory pricing. Services included preparing a damage analysis.

- Consulting services on behalf of defendant's counsel in a patent infringement matter in the computer hardware industry. Services included researching transfer pricing issues and analyzing complex company P&L data in preparation for damages calculation.

- Consulting services on behalf of defendant's counsel in a real estate financing dispute. Dispute revolved around the financing of a major New York office property. Services included analysis of interest rates and their relationship to potential damages at various points in time, as well as the construction of a financial model of the property with the but-for financing in place.

- Consulting services on behalf of plaintiff's counsel in an antitrust matter involving allegations of non-competitive practices and predatory pricing in the home cable television market. Services included an analysis of "raising rivals costs", as well as a statistical analysis of pricing of complex products over time.

Summary of Employment Experience

Resolution Economics LLC:
Managing Partner, October 1998 to date.

University of Southern California
Adjunct Associate Professor in the Department of Economics, January 1999 to September 2001.

Deloitte & Touche, LLP:
Partner, Dispute Consulting Services, (Los Angeles), 1998.

Altschuler, Melvoin and Glasser LLP:
Partner, Economics and Litigation Services, (Los Angeles), 1995 to 1998.

Price Waterhouse LLP:
Senior Manager, Manager, Litigation and Corporate Recovery Services Group, (New York and Los Angeles), January 1989 – November 1989, June 1990 to 1995.

Olympia & York Companies (USA):
Assistant VP and Senior Economist, (New York), November 1989 - June 1990.

Baruch College, City University of New York (CUNY):

Instructor and Assistant Professor of Economics, Department of Economics and Finance, 1982-1988; Center for the Study of Business and Government, Research Associate, 1983-1986; U.S. Small Business and Veterans Administrations, Consultant, 1985-1986.

Education

Ph.D., Economics, The University of Chicago.

B.A., History, Economics, The University of Pennsylvania

Publications

Financial Success and Business Ownership among Vietnam and other Veterans (with S. Lustgarten) SBA - 7210 - VA - 83, 1986.

"Schooling and Occupational Choice in 19th Century Urban America", Journal of Economic History, vol. 49, no. 2, June 1989.

"Employment Discrimination Litigation", chapter in Litigation Services Handbook, ed. by Roman Weil, et al., 1995, 2001, 2006, 2012, 2017.

"Employment Discrimination", chapter in Litigation Support Report Writing, ed. by Jack P. Friedman, et al, 2003.

Paul Grossman, Paul Cane, and Ali Saad, "Lies, Damned Lies, and Statistics: How the Peter Principle Warps Statistical Analysis of Age Discrimination Claims", The Labor Lawyer, vol. 22, no. 3, Winter/Spring 2007, pp. 251-268.

Saad, Ali, "Beyond the Peter Principle – How Unobserved Heterogeneity in Employee Populations Affects Statistical Analysis in Age Discrimination Cases: Application to a Termination/RIF Case", AELC Conference Volume, 2007.

Saad, Ali, "Filling the Data Vacuum in Wage and Hour Litigation: The Example of Misclassification Cases, Emphasis on Class Certification", SIOP Annual Conference Proceedings, 2009.

Saad, Ali, "Wage and Hour Cases - Filling the Data Vacuum: Misclassification Cases and Other Observational Studies", SIOP Annual Conference Proceedings, 2012.

Presentations

Dr. Saad has delivered many presentations at professional conferences, to law firms and to industry groups.

Academic Honors

Finalist, Allan Nevins National Doctoral Dissertation Award
NIMH Doctoral Fellowship, The University of Chicago
Magna Cum Laude, The University of Pennsylvania
Honors in History, Economics, The University of Pennsylvania
Omicron Delta Epsilon, Honor Society in Economics

Professional Affiliations

American Economic Association
American Bar Association (associate membership)

Ali I. Saad, Ph.D.
Attachment to Resume

Testimony, Expert Reports, and Declarations:

In the matter of Smiles, et al., v. Walgreen Company, et al., Case No: RG-17862495 (Superior Court for the State of California) in connection with wage and hour claims. Report filed February 22, 2019, deposition testimony February 25, 2019.

In the matter of Kennard v. Reeves, Case No: BD 604 788 (Superior Court for the State of California) in connection with reasonable compensation issues. Reports filed January 28, 2019 and February 4, 2019. Arbitration Testimony February 22, 2019.

In the matter of PAntionne, et al., v. The School Board of Collier County, Florida, et al., Case No: 2:16-cv-00379-SPC-MRM, United States District Court for the Middle District of Florida) in connection with employment discrimination claims. Report filed February 5, 2019.

In the matter of Pineda v. Abbot Laboratories, et al., Case No: CV18-3395 SVW (RAOx) (United States District Court for the Central District of California) in connection with employment discrimination claims. Report filed November 16, 2018.

In the matter of EEOC, et al., v. Jackson National Life Insurance, et al., Case No: 16-CV-2472-PAB-SKC, (United States District Court for the District of Colorado) in connection with class action discrimination claims. Reports filed August 31, 2018 and October 26, 2018.

In the matter of Florida Education Association, et al., v. State of Florida Department of Education, et al., Case No: 4-17-cv-414-RH/CAS, (United States District Court For the Northern District of Florida) in connection with class action discrimination claims. Reports filed September 28, 2018 and July 5, 2018.

In the matter of Leanna Delgado v. California Commerce Club, Inc., et al., Case No: BC 586727, (Superior Court for the State of California for the County of Los Angeles) in connection with allegations of age discrimination. Deposition July 25, 2018.

In the matter of Hall v. Rite Aid Corporation, Case No. 37-2009-00087938-CU-OE-CTL, (Superior Court for the State of California for the Country of San Diego) in connection with suitable seating claims. Deposition January 20, 2012, Report filed on June 11, 2018.

In the matter of Harris, et al., v. Union Pacific, Case No: 8:16-cv-381, (United States District Court For the District of Nebraska) in connection with class action discrimination claims. Report filed May 3, 2018. Deposition May 23, 2018.

In the matter of Henderson, et al., v. JP Morgan Chase, Case No. 11-CV-03428 (PLAx), (United States District Court For the Central District of California) in connection with wage and hour claims. Report filed February 26, 2018. Deposition March 21, 2018.

In the matter of Moussouris, et al., v. Microsoft, Case No. 15-CV-1483 (JLR), (United States District Court for the Western District of Washington) in connection with class action claims of gender discrimination in pay, performance and promotions. Reports filed January 5, 2018, April 6, 2018 and April 25, 2018. Deposition January 30, 2018.

In the matter of Creative Artists Agency LLC, v. Martin Lesak, et al., JAMS Ref nos. 120032335, 336 and 337 (Arbitral Tribunal of JAMS) in connection with breach of contract claims. Deposition January 16 and 21, 2018 and March 19, 2018. Arbitration testimony March 26, April 16, and September 7, 2018.

In the matter of Negrete, et al., v. Conagra Foods, Inc., Case No. 2:16-cv-631-FMO-AJW, (United States District Court For the Central District of California) in connection with class action wage and hour claims. Report filed February 28, 2018. Deposition April 18, 2018. Revised report filed on June 18, 2018 to respond to a revised report filed by plaintiff's expert.

In the matter of Woods, et al., v. JFK Memorial Hospital, Inc., Case No. INC 1205209, (Superior Court of California, County of Riverside), in connection with wage and hour claims. Report filed October 13, 2017. Deposition November 29, 2017.

In the matter of Bridewell-Sledge, et al., v. Blue Cross of California, et al., Case No. BC 477 451 c/w BC 481 586, (Superior Court of California, County of Los Angeles), in connection with employment discrimination claims. Reports filed September 7, 2017 and June 13, 2018. Deposition October 30, 2017.

In the matter of Stewart, et al, v. Hat World, et al., Case No. CIV 533617, (Superior Court of California, County of San Mateo), in connection with wage and hour claims. Report filed September 7, 2017.

In the matter of Truitt, et al., v. Atlanta Independent School System, Case No. 1:15-cv-4295-SCJ-WEJ, (United States District Court, Northern District of Georgia, Atlanta Division), in connection with allegations of employment discrimination. Report filed August 31, 2017. Deposition September 20, 2017.

In the matter of Williams, et al., v. TGI Fridays, Inc. Case No. 15-cv-0426, (United States District Court, Northern District of Illinois), in connection with allegations of wage and hour violations. Report filed August 4, 2017, deposition August 25, 2017.

In the matter of Whitaker, et al., v. U.S. Renal Care, Inc., et al, Case No. 1:17-cv-02661-GJSx), (United States District Court, Central District of California), in connection with allegations of wage and hour violations. Report filed July 17, 2017.

In the matter of Victor Cejka, et al., v. Vectrus Systems Corporation, et al. Case No. 15-cv-02418-MEH, (United States District Court, District of Colorado), in connection with alleged employment damages. Report filed July 17, 2017, Rebuttal report filed August 14, 2017. Trial testimony June 18, 2018.

In the matter of EEOC, v. GMRI, Inc. Case No. 15-cv-20561-JAL, (United States District Court, Southern District of Florida, Miami Division), in connection with allegations of employment discrimination. Report filed April 21, 2017, deposition June 8, 2017.

In the matter of Coordinated Proceedings, Special Title, Staples Wage and Hour Cases (Included Actions: Lawson, et al. v. Staples Contract and Commercial, Inc., Los Angeles County Superior Court, Case No. BC542237, and Rosales v. Staples Contract and Commercial, Inc., San Bernardino Superior Court, Case No. CIVDS 1505146), in connection with wage and hour allegations. Report filed May 16, 2017.

In the matter of Curtis Patton, et al. v. Dollar Tree Stores, et al. Case No. 2:15-cv-03813 MWF-PJW, (United States District Court, Central District of California), in connection with wage and hour allegations. Rebuttal Report filed May 15, 2017.

In the matter of Bowerman, et al., v. FAS, Civil Action No. 13-00057-WHO, (United States District Court, Northern District of California), in connection with wage and hour allegations. Rebuttal Report filed April 6, 2017, deposition April 11, 2017.

In the matter of Romero, et al., v. Allstate Insurance Company, et al., Consolidated Cases, Civil Action No. 01-3894-MAK, (United States District Court, Eastern District of Pennsylvania), in connection with employment discrimination allegations. Rebuttal Report filed March 20, 2017, deposition March 29, 2017.

In the matter of Wall v. HP Inc., et al. Case No. 30-2012-00537897-CU-BT-CXC, (Superior Court of the State of California, County of Orange), in connection with wage and hour allegations. Declaration March 14, 2017.

In the matter of Controulis, et al., v. Anheuser-Busch, LLC, et al., Case No. BC-518518, (Superior Court of the State of California, County of Los Angeles), in connection with wage and hour allegations. Report filed December 12, 2016.

In the matter of Urbano, et al., v. SMG Holdings, et al., Case No.: 5:15-cv-00603-MMM (MRW), (United States District Court for the Central District of California), in connection with wage and hour allegations. Report filed October 14, 2016, deposition October 26, 2016.

In the matter of Williams, et al., v. Baker Hughes Oilfield Operations, Case No.: 1:15-cv-00049-RRE-ARS, (United States District Court for District of North Dakota), in connection with wage and hour allegations. Reports filed June 24, 2016, January 12, 2017.

In the matter of In re: AutoZone, Inc., Wage and Hour Employment Practices Litigation, Case No.: 3:10-cv-02159-CRB (JSC), (United States District Court for the Northern District of California), in connection with wage and hour allegations. Report filed April 29, 2016, deposition May 27, 2016.

In the matter of EEOC v. Texas Roadhouse, Inc., et al. Case No.:1:11-cv-11732 (United States District Court for the District of Massachusetts), in connection with allegations of age discrimination. Reports filed April 22, 2016 and July 20, 2016. Deposition June 17, 2016; trial testimony January 26, 2017.

In the matter of Luanna Scott, et al., v. Family Dollar Stores, Inc., Case No.:3:08-cv-540 (United States District Court for the Western District of North Carolina), in connection with allegations of gender discrimination. Reports filed January 28, 2016, May 31, 2016. Deposition February 10, 2016.

In the matter of Brown, et al., v. In-N-Out Burger, Inc., Case No.:RG12646351 (Superior Court for the State of California, County of Alameda), in connection with allegations of age discrimination. Report filed December 23, 2015.

In the matter of Valerie Horvath v. Western Refining Wholesale, Inc., Case no. Case No.:CIV-ds1311846 (Superior Court for the State of California, County of San Bernardino), in connection with allegations of age discrimination. Report filed November 19, 2015. Deposition January 14, 2016.

In the matter of Cortina, et al., v. North American Title Company, Case no. 07 CE CG 01169 JH, (Superior Court of the State of California, County of Fresno), in connection with class action employment matter. Reports filed May 11, 2012, June 25, 2012, and August 13, 19, 21, and 26, 2015. Deposition September 8 and 9, 2015. Trial testimony December 3 and December 10, 2015. Hearing testimony March 14, April 12, May 18, July 12, 2018, September 18, 2018 and November 26th, 2018.

In the matter of Curley, et al., v. Savemart, et al. Case no RG13685740, (Superior Court of California, County of Alameda), in connection with class action wage and hour matter. Report filed September 2, 2015. Deposition December 18, 2015 and January 20, 2016.

In the matter of Gamble, et al., v. Boyd Gaming Corporation, et al., Case no. 2:13-cv-01009-JCM-PAL, (US District Court, District of Nevada), in connection with class action wage and hours claim. Report filed January 26, 2015.

In the matter of Chea, et al., v. Best Buy Stores. Case no 4:14-cv-0020-PJH, (United States District Court, Northern District of California, Oakland Division), in connection with class action wage and hour matter. Report filed March 13, 2015.

Attachment B: Data and Documents Considered

Attachment B – Data and Documents Considered

I. Court Documents

Second Amended Class Action Complaint, in the matter of Rong Jewett, Sophy Wang, and Xian Murray, individually and on behalf of themselves and others similarly situated, v. Oracle America Inc. Superior Court of the State of California, County of San Mateo, filed October 16, 2017

Fourth Amended Class Action Complaint, in the matter of Rong Jewett, Sophy Wang, and Xian Murray, individually and on behalf of themselves and others similarly situated, v. Oracle America Inc. Superior Court of the State of California, County of San Mateo, filed September 7, 2018

Case Management Order 2, filed August 16, 2018

Compendium of Evidence in Support of Defendant Oracle America, Inc.'s Motions for Summary Judgment, or, in the Alternative, Summary Adjudication Vol VI, January 18, 2019

Plaintiff Elizabeth Sue Petersen's Amended Responses to Defendant's Special Interrogatories to Plaintiff Sue Petersen (Set One), September 24, 2018

Plaintiff Manjari Kant's Responses to Defendant Oracle America, Inc.'s Special Interrogatories to Plaintiff Manjari Kant (Set One), October 8, 2018

Plaintiff Marilyn Clark's Responses to Defendant Oracle America, Inc.'s Special Interrogatories to Plaintiff Marilyn Clark (Set One), August 31, 2018

Plaintiff Rong Jewett's Amended Responses to Oracle America, Inc.'s Special Interrogatories to Plaintiff Rong Jewett (Set One), January 22, 2018

Plaintiff Sophy Wang's Amended Responses to Defendant Oracle America, Inc.'s Special Interrogatories to Plaintiff Sophy Wang (Set One), January 22, 2018

Plaintiff Xian Murray's Amended Responses to Defendant Oracle America, Inc.'s Special Interrogatories to Plaintiff Xian Murray (Set One), January 22, 2018

Representative Plaintiffs' Memorandum of Points and Authorities in Support of Motion for Class Certification, January 18, 2019

II. Depositions and Declarations

Videotaped PMK Deposition of Oracle America, Inc., By: Anje Dodson, July 17, 2018

- Exhibits 1 - 22

Videotaped PMK Deposition of Oracle America, Inc., By: Kate Waggoner Volume 1, July 26, 2018

- Exhibits 23 to 46

Videotaped PMK Deposition of Oracle America, Inc., By: Kate Waggoner Volume 2, July 27, 2018

- Exhibits 47 to 64

Videotaped PMK Deposition of Oracle America, Inc., By: Kristina Karstensson Edwards, October 16, 2018

- Exhibits 65 to 73

Videotaped PMK Deposition of Oracle America, Inc., By: Chad Wayne Kidder, October 23, 2018

- Exhibits 74 to 75

Videotaped Deposition of Elizabeth Peterson, September 26, 2018

Videotaped Deposition of Manjari Kant Volume 1, October 19, 2018

Videotaped Deposition of Manjari Kant Volume 2, November 2, 2018

Videotaped Deposition of Marilyn Clark, September 14, 2018

Videotaped Deposition of Rong Jewett, March 23, 2018

Videotaped Deposition of Sophy Wang Volume 1, March 9, 2018

Videotaped Deposition of Sophy Wang Volume 2, March 30, 2018

Videotaped Deposition of Xian Murray, May 11, 2018

Declaration of Srividhya Subramanian in Support of Representative Plaintiffs' Motion for Class Certification, filed January 22, 2019

Videotaped Deposition of David Neumark, Ph.D., February 8, 2019

- Exhibits 1 to 19

Videotaped Deposition of Leaetta M. Hough, Ph.D., January 30, 2019

- Exhibits 1 to 10

Videotaped Deposition of Expert David Neumark, Rabin and Chapman et al. v. Pricewaterhousecoopers, LLP, United States District Court Northern District of California San Francisco Division, Case No. 16-cv-02276-JST, January 12, 2018

III. Oracle Documents

ORACLE_JEWETT_00001550_Reference Guide - Allocations and Changes in Workforce Compensation.pdf

ORACLE_JEWETT_00004881_Alternate Equity Awards HR Presentation 20140430 v2.pptx

ORACLE_JEWETT_00005276_Native_USWorkforceCompHRTrainingFocal.pptx

ORACLE_JEWETT_00005426_Native_WorkforceCompManagerTrainingEquity.pptx

ORACLE_JEWETT_00005427_Native_WorkforceCompHRTrainingFocal.pptx

Documents from Production 22:

- ORACLE_JEWETT_00030956_Native_Taleo and iRecruitment-How to Create Offer.ppt
- ORACLE_JEWETT_00030957_Hiring Process_iRecruitment 2009.pdf
- ORACLE_JEWETT_00030967_Native_Hiring Manager Recruiting Process start to finish.pptx
- ORACLE_JEWETT_00030968_RPM Training Manual.pdf
- ORACLE_JEWETT_00030994_Native_MODULE6_RecruitHire_HowToCreate an Offer_iRecruit_052017.ppt
- ORACLE_JEWETT_00030995_Candidate Screen Form_US_Canada_July2018.pdf
- ORACLE_JEWETT_00030997_Compensation Collection Tool-User Guidelines.pdf
- ORACLE_JEWETT_00031007_Compensation Collection Tool Changes.pdf
- ORACLE_JEWETT_00031008_COMPENSATION COLLECTION TOOL_User Manual.pdf
- ORACLE_JEWETT_00031016_Native_Global Rehire GuidelineTraining June 2018 V3.pptx
- ORACLE_JEWETT_00031017_Role Checklist for a Successful Hire-New US Legislation.pdf
- ORACLE_JEWETT_00031472_August_2017_.pdf

- ORACLE_JEWETT_00031475_Irec_.pdf
- ORACLE_JEWETT_00031478_iRec2_.pdf
- ORACLE_JEWETT_00031480_Old_ScreenShot.pdf

ORACLE_JEWETT_00068776_Amended_and Restated_2000_LTIP_2.1.18.pdf

ORACLE_JEWETT_00068811_Amended_and Restated_2000_LTIP_5.31.11.pdf

ORACLE_JEWETT_00068853_Amended_and Restated_2000_LTIP_6.14.14.pdf

ORACLE_JEWETT_00068897_Amended_and Restated_2000_LTIP_6.30.16.pdf

ORACLE_JEWETT_00068935_Amended_and Restated_2000_LTIP_12.1.17.pdf

Equity Choice FAQ (WANG_00001)

IV. Opposing Expert Reports

Declaration of David Neumark in Support of Plaintiffs' Motion for Class Certification, January 16, 2019

Expert Report of David Neumark in the Matter of Jewett et al. v. Oracle America, Inc. January 2019

- NEUMARK00001-NEUMARK00112: Backup production files
- NEUMARK00113-NEUMARK00116
- Neumark Cross-Reference.xlsx

Declaration of Leaetta M. Hough, Ph.D., in Support of Representative Plaintiffs' Motion for Class Certification, January 15, 2019

Expert Report of Leaetta M. Hough, Ph.D., In the Matter of: Jewett, Wang, and Murray on behalf of themselves, and Petersen, Clark, & Kant, on behalf of themselves and a proposed class v. Oracle America, Inc., January 15, 2019

V. Data Correspondence

2018.04.13 [Oracle] Mantoan ltr to Finberg re Initial Data Production (ORA_JEWETT_007).pdf

2018.04.25 [Jewett] [Finberg] Email to [Oracle] Mantoan w Data Questions Nos 1-6.pdf

2018.04.25 [Oracle] Mantoan Email to [Jewett] Finberg re Questions re Data.pdf

2018.05.01 [Oracle] Mantoan Ltr to [Jewett] re 1st Suppl Data Production.pdf

2018.05.02 [Jewett] [Finberg] Email to [Oracle] Mantoan Another Data Question.pdf

2018.05.03 [Jewett] [Finberg] Email to [Oracle] Mantoan w add'l Data Questions.pdf

2018.05.04 [Jewett] [Finberg] Email to [Oracle] Mantoan w Data Question No 10.pdf

2018.05.04 [Oracle] [Mantoan] Email to [Jewett] Finberg re add'l Data Questions Nos 7-9.pdf

2018.05.07 [Jewett] [Finberg] Email to [Oracle] Mantoan w add'l data questions nos. 11-17.pdf

2018.05.31 [Finberg] Email to [Oracle] Mantoan re Data Questions.pdf

2018.05.31 [Finberg] Email to [Oracle] Mantoan-ATTACHMENT-Questions for Attys 5_31_18..xlsx

2018.05.31 [Jewett] [Finberg] Email to [Oracle] Mantoan re Additional Data Questions.pdf

2018.05.31 [Jewett] Finberg Email to [Oracle] Mantoan re Add'l Data Questions Nos 18-19.pdf

2018.06.01 [Jewett] [Finberg] Email to [Oracle] Mantoan w add'l Data Questions Nos. 20-21.pdf

2018.06.06 [Oracle] Mantoan Ltr to [Jewett] Finberg resp to data Qs Nos. 1-17.pdf

2018.06.15 [Jewett] [Finberg] Email to [Oracle] Mantoan w add'l Qs Oracle data.pdf

2018.06.21 [Jewett] [Finberg] Email to [Oracle] Mantoan Re Add'l Data Qs.pdf

2018.06.29 [Oracle] Mantoan ltr to Finberg re responses to Data Qs (No. 22).pdf

2018.07.01 [Jewett] [Finberg] Email to [Oracle] Mantoan Re Data Question Priority.pdf

2018.07.06 [Oracle] Mantoan ltr to [Jewett] Finberg Continuing Resps to Data Qs.pdf

2018.07.17 [Jewett] Finberg email to [Oracle] Mantoan re Data Qs Nos. 24, 25, 26, 27.pdf

2018.08.03 [Jewett] Finberg email to Mantoan re data questions.pdf

2018.08.03 [Oracle] [Mantoan] Email to [Jewett] Finberg re Remaining Data Qs.pdf

2018.08.03 [Oracle] Mantoan ltr to [Jewett] Finberg continued resp to Data Qs.pdf

2018.08.13 [Jewett] [Finberg] Email to [Oracle] Mantoan re Data Questions.pdf

2018.08.13 [Jewett] [Finberg] Email to [Oracle] Mantoan-List of Oracle Questions Pending as of 8.10.18.docx

2018.08.14 [Oracle] [Mantoan] Email to [Jewett] Finberg, et al re Data Questions.pdf

2018.08.17 [Oracle] Mantoan ltr to [Jewett] Finberg resp to data Qs 21, 22, 26, prod vol 17.pdf

2018.09.04 [Jewett] Finberg email to [Oracle] Mantoan re Follow up on Answers to Data Questions.pdf

2018.09.04 [Jewett] Finberg email to [Oracle] Mantoan-List of Oracle questions for stocks and bonuses 9.4.18.docx

2018.09.04 [Oracle] Mantoan ltr to [Jewett] Mullan re Oracle prod vol 18.pdf

2018.09.07 [Oracle] Mantoan ltr to [Jewett] Finberg re Oracle Production Vol 19.pdf

2018.09.12 [Oracle] Mantoan Email to [Jewett] Finberg in resp to data Qs.pdf

2018.09.14 [Oracle] Mantoan ltr to [Jewett] Finberg re Oracle prod vol 20.pdf

2018.09.18 [Oracle] [Mantoan] Ltr to [Jewett] Finberg re Continuing Resp....pdf

2018.09.21 [Oracle] Mantoan ltr to [Jewett] Finberg re Oracle prod vol 21.pdf

2018.10.05 [Oracle] Mantoan ltr to [Jewett] Finberg re Qs re Oracle recent prods.pdf

2018.10.05 [Oracle] Mantoan ltr to [Jewett] Finberg resp to remaining data Qs.pdf

2018.10.10 [Jewett] [Finberg] Ltr to [Oracle] Mantoan re Resp to 10.05.2.._.pdf

2018.10.12 [Oracle] Mantoan ltr to [Jewett] Finberg re Oracle prod vols 24 and 25.pdf

[Oracle-Jewett] 2018.10.25 Ltr to Finberg from Mantoan re 2018.10.11 Dat....pdf

2018.10.26 [Oracle] Mantoan Ltr to [Jewett] Finberg re Suppl Data Prod Vol 28.pdf

2018.10.31 [Oracle] Mantoan Ltr to [Jewett] Finberg re Suppl Data Prod Vol 29.pdf

2018.11.08 [Oracle] Mantoan Ltr to [Jewett] Finberg re Suppl Data Prod Vol 30.pdf

VI. Data

ORACLE_JEWETT_00001110_Jewett- Posting Job Descriptions & Requirements.xlsx

ORACLE_JEWETT_00001166_Jewett_Merged Assignment History, Medicare and Sal Admin.xlsx

ORACLE_JEWETT_00001167_Jewett_AllEarnings.xlsx

ORACLE_JEWETT_00001168_jewett_adhoc_comp_total.xlsx

ORACLE_JEWETT_00001169_jewett_adhoc_comp_wf.xlsx

ORACLE_JEWETT_00001170_hcm_jewett_wfc_audit.xlsx

ORACLE_JEWETT_00001171_gsi_jewett_bonus_audit.xlsx

ORACLE_JEWETT_00001172_gsi_jewett_focal_audit.xlsx

ORACLE_JEWETT_00001173_gsi_jewett_stock_audit.xlsx

ORACLE_JEWETT_00001179_Salary Range History.xlsx

ORACLE_JEWETT_00001180_Jewett_Emp_Personal_Experience_Qualification_Assign_Details.xlsx

ORACLE_JEWETT_00007169_ Data Dictionary -ORG_LOB_HIERARCHY_DATA Codes.xlsx

ORACLE_JEWETT_00007170_ Jewett - Consolidated Location Codes.xlsx

ORACLE_JEWETT_00007171-00007289 (Production 18: Organization Hierarchy Files)

ORACLE_JEWETT_00007290_Jewett_Application_CandidateSkills.xlsx

ORACLE_JEWETT_00007291_Jewett_Application_CSWHistory.xlsx

ORACLE_JEWETT_00007292_Jewett_Application_Data.xlsx

ORACLE_JEWETT_00007293_Jewett_Application_Education.xlsx

ORACLE_JEWETT_00007294_Jewett_Application_Experience.xlsx

ORACLE_JEWETT_00007295_Jewett_Application_History.xlsx

ORACLE_JEWETT_00007296_Jewett_Application_Source.xlsx

ORACLE_JEWETT_00007297_Jewett_Candidate_Demographics.xlsx

ORACLE_JEWETT_00007298_Jewett_Candidate_Languages.xlsx

ORACLE_JEWETT_00007299_Jewett_Candidate_Preferences_Job_Field.xlsx

ORACLE_JEWETT_00007300_Jewett_Candidate_Preferences_Location.xlsx

ORACLE_JEWETT_00007301_Jewett_Candidate_Preferences_Organization.xlsx

ORACLE_JEWETT_00007302_Jewett_Candidate_Referrals.xlsx

ORACLE_JEWETT_00007303_Jewett_CWB_details.xlsx

- Production 28: ORACLE_JEWETT_00062012 - ORACLE_JEWETT_00065615 (963 files referenced in ORACLE_JEWETT_00007303)
- Production 29: ORACLE_JEWETT_00065620 - ORACLE_JEWETT_00065627 (2 files referenced in ORACLE_JEWETT_00007303)

ORACLE_JEWETT_00007304_JEWETT_IREC_Data.xls

- Production 24: ORACLE_JEWETT_00031755 - ORACLE_JEWETT_00059677 (10,100 files referenced in ORACLE_JEWETT_00007304)

ORACLE_JEWETT_00007305_Jewett_Requisition_Collaborators_Data.xlsx

ORACLE_JEWETT_00007306_Jewett_Requisition_Data.xlsx

ORACLE_JEWETT_00007307_Jewett_Requisition_Description_and_Qualification_Data.xlsx

ORACLE_JEWETT_00007308_Jewett_Requisition_Other_Locations.xlsx

ORACLE_JEWETT_00007309_Jewett_SalesFocal_AUDIT.xlsx

ORACLE_JEWETT_00007310_Jewett_Txn_Report_R12_sets.xlsx

ORACLE_JEWETT_00007311_Jewett_Txn_Report_R13sets.xlsx

ORACLE_JEWETT_00007312_Jewett_hcm_wfc_details.xlsx

- Production 25: ORACLE_JEWETT_00059678 - ORACLE_JEWETT_00060152 (151 files referenced in ORACLE_JEWETT_00007312)

ORACLE_JEWETT_00007313_Jewett_File_Attachments_By_Candidate.xlsx

ORACLE_JEWETT_00007314_Jewett_File_Attachments_By_Requisition.xlsx

- Production 20: ORACLE_JEWETT_0000007316 - ORACLE_JEWETT_00030953 (8,684 files referenced in ORACLE_JEWETT_00007313 and ORACLE_JEWETT_00007314)

ORACLE_JEWETT_00030954_Jewett_All_AppraisalData.xlsx

ORACLE_JEWETT_00030955_Jewett_gsi_comp_history.xlsx

ORACLE_JEWETT_00060153_Jewett_Talent_Review_Audit.xlsx

ORACLE_JEWETT_00065616_Jewett_supplement_hcm_wfc_audit.xlsx

ORACLE_JEWETT_00065617_Jewett_supplement1_hcm_wfc_details.xlsx

ORACLE_JEWETT_00065618_Jewett_supplement2_hcm_wfc_details.xlsx

ORACLE_JEWETT_00065619_Jewett_Talent_Review_Notes.xlsx

ORACLE_JEWETT_00066965_Jewett_adhoc_comp_attachments.xlsx

- Production 29: ORACLE_JEWETT_0065628 - ORACLE_JEWETT_00068775 (760 files referenced in ORACLE_JEWETT_00066965)

ORACLE_JEWETT_00066966_Jewett_Talent_Profile_Data.xlsx

ORACLE_JEWETT_00068971_Jewett_Stock_Data_Product_Statement_Combined.xlsx

ORACLE_JEWETT_00068972_ Jewett_Supplement_Talent_Profile_Data.xlsx

ORACLE_JEWETT_00076871_ Oracle Budget 1_Highly Confidential Attorneys Eyes Only Information.xlsx

ORACLE_JEWETT_00076872_ Oracle Budget 2_Highly Confidential Attorneys Eyes Only Information.xlsx

ORA_JEWETT_020.dat

ORA_JEWETT_024.dat

ORA_JEWETT_025.dat

ORA_JEWETT_028.dat

ORA_JEWETT_029_PDF.dat

VII. Literature

Baker, Scott R., Nicholas Bloom and Steven J. Davis (2016), Measuring Economic Policy Uncertainty. *The Quarterly Journal of Economics*, 131(4), pp. 1593-1636.

Bertrand, Marianne, Claudia Goldin, and Lawrence F. Katz. (2010). Dynamics of the gender gap for young professionals in the financial and corporate sectors. *American Economic Journal: Applied Economics*, 2(3), pp. 228-55.

Born, Benjamin, Michael Ehrmann and Marcel Fratzscher. (2013) Central Bank Communication on Financial Stability. *The Economic Journal*, 124, pp. 701–734.

Cahuc, Pierre and Andre Zylberberg, *Labor Economics*, Cambridge: The MIT Press, 2001, p. 175.

Chae, David H., Sean Clouston, Mark L. Hatzenbuehler, Michael R. Kramer, Hannah LF Cooper, Sacoby M. Wilson, Seth I. Stephens-Davidowitz, Robert S. Gold, and Bruce G. Link. (2015) Association between an Internet-Based Measure of Area Racism and Black Mortality. *PLoS ONE* 10(4):e0122963.

Cotton, Jeremiah (1988). On the decomposition of wage differentials. *The Review of Economics and Statistics*, 236-243.

Einav, Liran and Jonathan D. Levin (2014) The Data Revolution and Economic Analysis. *Innovation Policy and the Economy*, 14, pp. 1-24.

Gentzkow, Matthew and Jesse M. Shapiro (2010), What Drives Media Slant? Evidence from U.S. Daily Newspapers. *Econometrica*, 78(1), 35-71.

Gentzkow, Matthew, Bryan T. Kelly and Matt Taddy. (Forthcoming) Text as Data. *Journal of Economic Literature.*

Gentzkow, Matthew, Jesse M. Shapiro and Michael Sinkinson (2014) Competition and Ideological Diversity: Historical Evidence from US Newspapers. *American Economic Review*, 104(10): 3073-3114.

Greene, William H. (1993) *Econometric Analysis*, 2nd Edition, NY: Macmillan Publishing Company, pp. 245-248.

Griliches, Zvi. (1977) Estimating the Returns to Schooling: Some Econometric Problems. *Econometrica*, 45(1), pp.1-22.

Groseclose, Tim and Jeffrey Milyo, A Measure of Media Bias. *The Quarterly Journal of Economics*, 120(4), pp. 1191-1237.

Hellerstein, Judith K., and David Neumark (2006). Using Matched Employer-Employee Data to Study Labor Market Discrimination. In *Handbook on the Economics of Discrimination*, ed. William Rodgers (pp. 29–60). Cheltenham, UK: Edgar Elgar.

Hellerstein, Judith K., David Neumark, and Kenneth Troske, 1999, "Wages, Productivity, and Worker Characteristics: Evidence from Plant-Level Production Functions and Wage Equations," *Journal of Labor Economics*, 17(3), pp. 409-446

Leamer, Edward, *Specification Searches: Ad Hoc Inferences with Non-Experimental Data*, New York, Wiley, 1978.

Litecky, Chuck, et al., (January/February 2010), Mining for Computing Jobs, IEEE Software.

McCloskey, Donald N. (1985). The Loss Function Has Been Mislaid: The Rhetoric of Significance Tests. *American Economic Review Papers and Proceedings*, 75(2), pp. 201-205.

Mincer, Jacob. *Schooling, Experience and Earnings*, National Bureau of Economic Research, 1974. pp.83-96.

Mincer, Jacob, and Solomon Polachek (1978). An Exchange: The Theory of Human Capital and the Earnings of Women: Women's Earnings Reexamined. *Journal of Human Resources*, 13(1), pp.118-134.

Neumark, David (1998). Employers' Discriminatory Behavior and the Estimation of Wage Discrimination. *Journal of Human Resources,* 279-295.

Piette, Michael J. and Paul F. White (1999). Approaches for Dealing with Small Sample Sizes in Employment Discrimination Litigation. *Journal of Forensic Economics*, 12(1), pp. 43-56.

Romer, Christina D. and David H. Romer (2017) New Evidence on the Aftermath of Financial Crises in Advanced Countries. *American Economic Review*, 107(10), pp. 3072-3118.

Rubenfeld, Daniel, Reference Guide on Multiple Regression, *Reference Manual on Scientific Evidence: Third Edition* (p.318), Washington D.C.: The National Academies Press.

Saiz, Albert and Uri Simonsohn (2013) Proxying for Unobserved Variables with Internet Document-Frequency. *Journal of the European Economic Association*, 11(1), pp. 137-165.

Taddy, Matt. (2013) Measuring Political Sentiment on Twitter: Factor Optimal Design for Multinomial Inverse Regression. *Technometrics*, 55(4), Special Issue (November 2013), pp. 415-425.

Taddy, Matt. (2015) Distributed Multinomial Regression. *The Annals of Applied Statistics*, 9(3), pp. 1394-1414.

Willis, Robert. (1986) Wage Determinants: A Survey and Reinterpretation of Human Capital Earnings Functions, Chapter 10, *Handbook of Labor Economics, Volume 1*, Edited by Orley Ashenfelter and Richard Layard. Elsevier Science Publishers BV.

Woweczko, Izabella A. (2015) Skills and Vacancy Analysis with Data Mining Techniques, *Informatics*, 2, pp. 31-49.

VIII. Other Data and Online Cites

California Equal Pay Act, Labor Code section 1197.5 (https://www.dir.ca.gov/dlse/California_Equal_Pay_Act.htm), retrieved on February 26, 2019.

EDD State of California - OES Employment and Wages (https://www.labormarketinfo.edd.ca.gov/data/oes-employment-and-wages.html#OES), retrieved on February 26, 2019.

Linkedin: Thomas Kurian CEO at Google Cloud (https://www.linkedin.com/in/thomas-kurian-469b6219/), retrieved March 1, 2019.

Oracle Fact Sheet (http://www.oracle.com/us/corporate/oracle-fact-sheet-079219.pdf), retrieved on February 26, 2019.

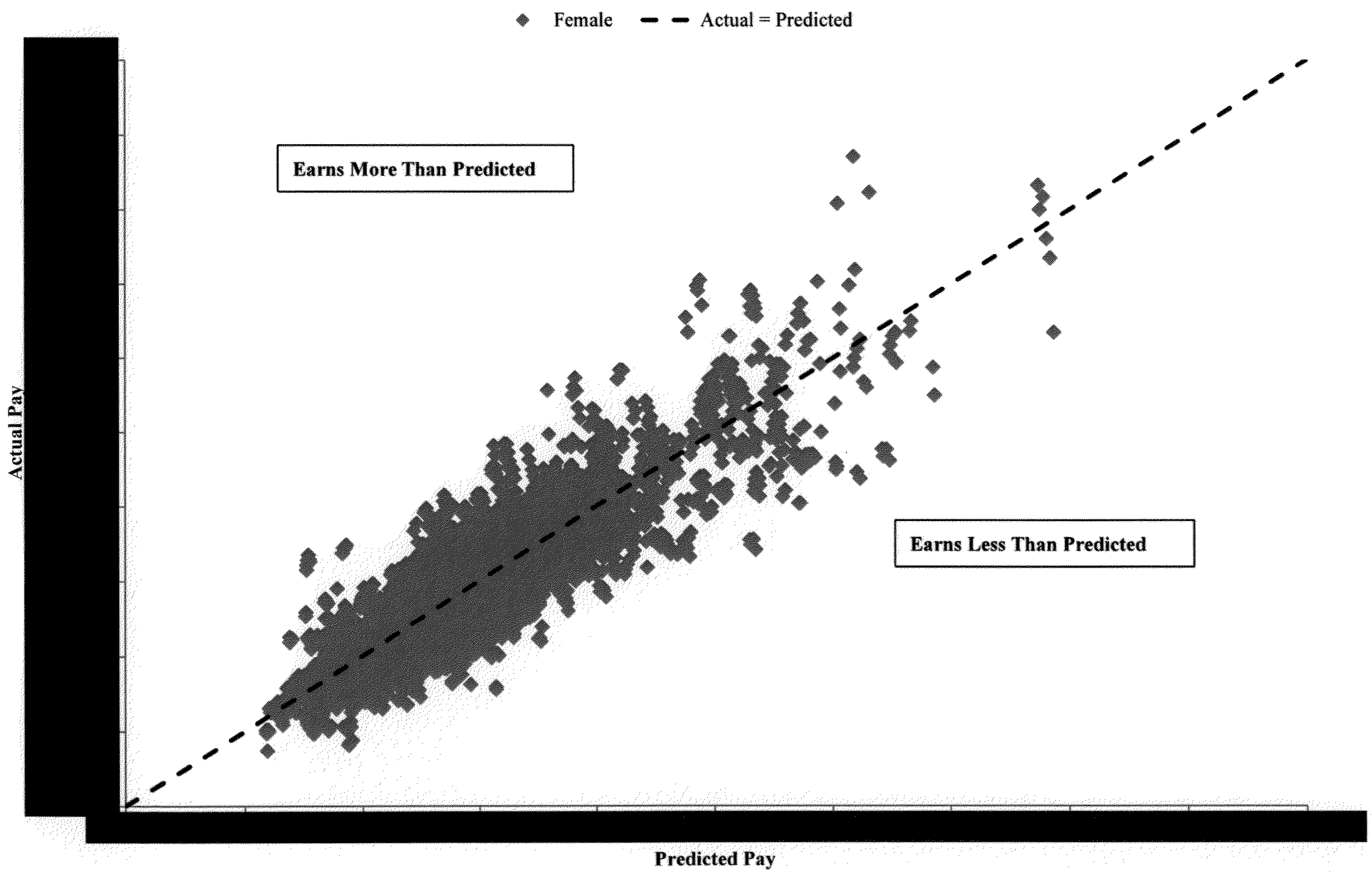
United States Census Bureau: Industry and Occupation (https://www.census.gov/topics/employment/industry-occupation/about/occupation.html), retrieved February 28, 2019.

1997 National Longitudinal Survey (NLS97)

Attachment C: Variability

Exhibit C1

PRODEV: Actual Base Pay vs. Predicted Base Pay
- Prediction Based on Dr. Neumark's Data and Model, Without a Gender Control -
- Female Incumbents in Neumark's Dataset, 2013-2018 -



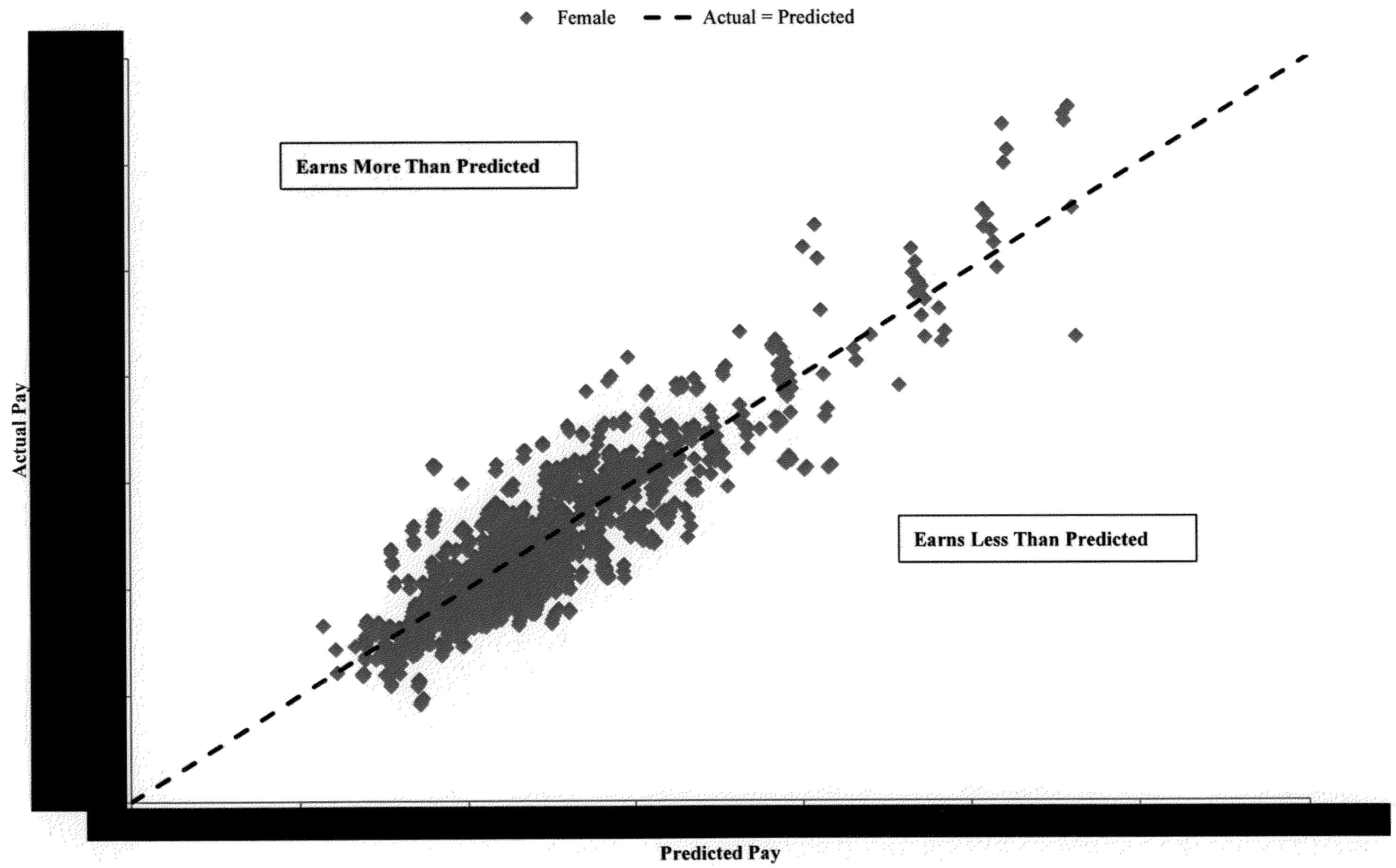
C1

Exhibit C2

INFTECH: Actual Base Pay vs. Predicted Base Pay

- Prediction Based on Dr. Neumark's Data and Model, Without a Gender Control -

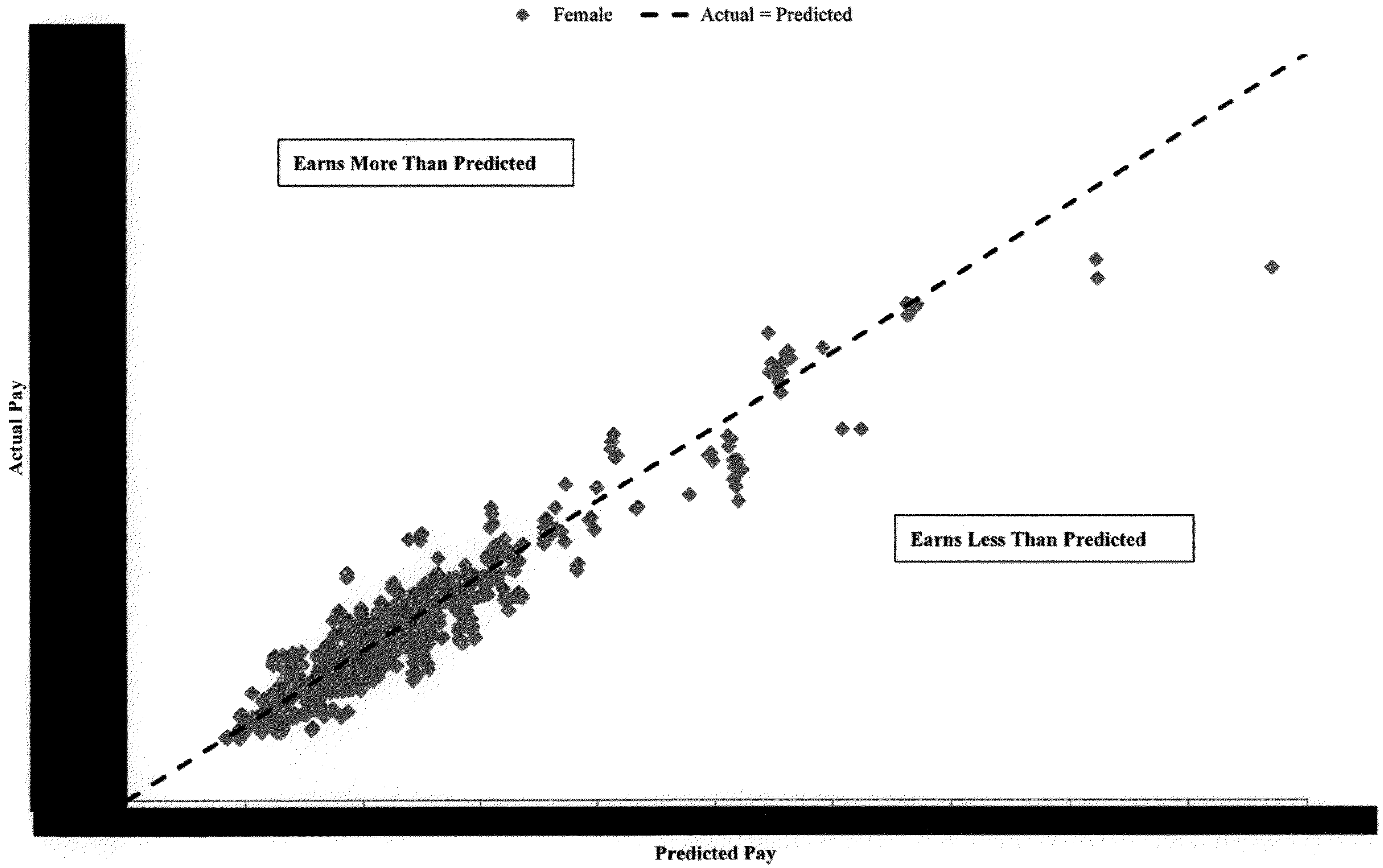
- Female Incumbents in Neumark's Dataset, 2013-2018 -



C2

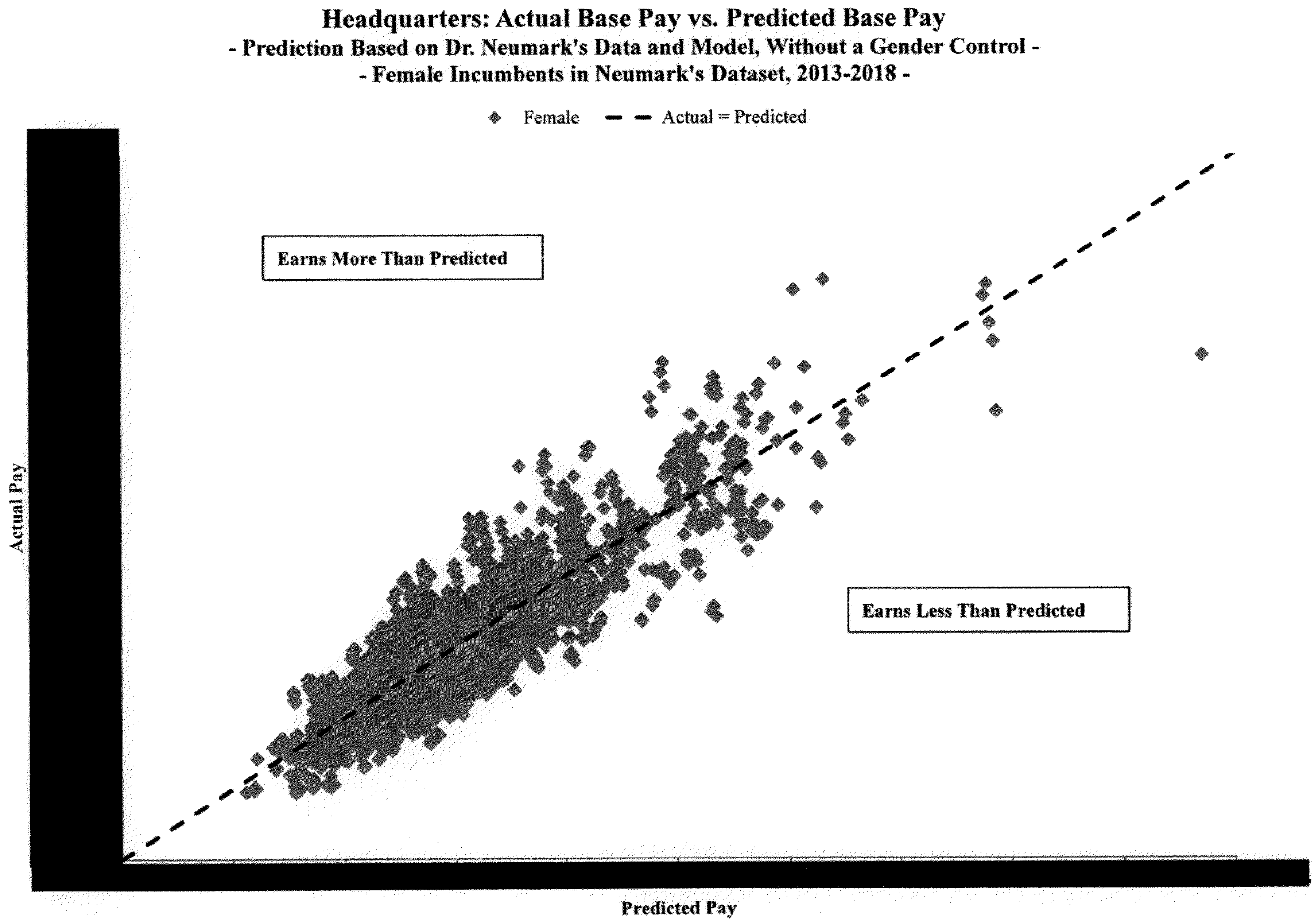
Exhibit C3

SUPP: Actual Base Pay vs. Predicted Base Pay
- Prediction Based on Dr. Neumark's Data and Model, Without a Gender Control -
- Female Incumbents in Neumark's Dataset, 2013-2018 -



C3

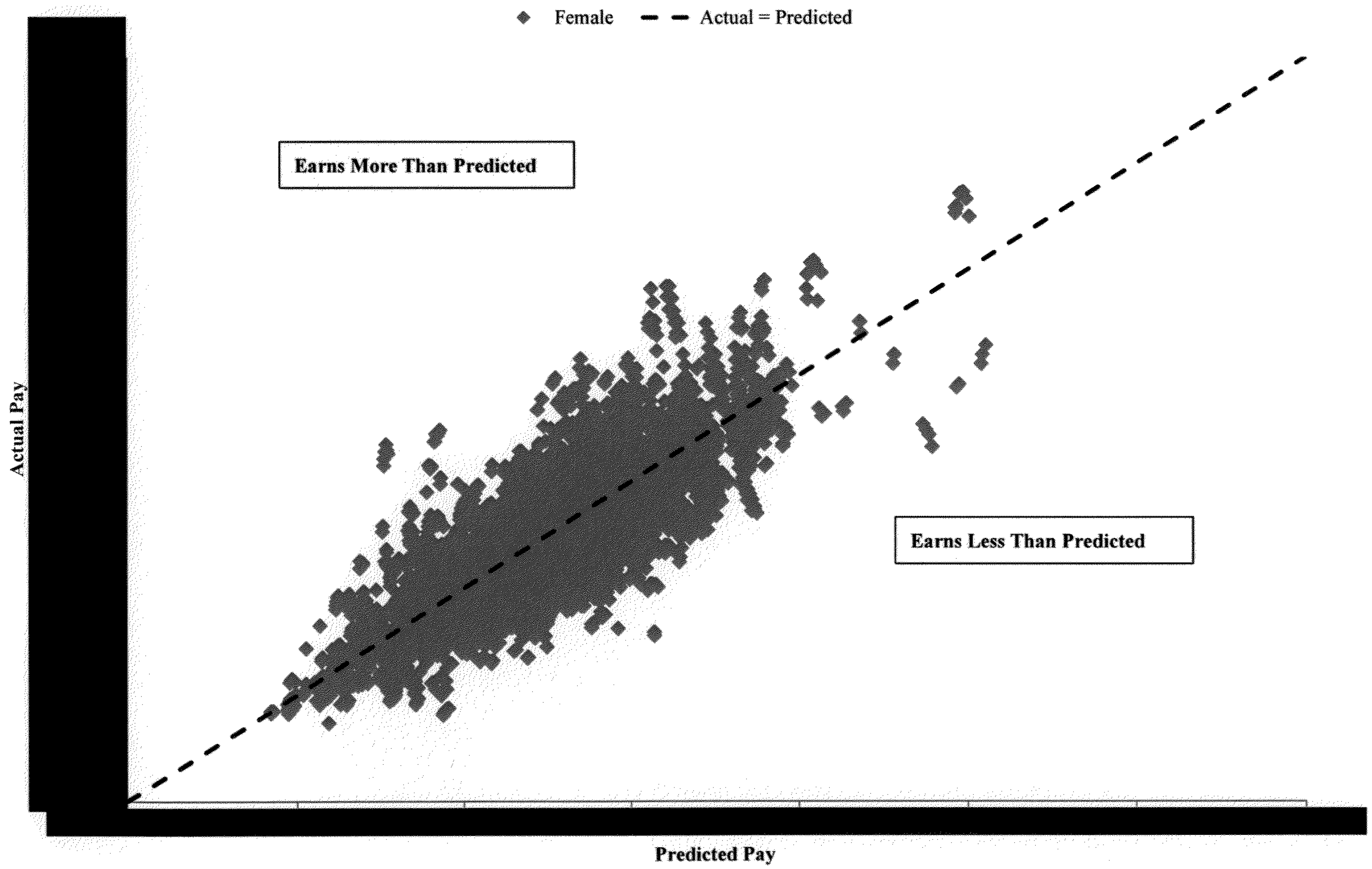
Exhibit C4



C4

Exhibit C5

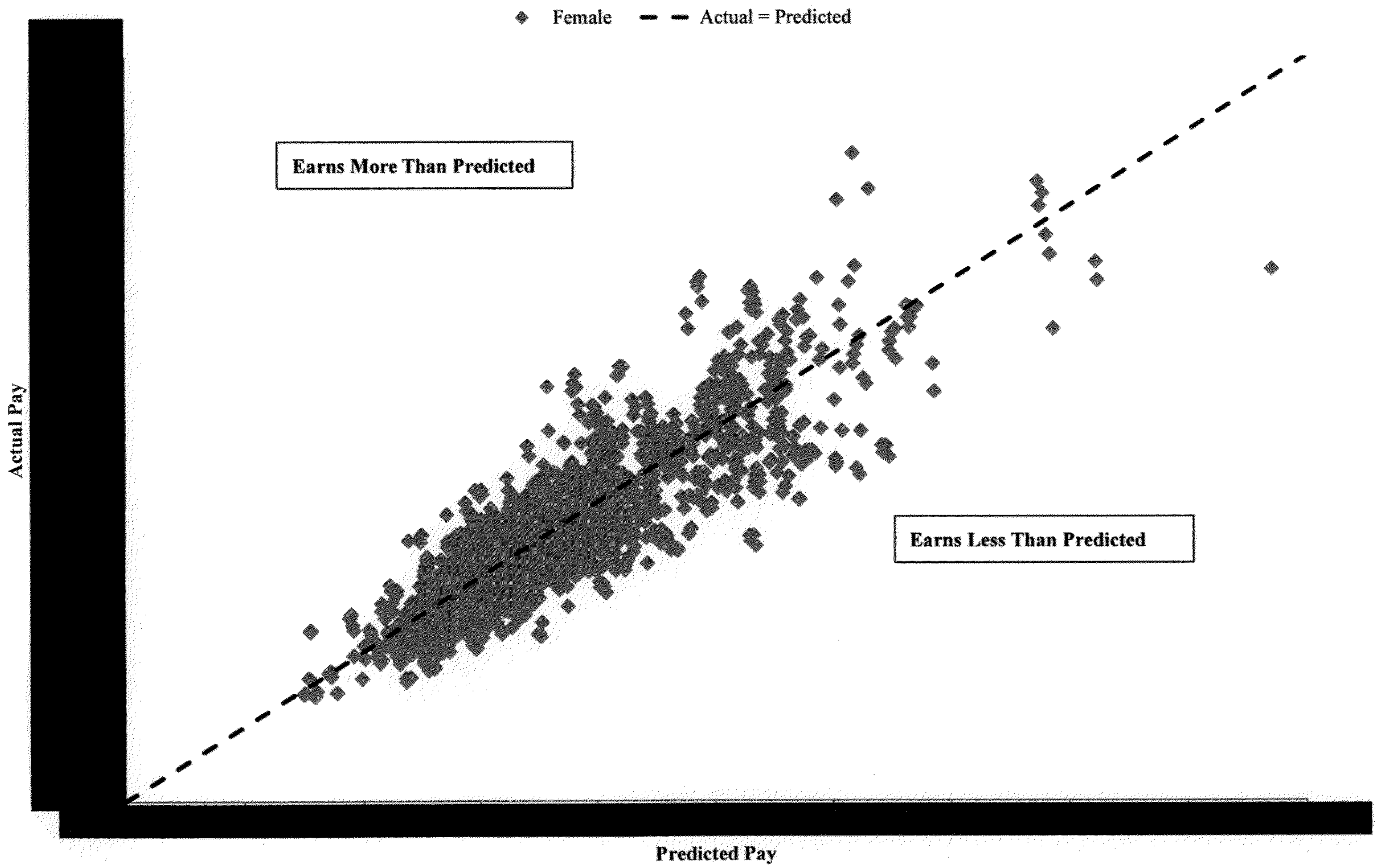
Individual Contributors: Actual Base Pay vs. Predicted Base Pay
- Prediction Based on Dr. Neumark's Data and Model, Without a Gender Control -
- Female Incumbents in Neumark's Dataset, 2013-2018 -



C5

Exhibit C6

Managers: Actual Base Pay vs. Predicted Base Pay
- Prediction Based on Dr. Neumark's Data and Model, Without a Gender Control -
- Female Incumbents in Neumark's Dataset, 2013-2018 -



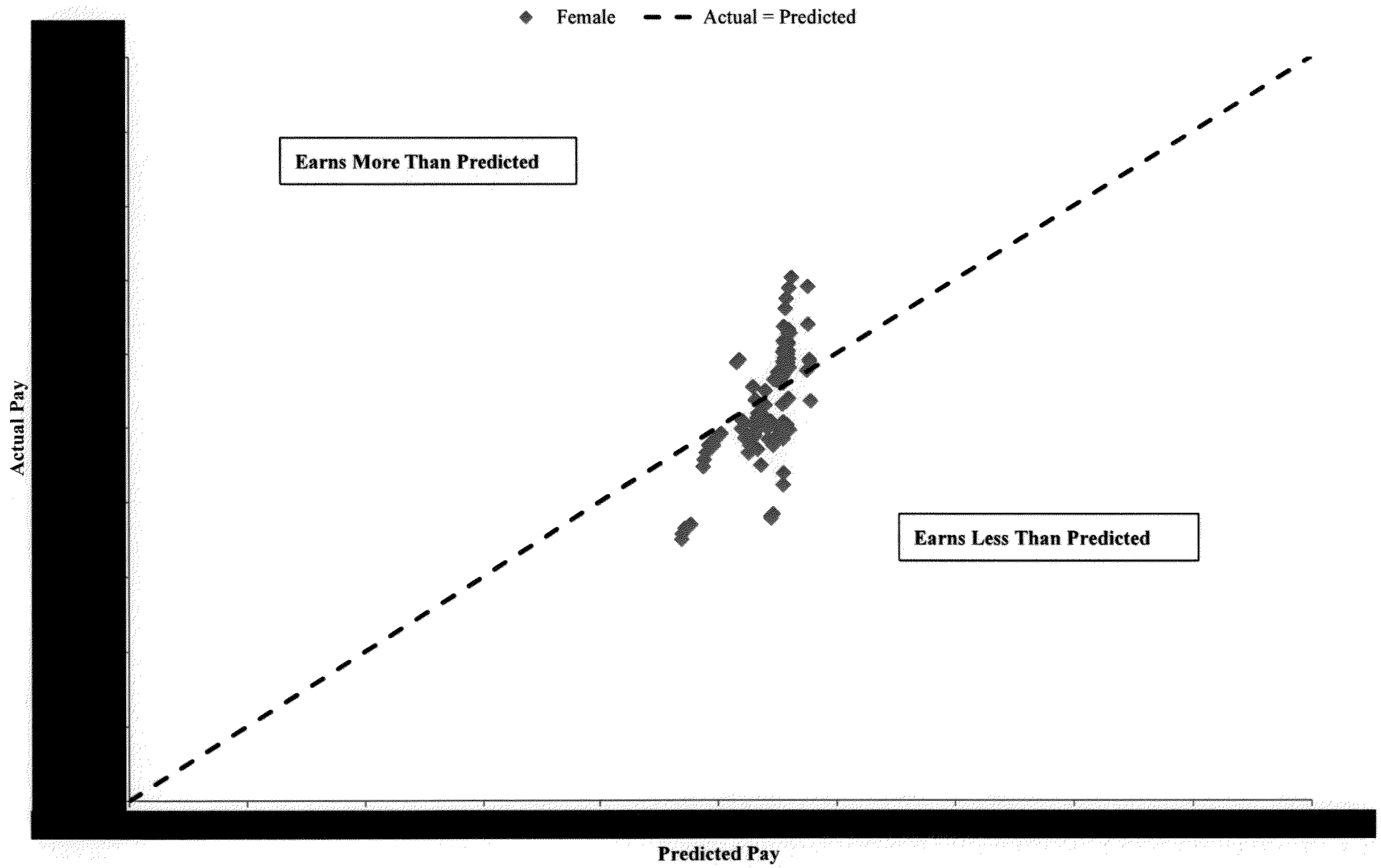
C6

Exhibit C7

Applications Developer 2: Actual Base Pay vs. Predicted Base Pay

- Prediction Based on Dr. Neumark's Data and Model, Without a Gender Control -

- Female Incumbents in Neumark's Dataset, 2013-2018 -



C7

Exhibit C8

Applications Developer 3: Actual Base Pay vs. Predicted Base Pay
- Prediction Based on Dr. Neumark's Data and Model, Without a Gender Control -
- Female Incumbents in Neumark's Dataset, 2013-2018 -

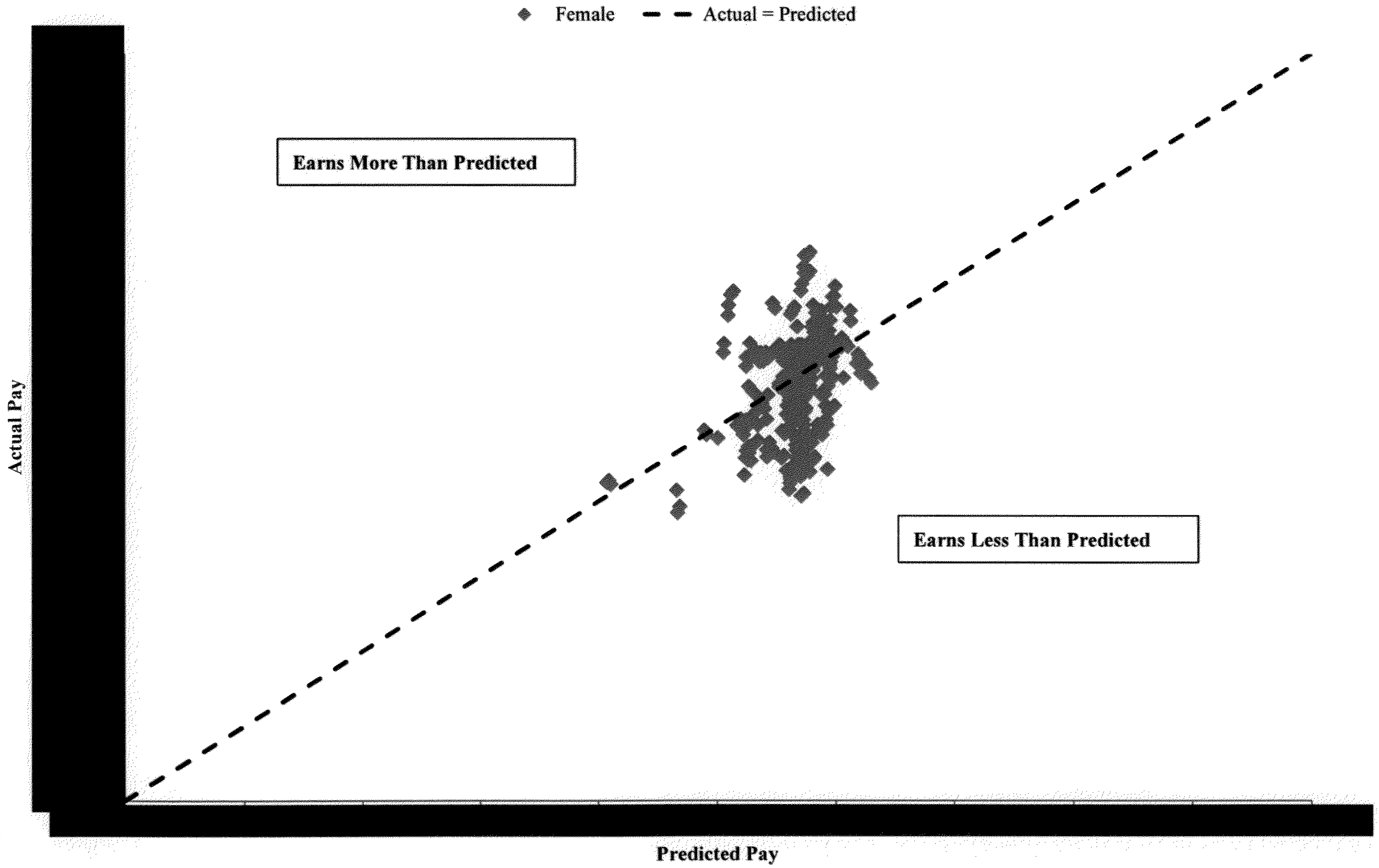
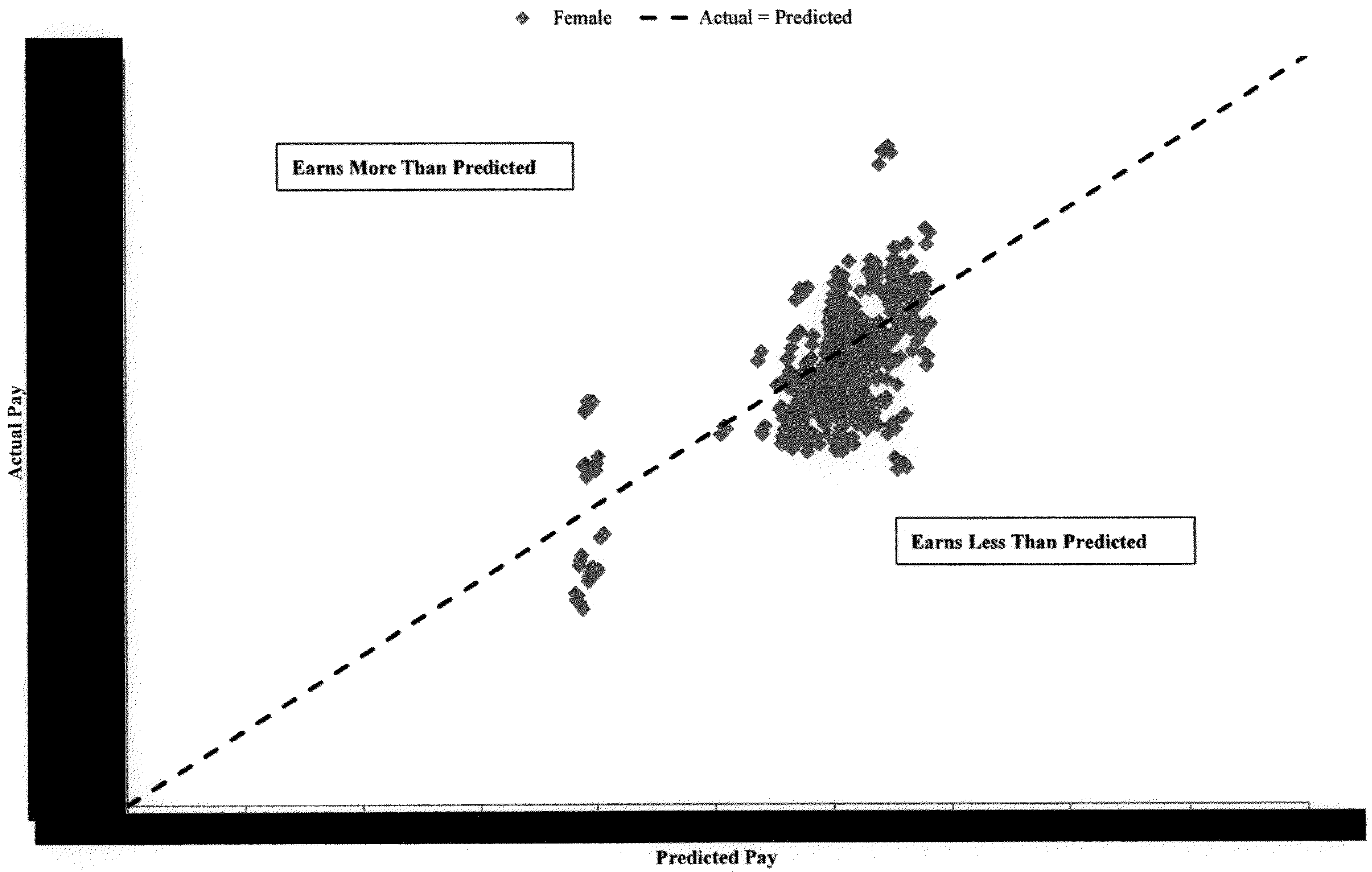


Exhibit C9

Applications Developer 4: Actual Base Pay vs. Predicted Base Pay

- Prediction Based on Dr. Neumark's Data and Model, Without a Gender Control -

- Female Incumbents in Neumark's Dataset, 2013-2018 -



C9

Exhibit C10

Applications Developer 5: Actual Base Pay vs. Predicted Base Pay
- Prediction Based on Dr. Neumark's Data and Model, Without a Gender Control -
- Female Incumbents in Neumark's Dataset, 2013-2018 -

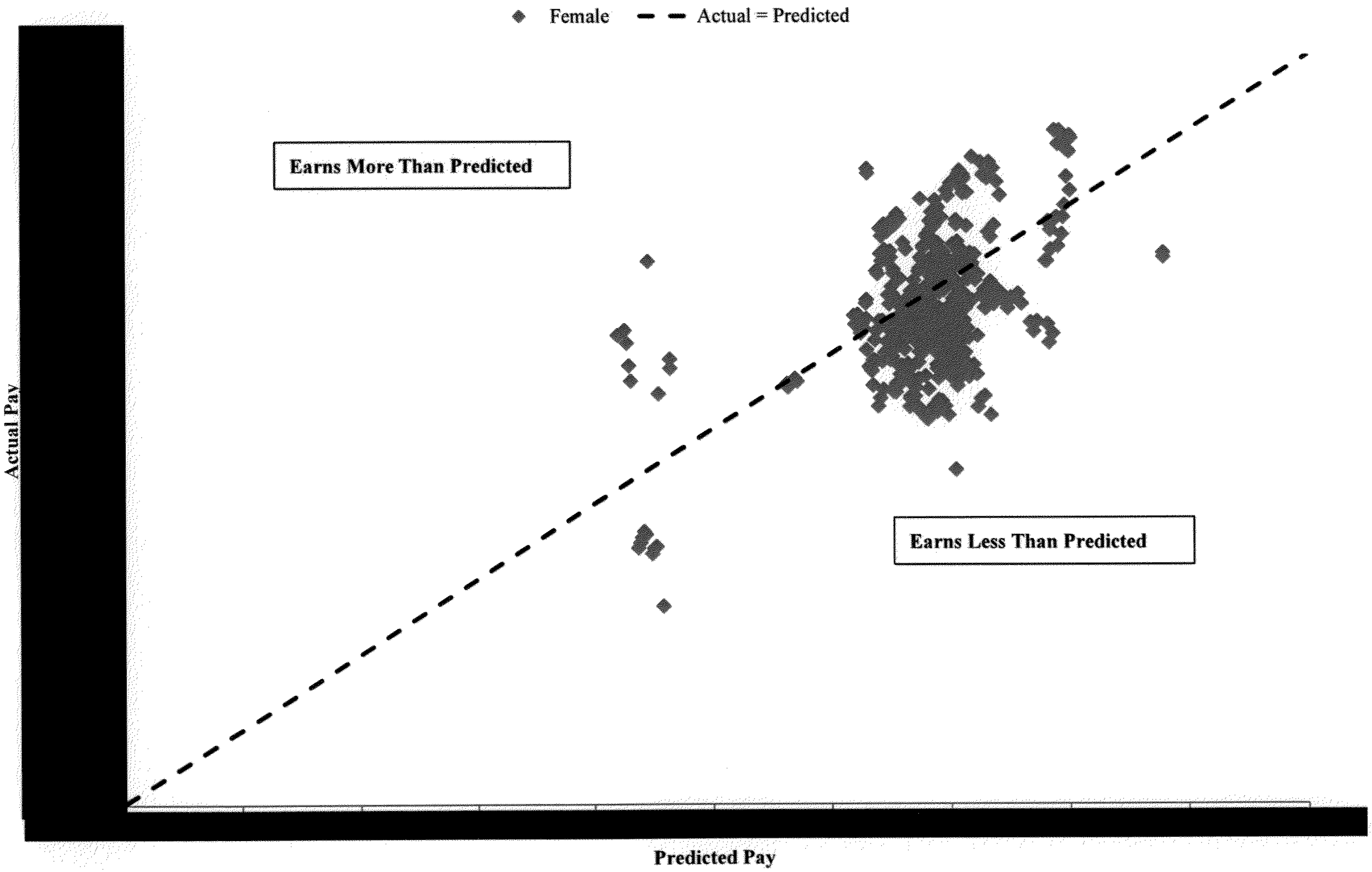


Exhibit C11

Hardware Developer 2: Actual Base Pay vs. Predicted Base Pay
- Prediction Based on Dr. Neumark's Data and Model, Without a Gender Control -
- Female Incumbents in Neumark's Dataset, 2013-2018 -

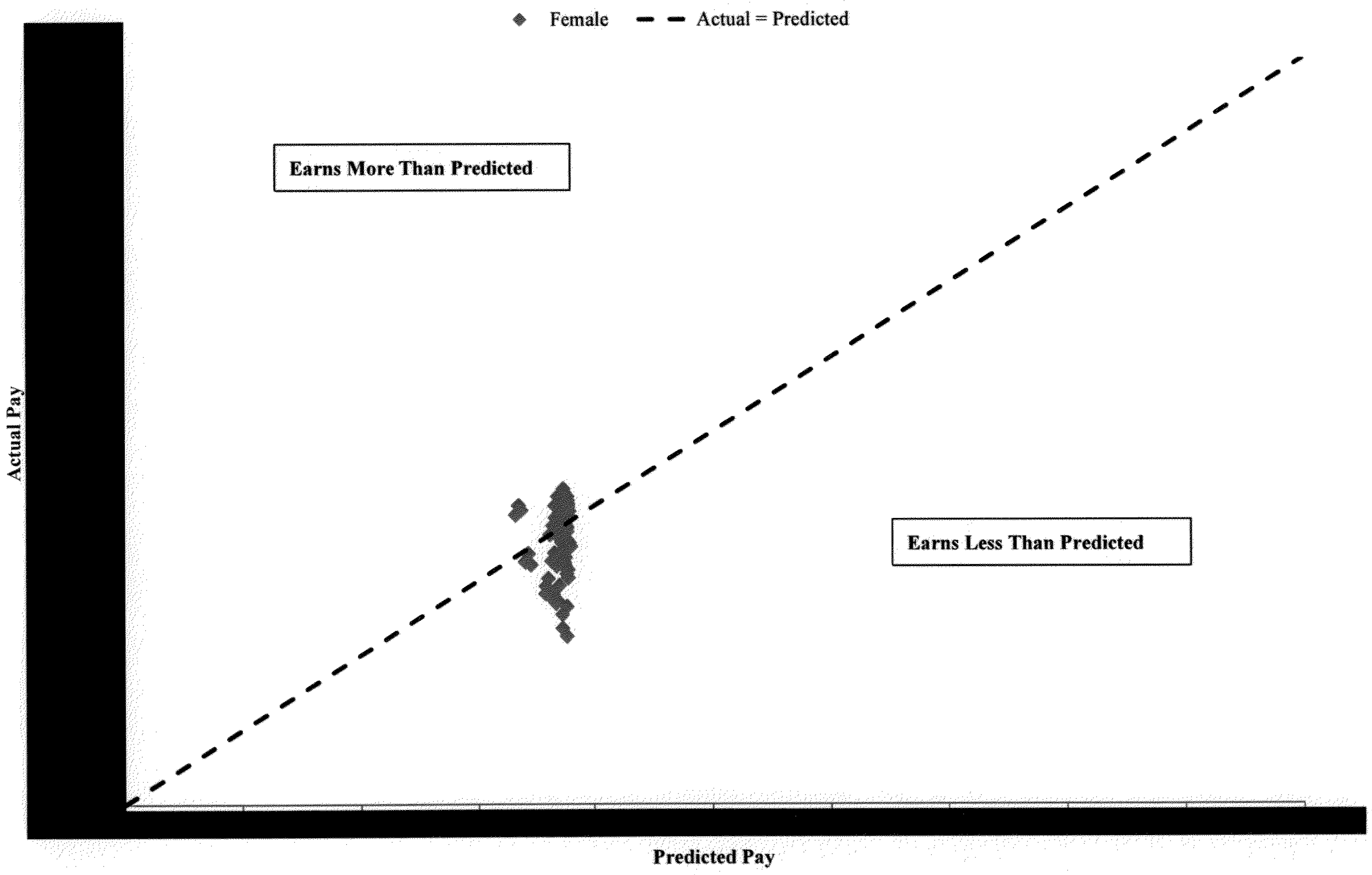


Exhibit C12

Hardware Developer 3: Actual Base Pay vs. Predicted Base Pay

- Prediction Based on Dr. Neumark's Data and Model, Without a Gender Control -

- Female Incumbents in Neumark's Dataset, 2013-2018 -

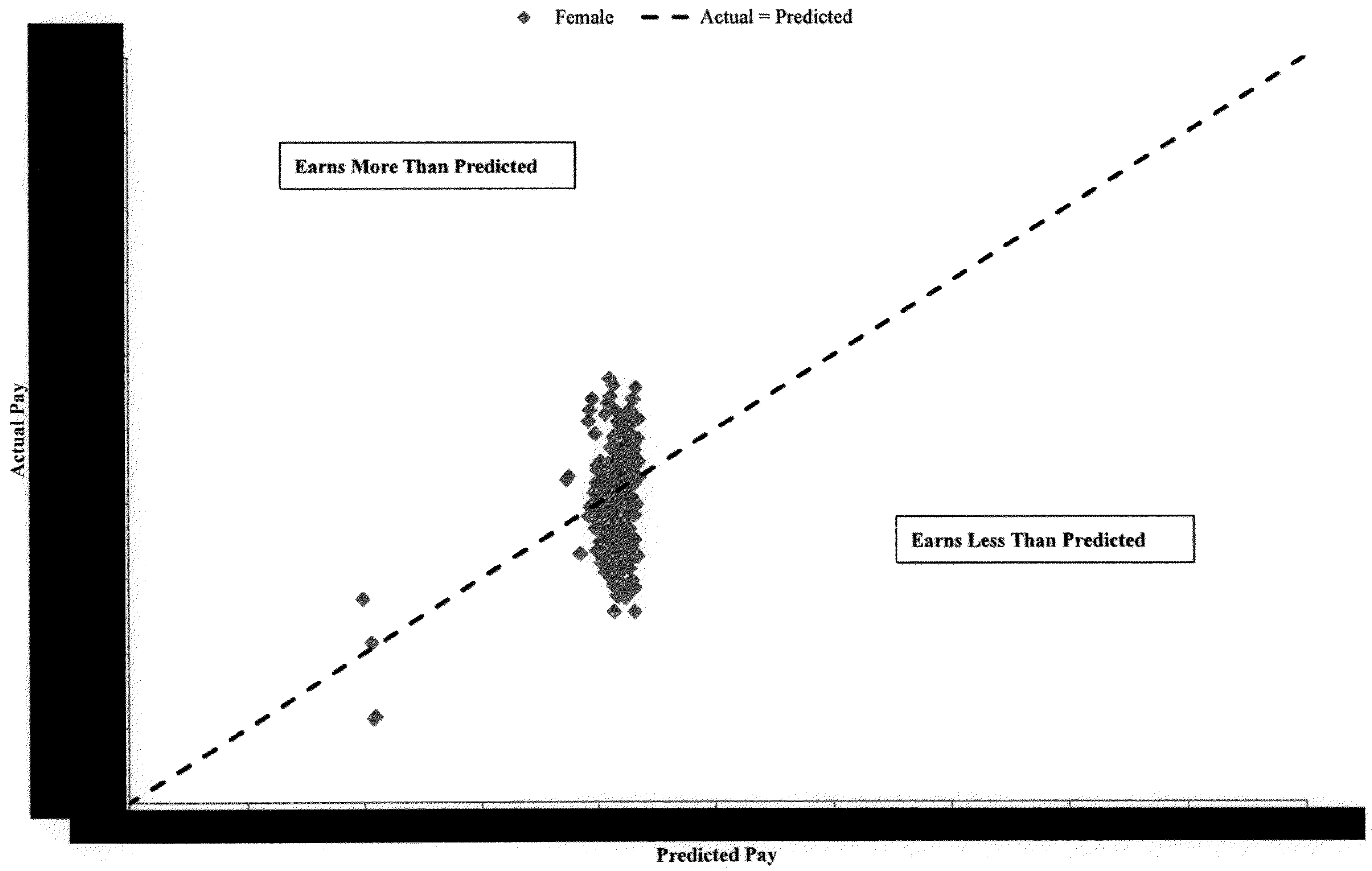
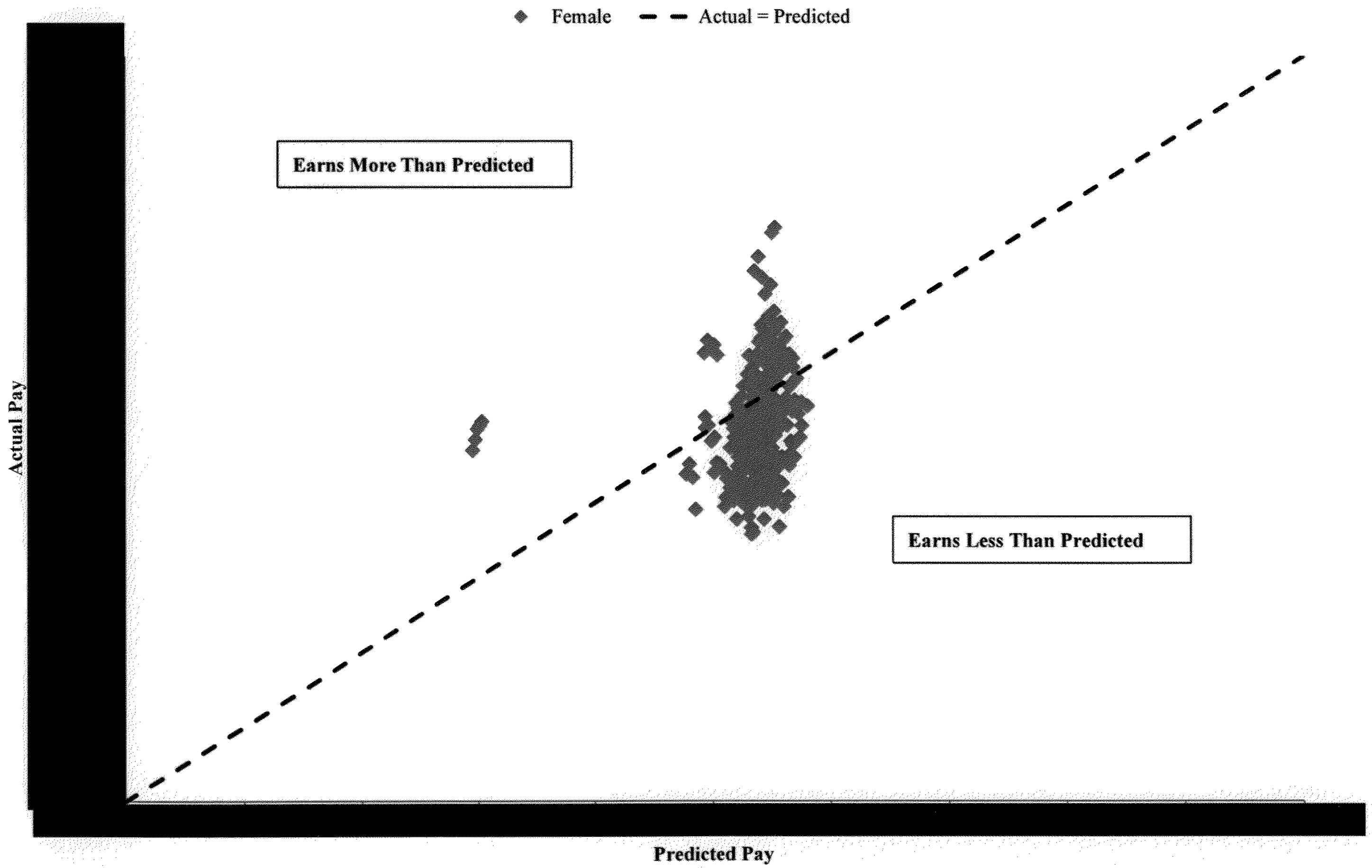


Exhibit C13

Hardware Developer 4: Actual Base Pay vs. Predicted Base Pay
- Prediction Based on Dr. Neumark's Data and Model, Without a Gender Control -
- Female Incumbents in Neumark's Dataset, 2013-2018 -



C13

Exhibit C14

Hardware Developer 5: Actual Base Pay vs. Predicted Base Pay

- Prediction Based on Dr. Neumark's Data and Model, Without a Gender Control -

- Female Incumbents in Neumark's Dataset, 2013-2018 -

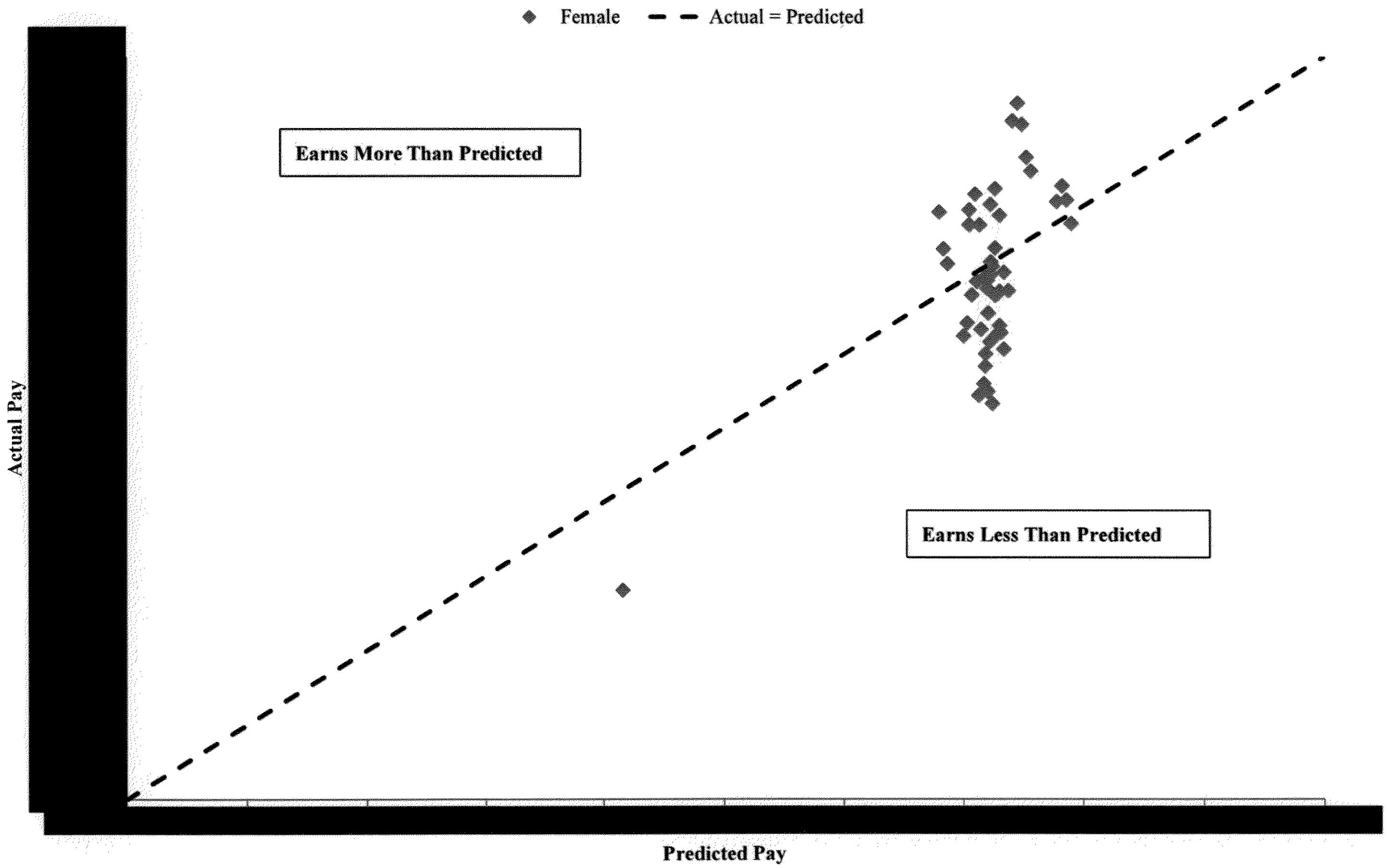


Exhibit C15

Software Developer 1: Actual Base Pay vs. Predicted Base Pay

- Prediction Based on Dr. Neumark's Data and Model, Without a Gender Control -

- Female Incumbents in Neumark's Dataset, 2013-2018 -

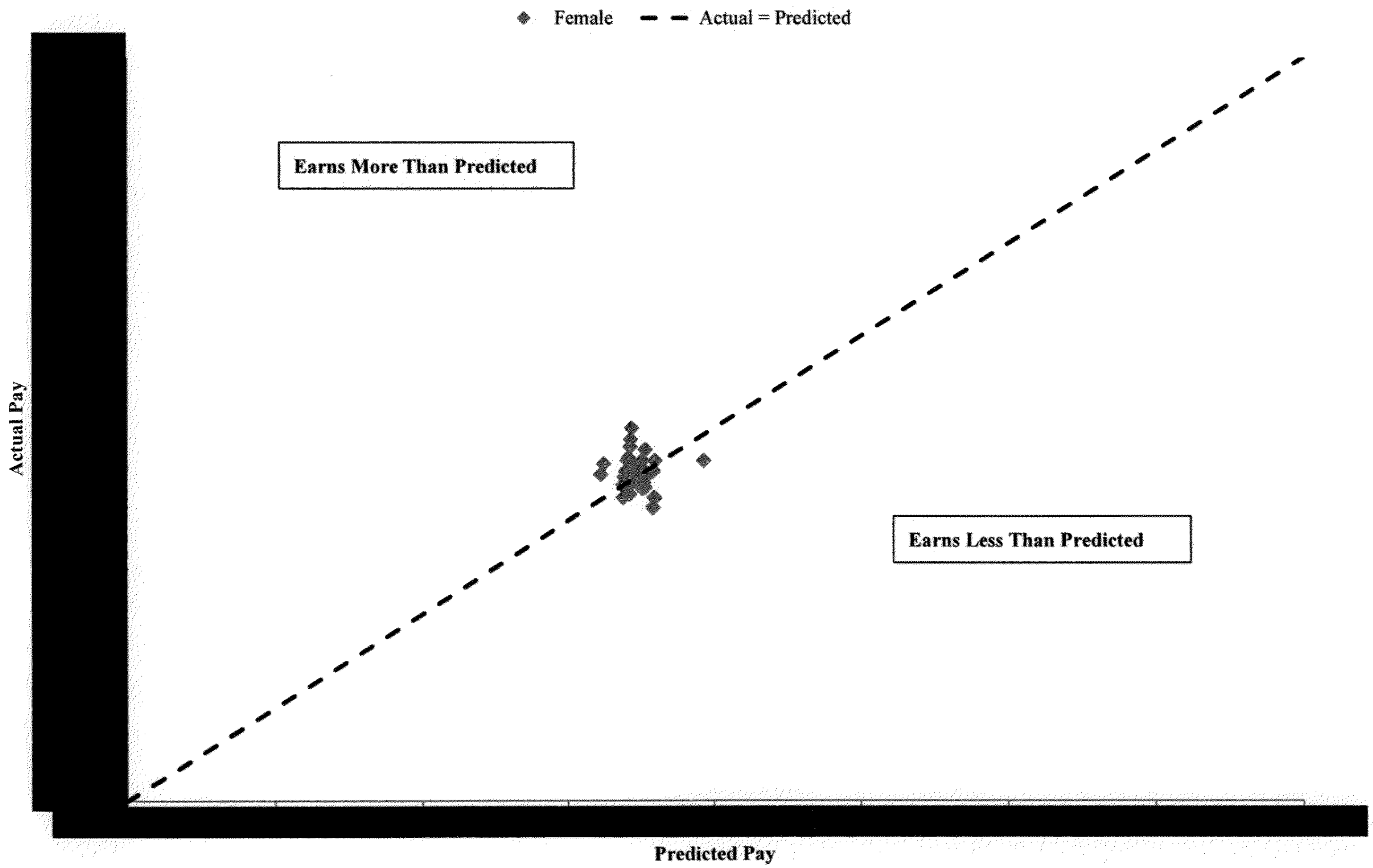
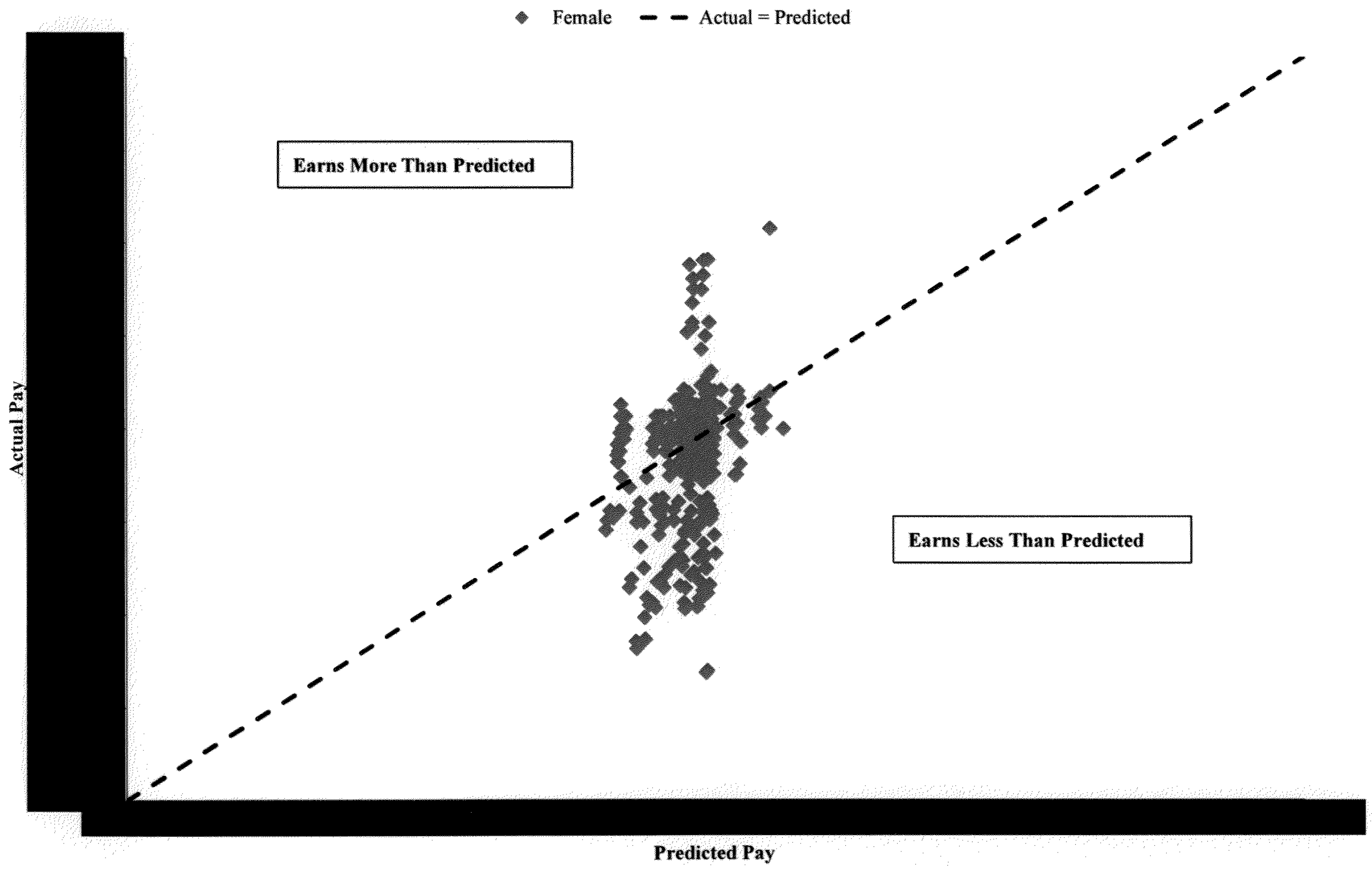


Exhibit C16

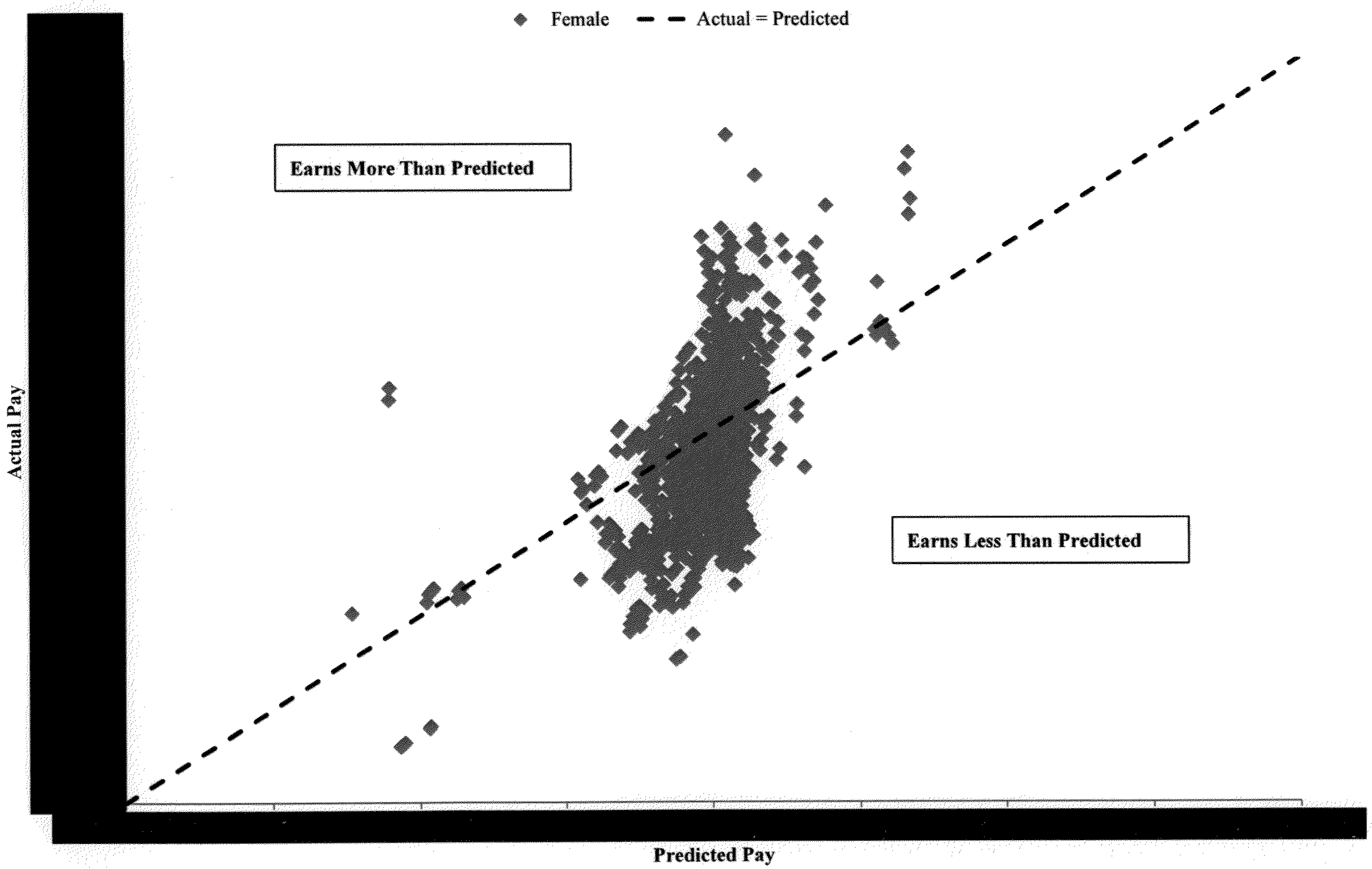
Software Developer 2: Actual Base Pay vs. Predicted Base Pay
- Prediction Based on Dr. Neumark's Data and Model, Without a Gender Control -
- Female Incumbents in Neumark's Dataset, 2013-2018 -



C16

Exhibit C17

Software Developer 3: Actual Base Pay vs. Predicted Base Pay
- Prediction Based on Dr. Neumark's Data and Model, Without a Gender Control -
- Female Incumbents in Neumark's Dataset, 2013-2018 -



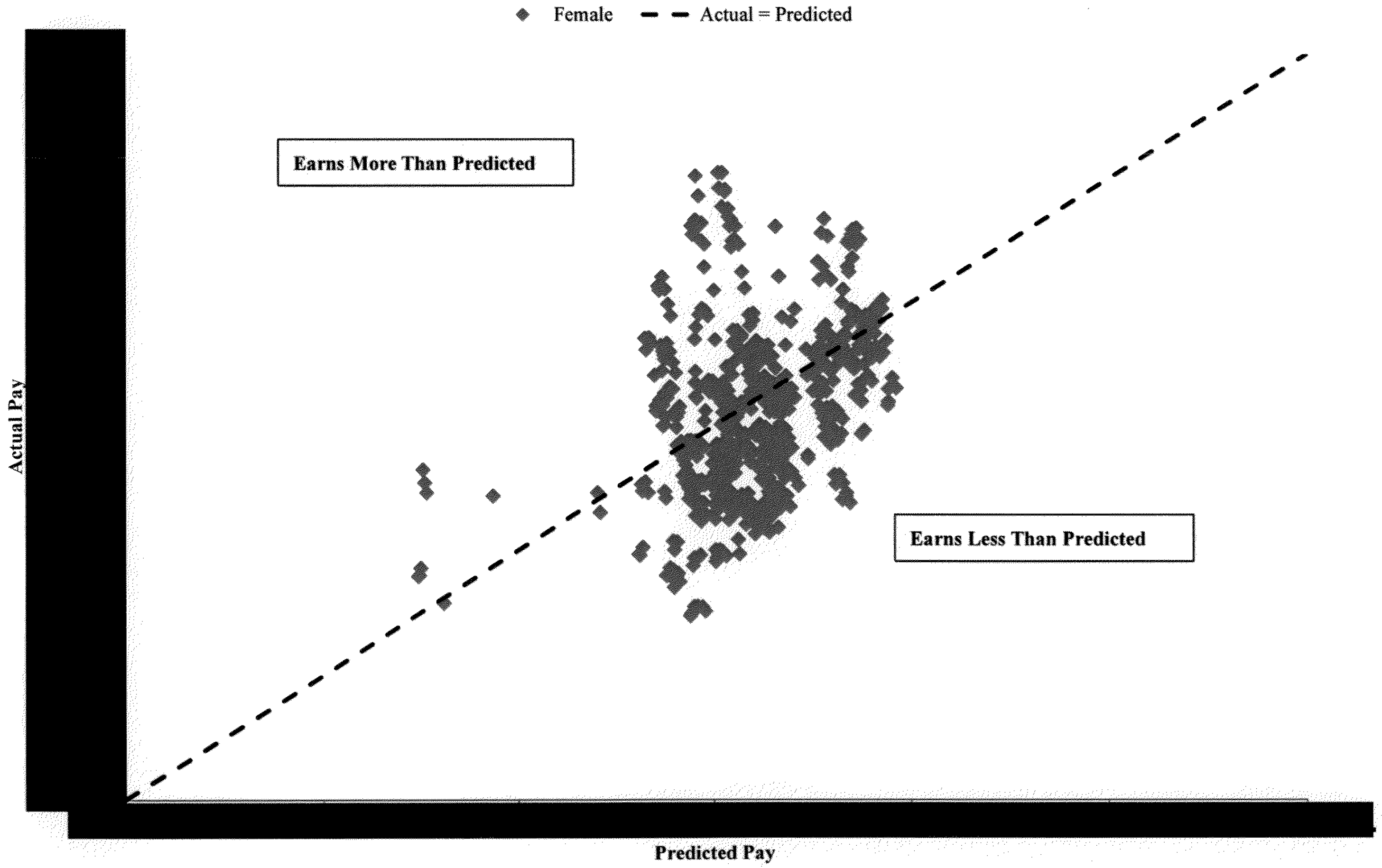
C17

Exhibit C18

Software Developer 5: Actual Base Pay vs. Predicted Base Pay

- Prediction Based on Dr. Neumark's Data and Model, Without a Gender Control -

- Female Incumbents in Neumark's Dataset, 2013-2018 -



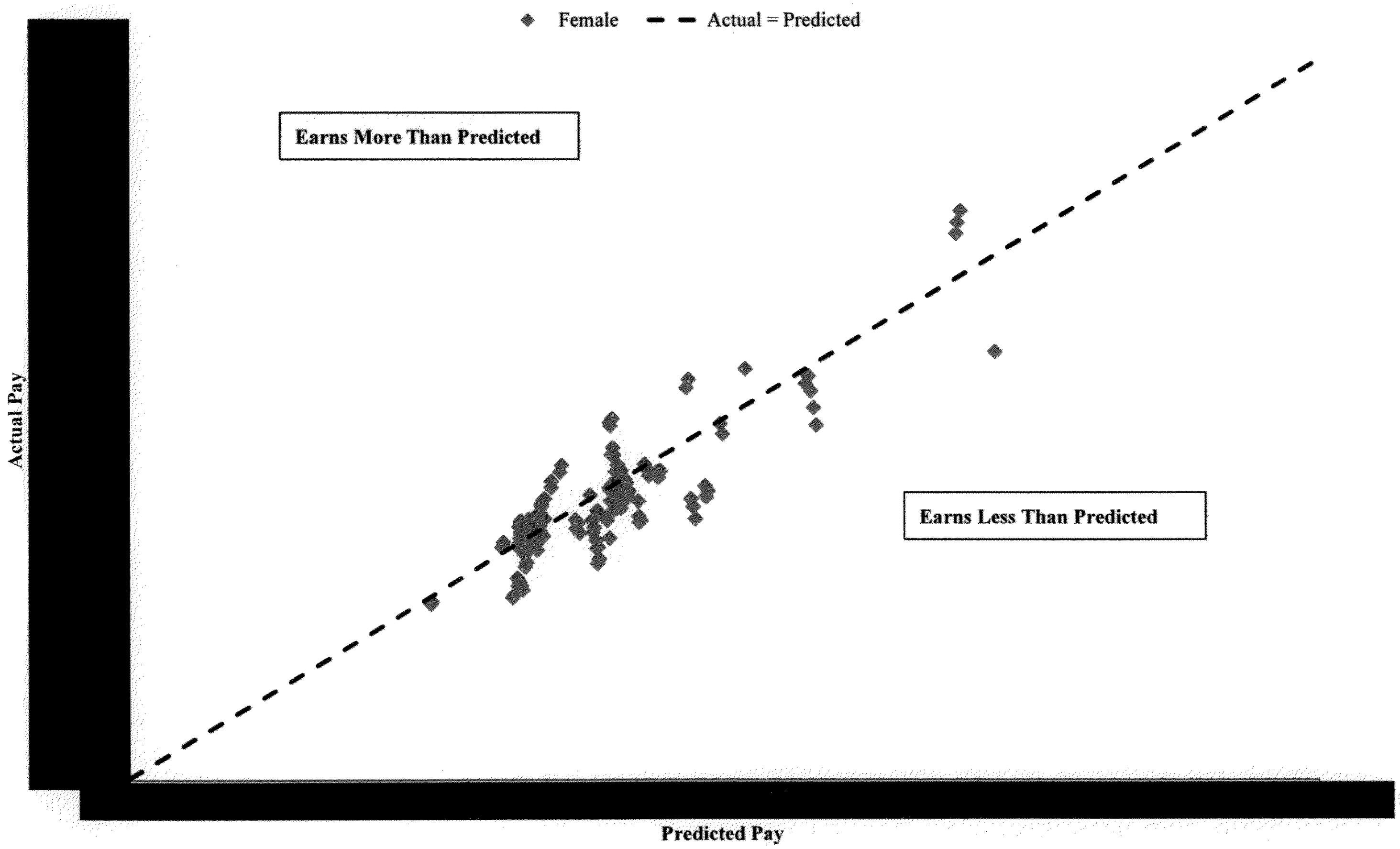
C18

Exhibit C19

0U30 - Corp Architecture - Development - ORCL USA: Actual Base Pay vs. Predicted Base Pay

- Prediction Based on Dr. Neumark's Data and Model, Without a Gender Control -

- Female Incumbents in Neumark's Dataset, 2013-2018 -



C19

Exhibit C20

AC80 - Pillar Development - ORCL USA: Actual Base Pay vs. Predicted Base Pay
- Prediction Based on Dr. Neumark's Data and Model, Without a Gender Control -
- Female Incumbents in Neumark's Dataset, 2013-2018 -

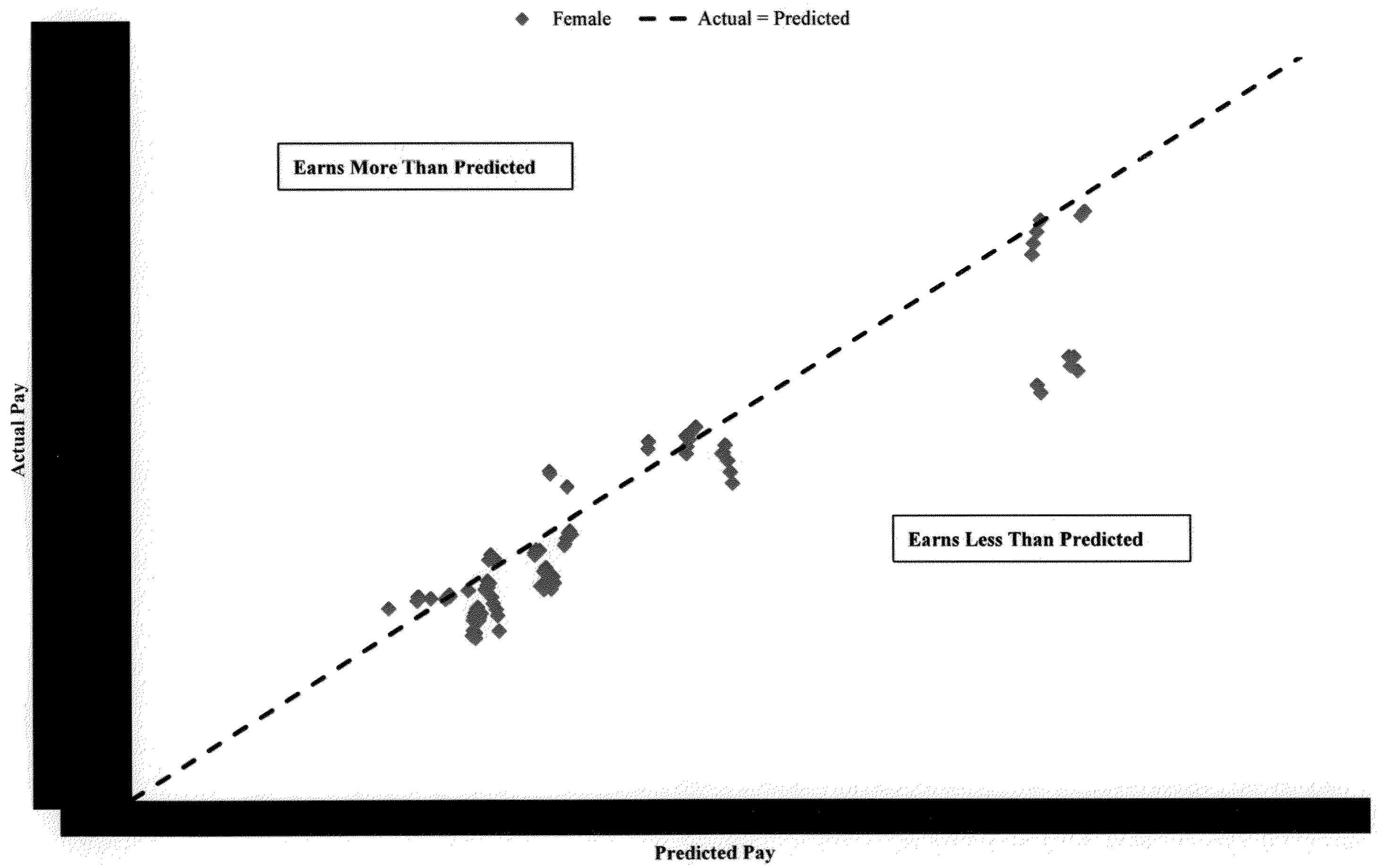


Exhibit C21

AW37 - Yosemite Falls - ORCL USA: Actual Base Pay vs. Predicted Base Pay
- Prediction Based on Dr. Neumark's Data and Model, Without a Gender Control -
- Female Incumbents in Neumark's Dataset, 2013-2018 -

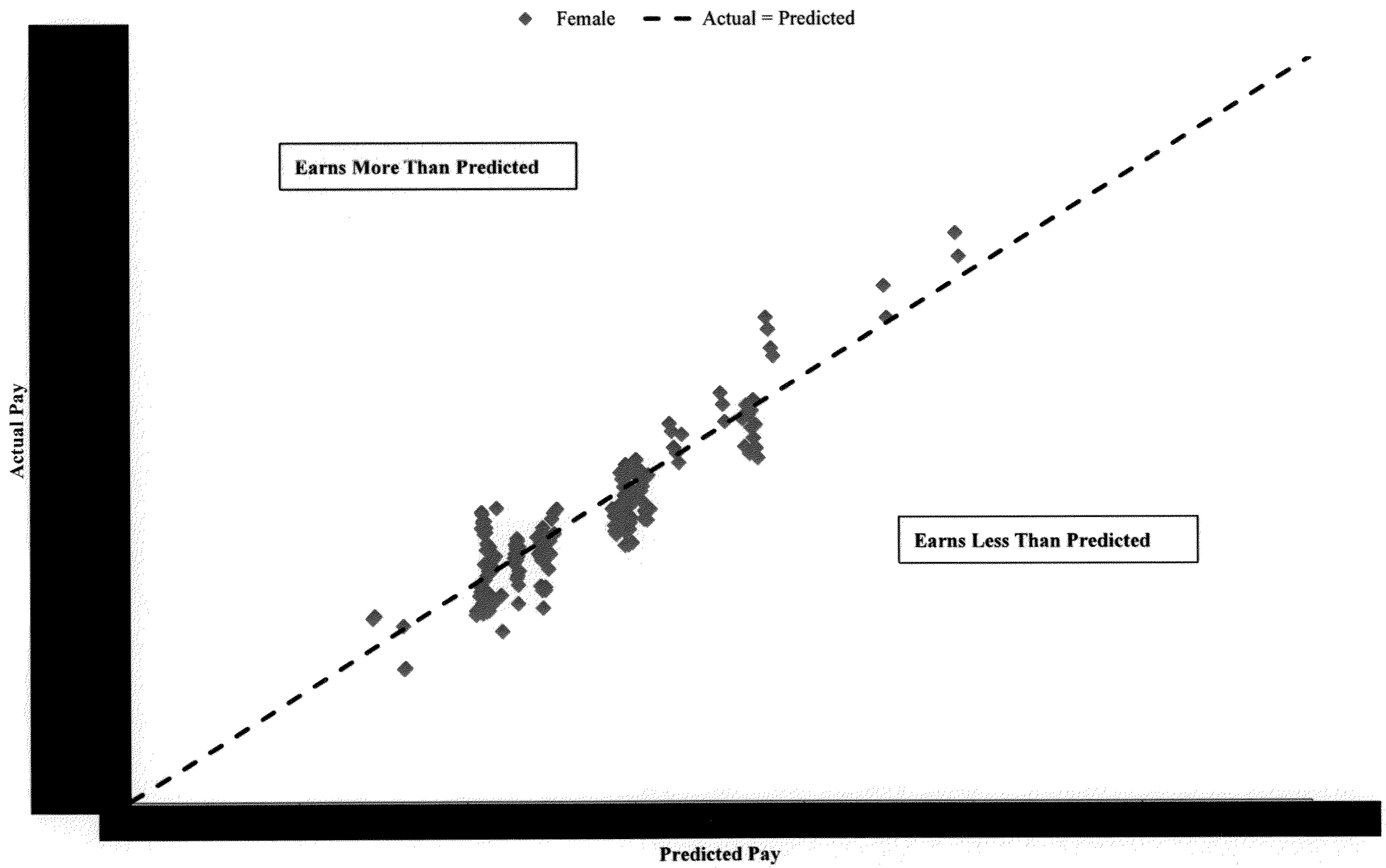
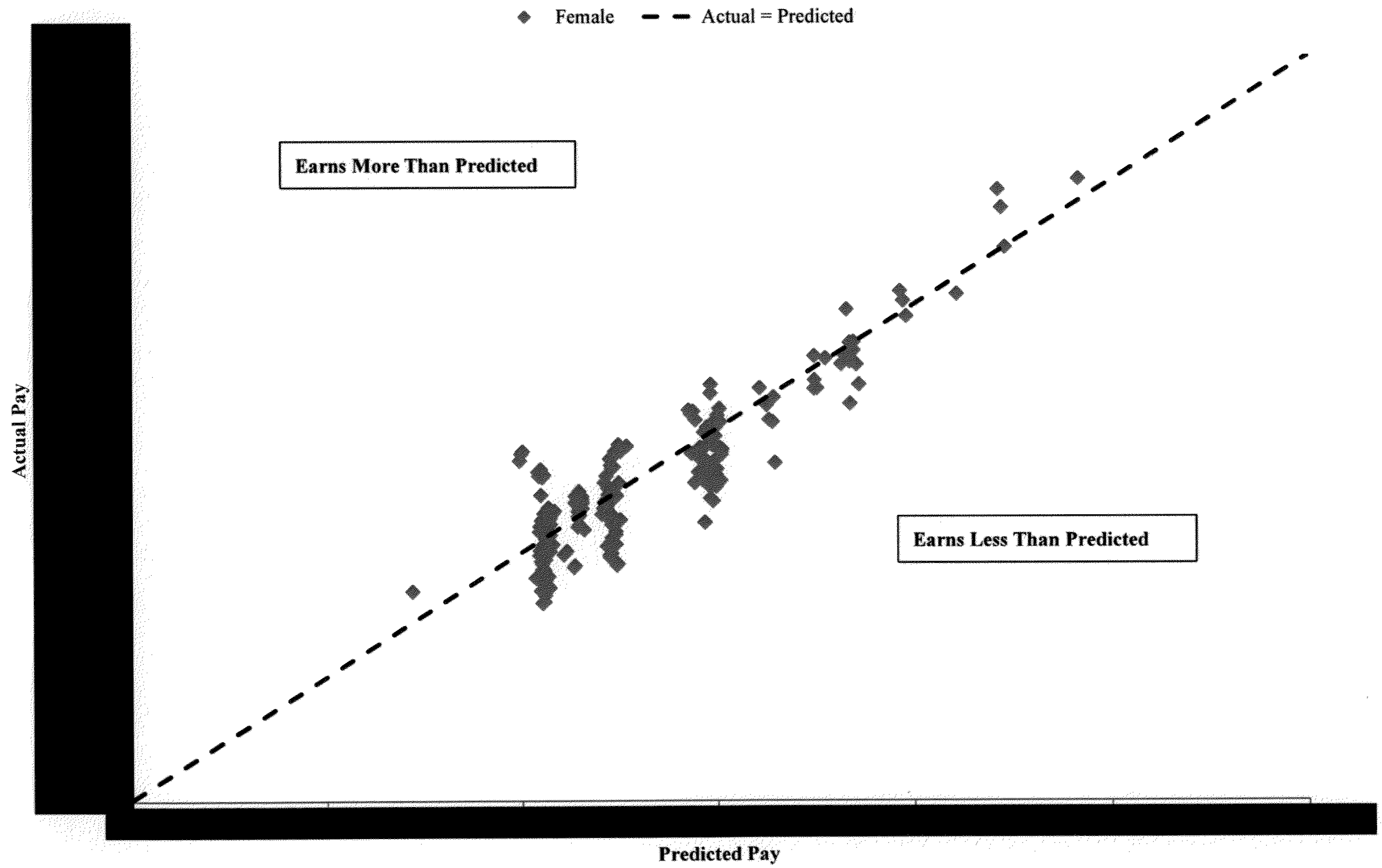


Exhibit C22

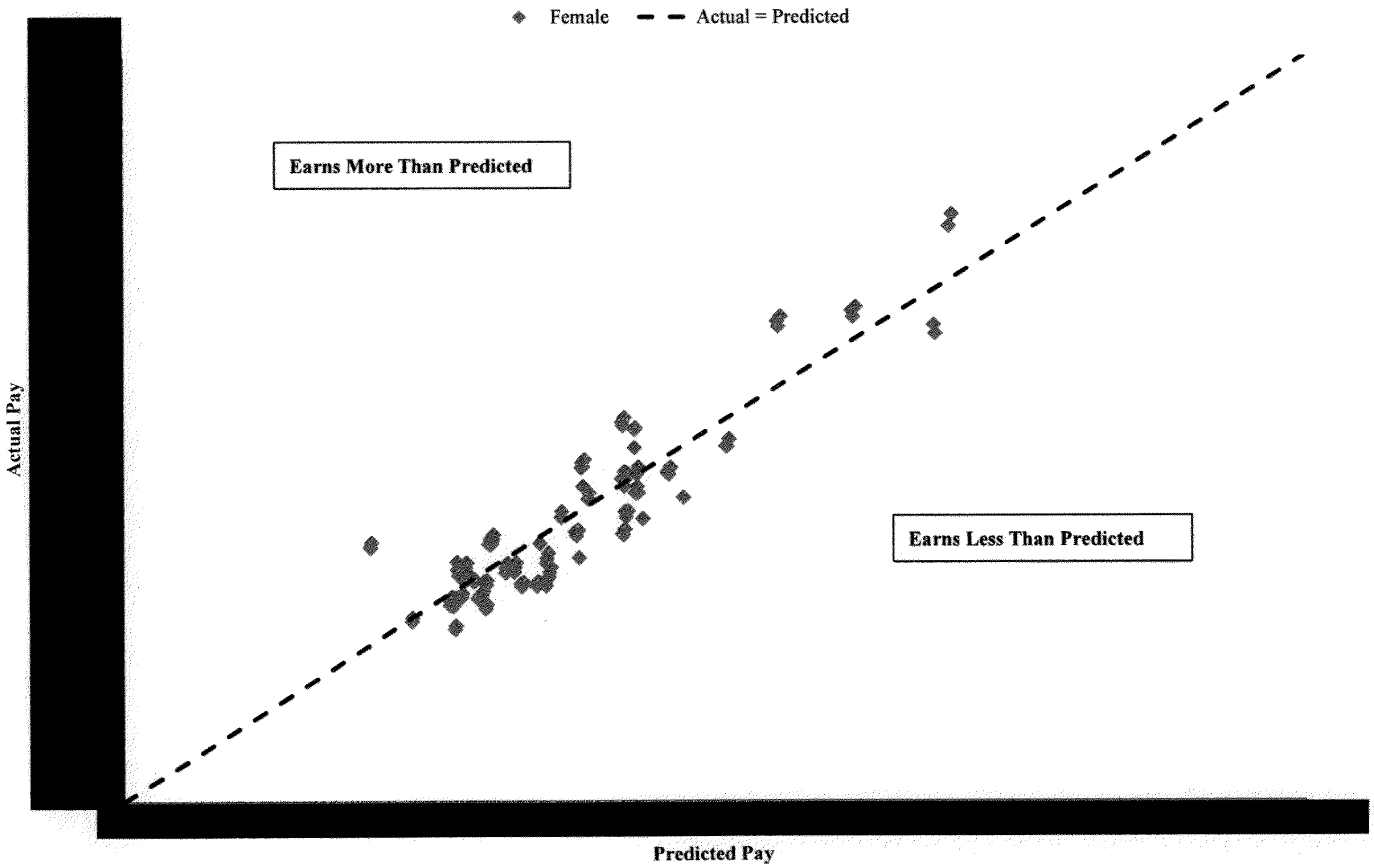
AW38 - Yellowstone Falls - ORCL USA: Actual Base Pay vs. Predicted Base Pay
- Prediction Based on Dr. Neumark's Data and Model, Without a Gender Control -
- Female Incumbents in Neumark's Dataset, 2013-2018 -



C22

Exhibit C23

BC78 - Management Cloud - ORCL USA: Actual Base Pay vs. Predicted Base Pay
- Prediction Based on Dr. Neumark's Data and Model, Without a Gender Control -
- Female Incumbents in Neumark's Dataset, 2013-2018 -



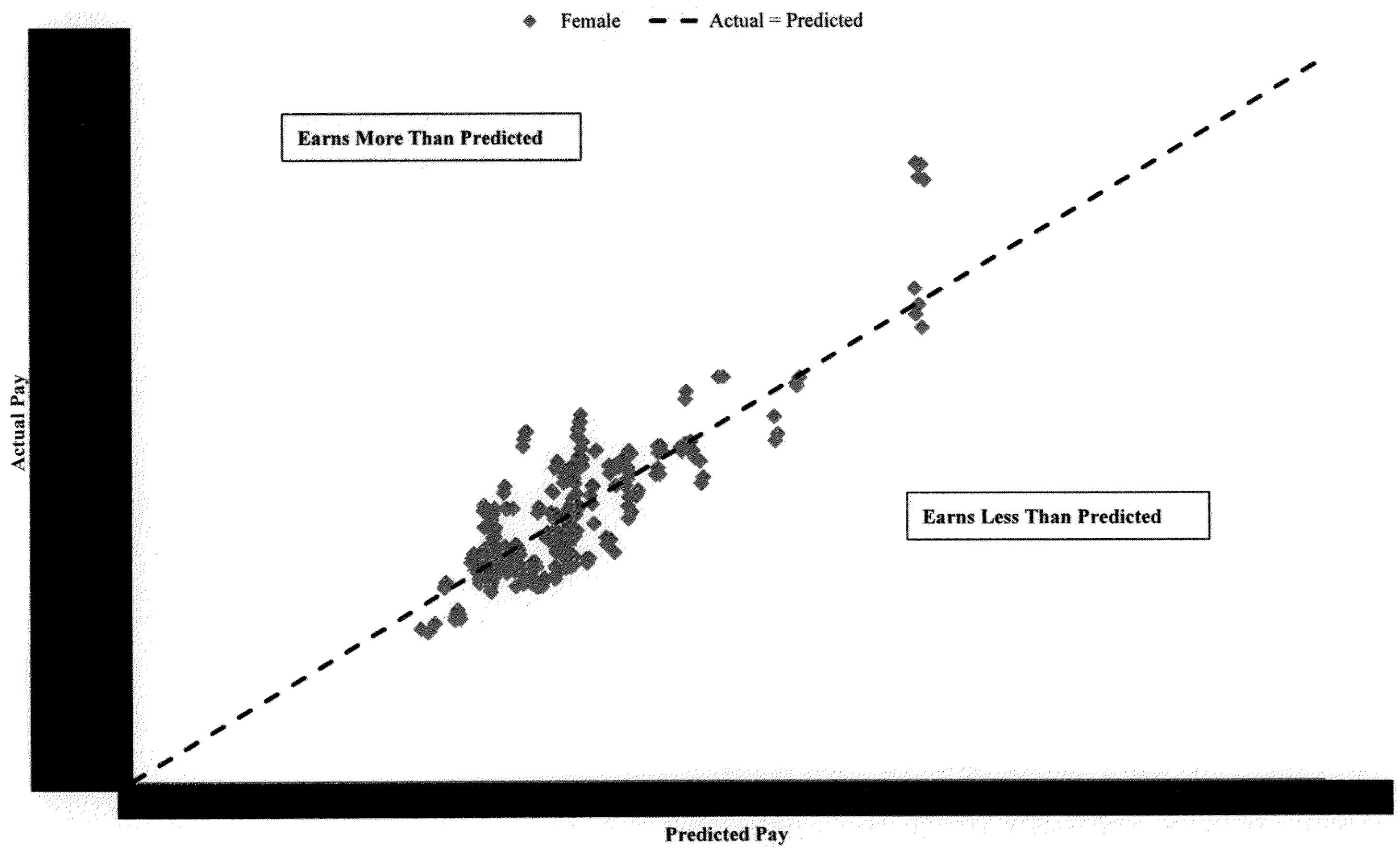
C23

Exhibit C24

BG16 - Public Cloud Platform Development - ORCL USA: Actual Base Pay vs. Predicted Base Pay

- Prediction Based on Dr. Neumark's Data and Model, Without a Gender Control -

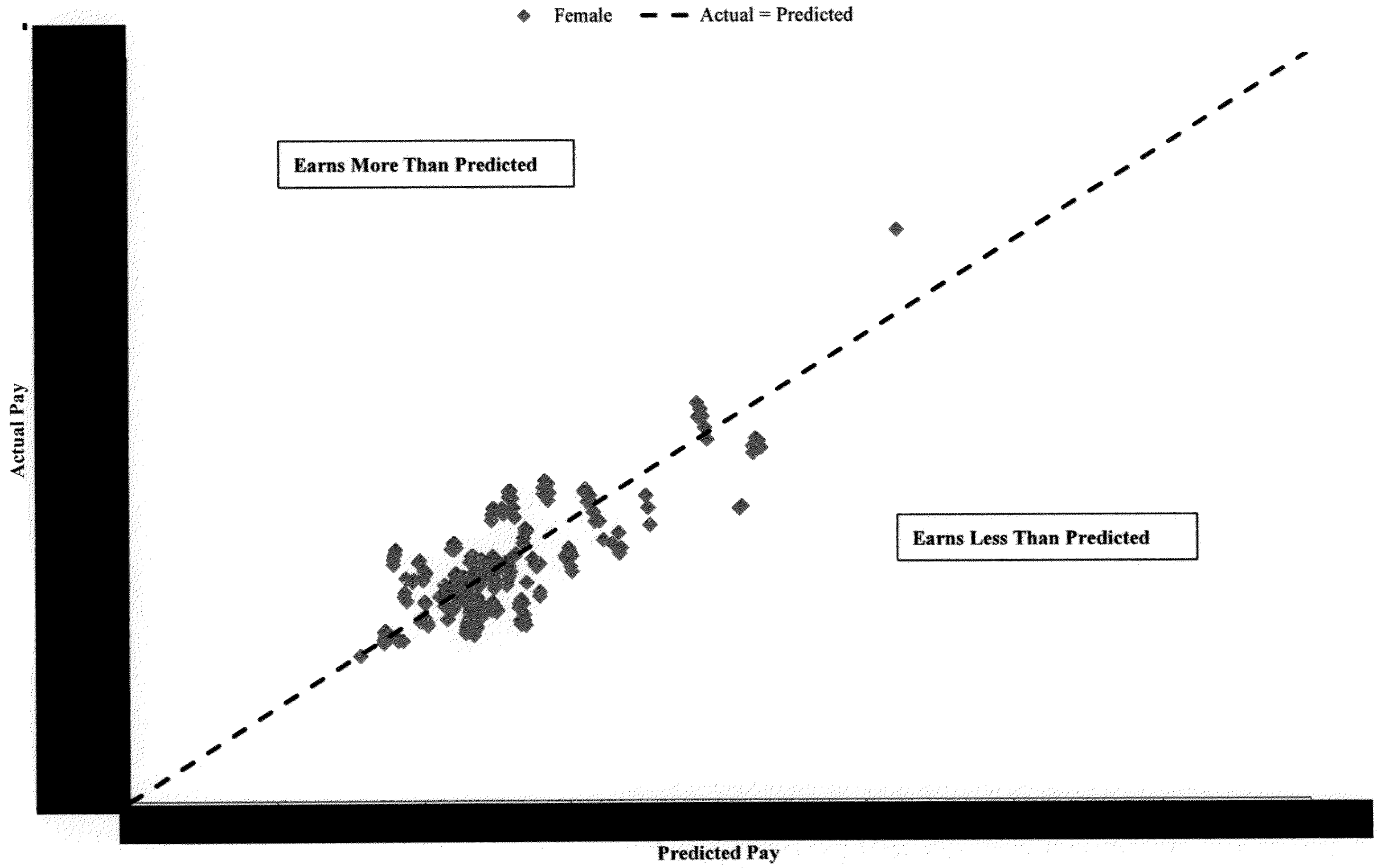
- Female Incumbents in Neumark's Dataset, 2013-2018 -



C24

Exhibit C25

DV14 - Engineering IT - ORCL USA: Actual Base Pay vs. Predicted Base Pay
- Prediction Based on Dr. Neumark's Data and Model, Without a Gender Control -
- Female Incumbents in Neumark's Dataset, 2013-2018 -



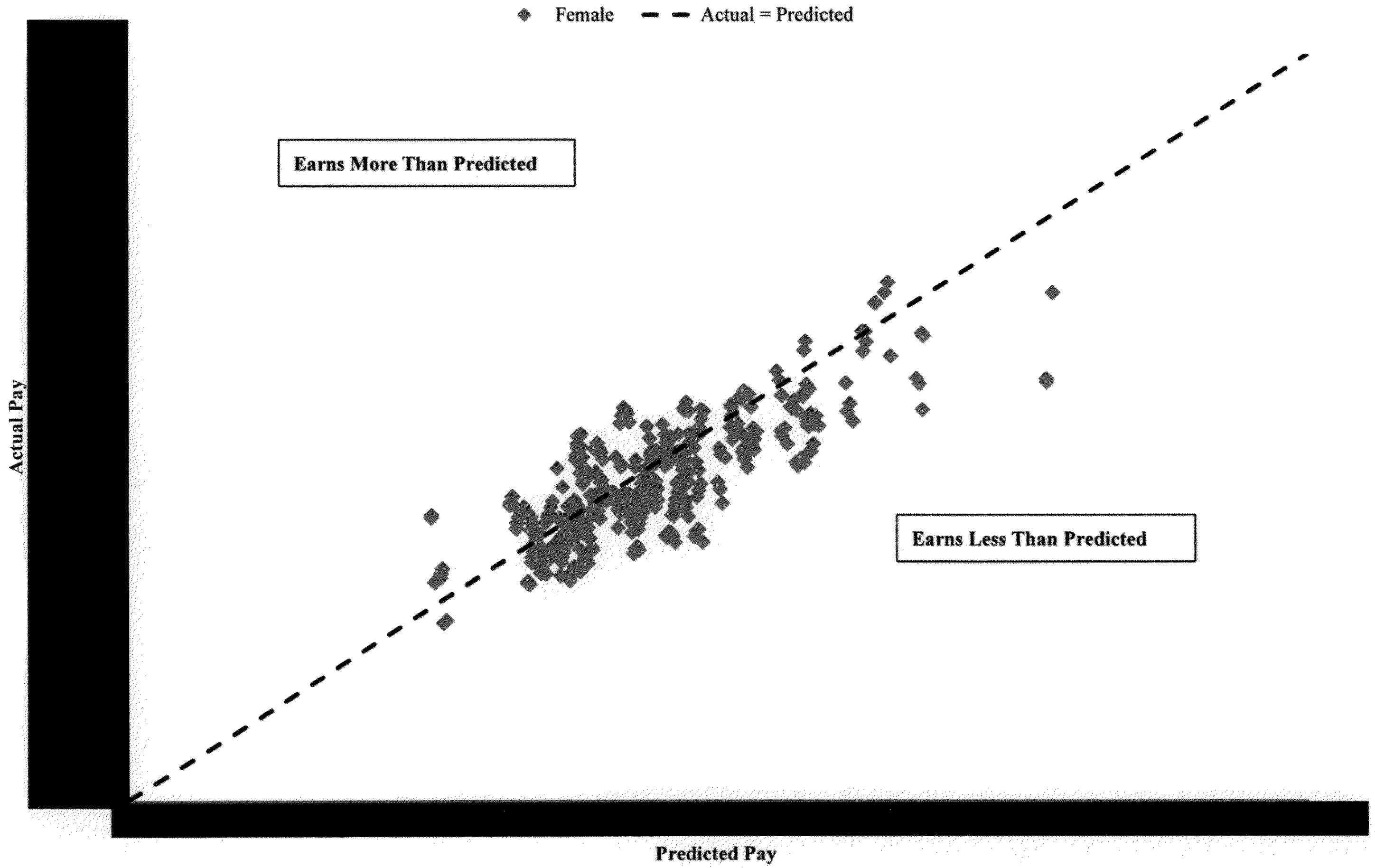
C25

Exhibit C26

PD98 - Fusion HCM Development - ORCL USA: Actual Base Pay vs. Predicted Base Pay

- Prediction Based on Dr. Neumark's Data and Model, Without a Gender Control -

- Female Incumbents in Neumark's Dataset, 2013-2018 -



C26

Exhibit C27

PL05 - Fusion Financials Development - ORCL USA: Actual Base Pay vs. Predicted Base Pay

- Prediction Based on Dr. Neumark's Data and Model, Without a Gender Control -

- Female Incumbents in Neumark's Dataset, 2013-2018 -

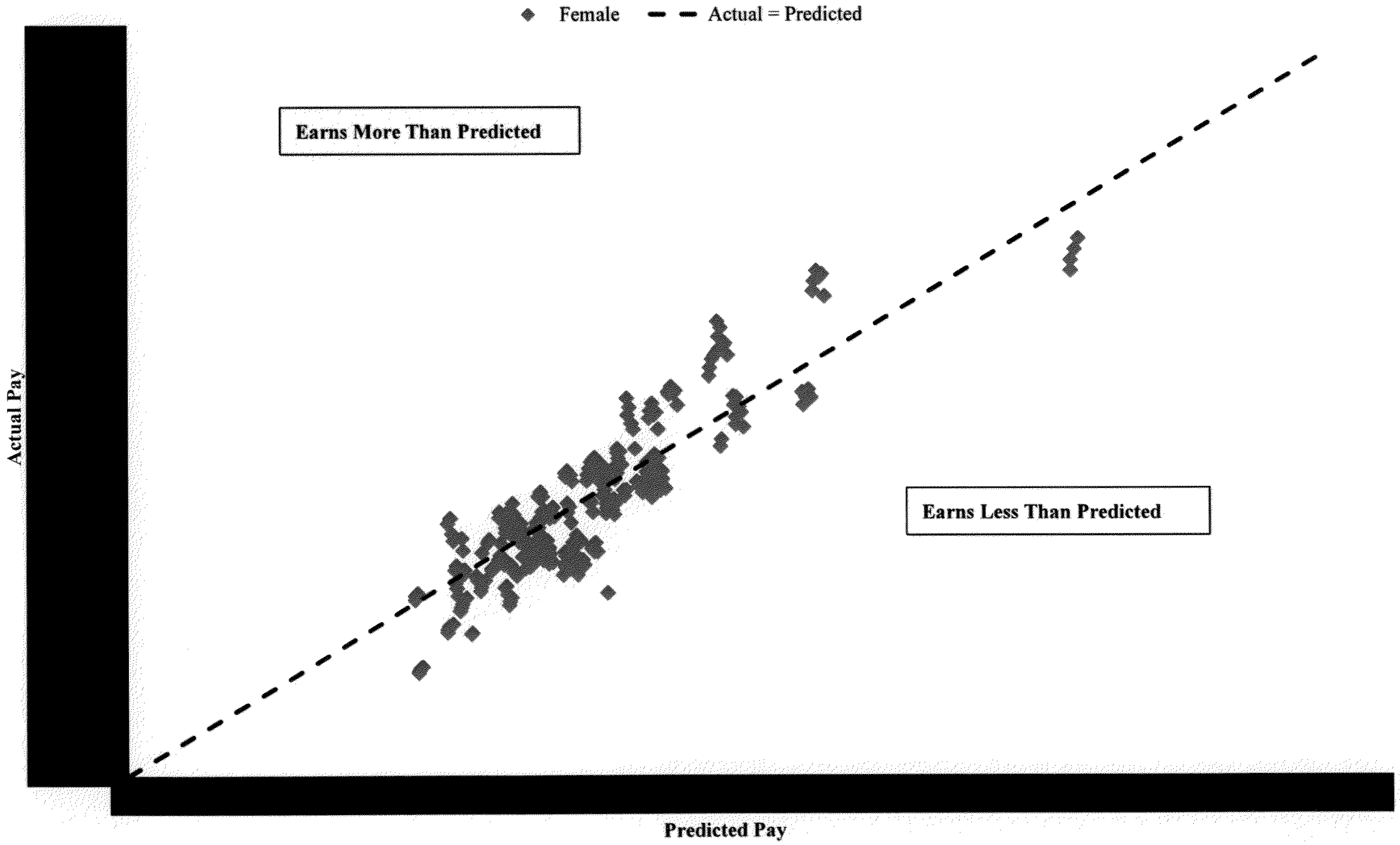
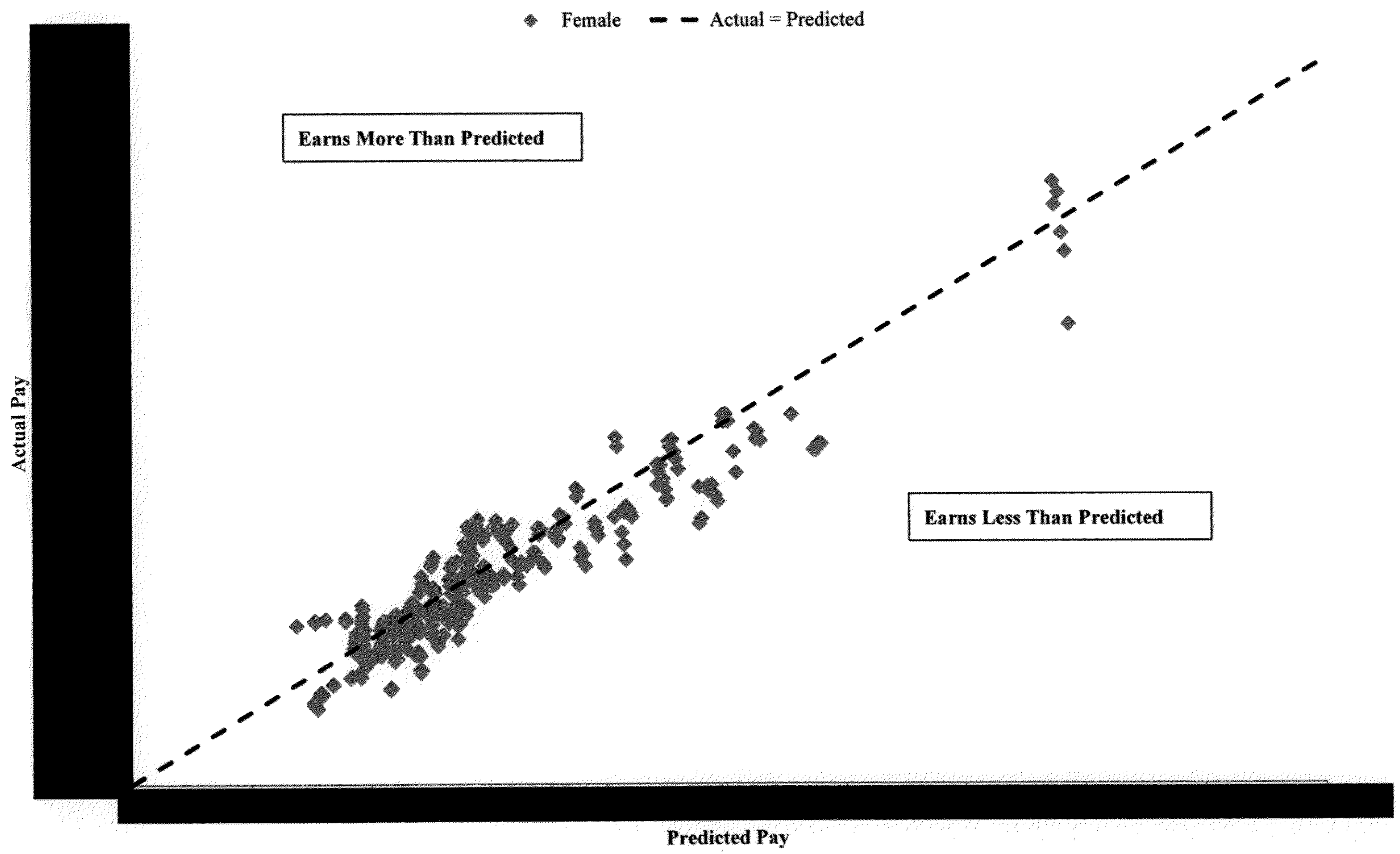


Exhibit C28

PL07 - Fusion Development Management - ORCL USA: Actual Base Pay vs. Predicted Base Pay

- Prediction Based on Dr. Neumark's Data and Model, Without a Gender Control -

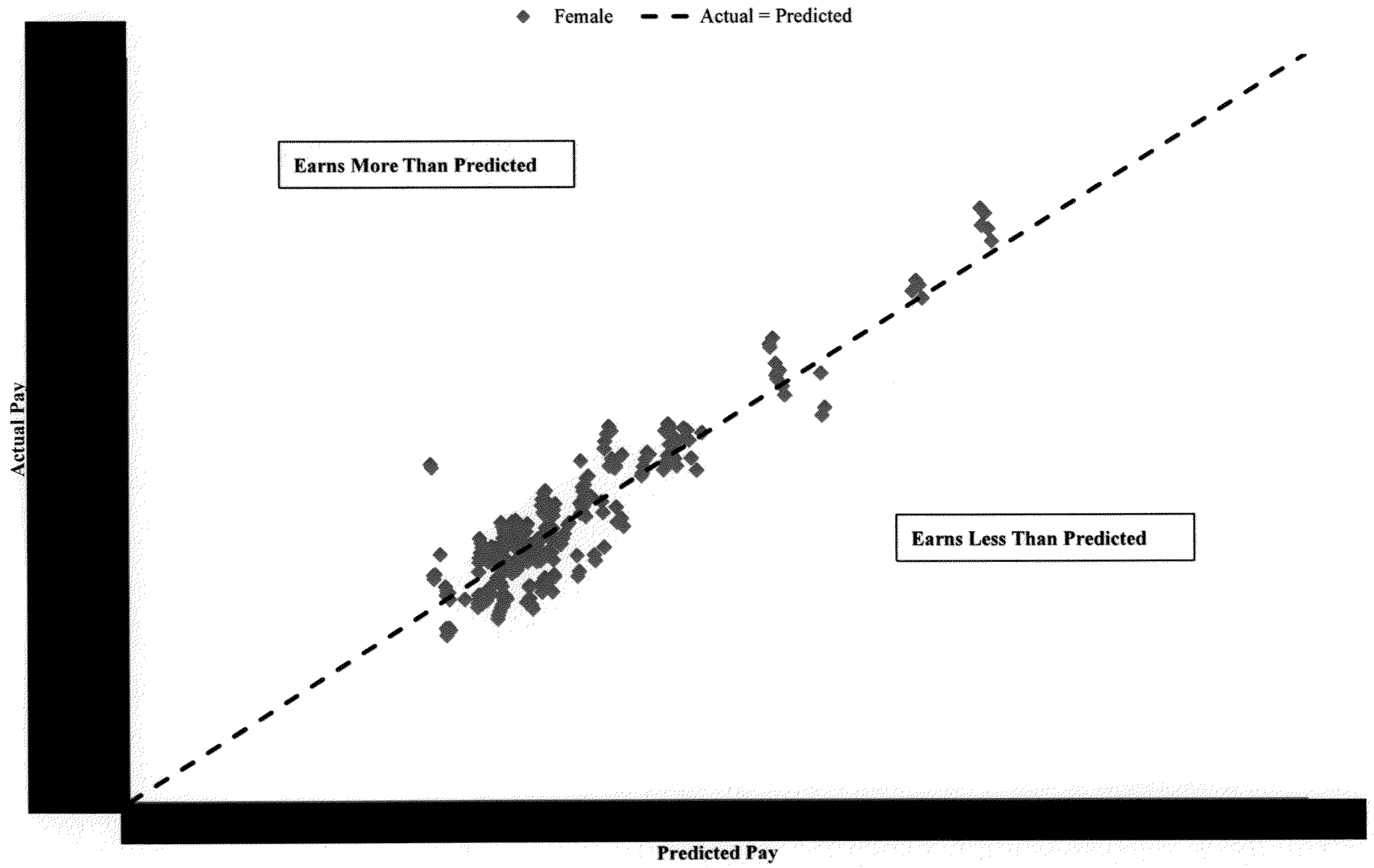
- Female Incumbents in Neumark's Dataset, 2013-2018 -



C28

Exhibit C29

RS67 - OAL - ORCL USA: Actual Base Pay vs. Predicted Base Pay
- Prediction Based on Dr. Neumark's Data and Model, Without a Gender Control -
- Female Incumbents in Neumark's Dataset, 2013-2018 -



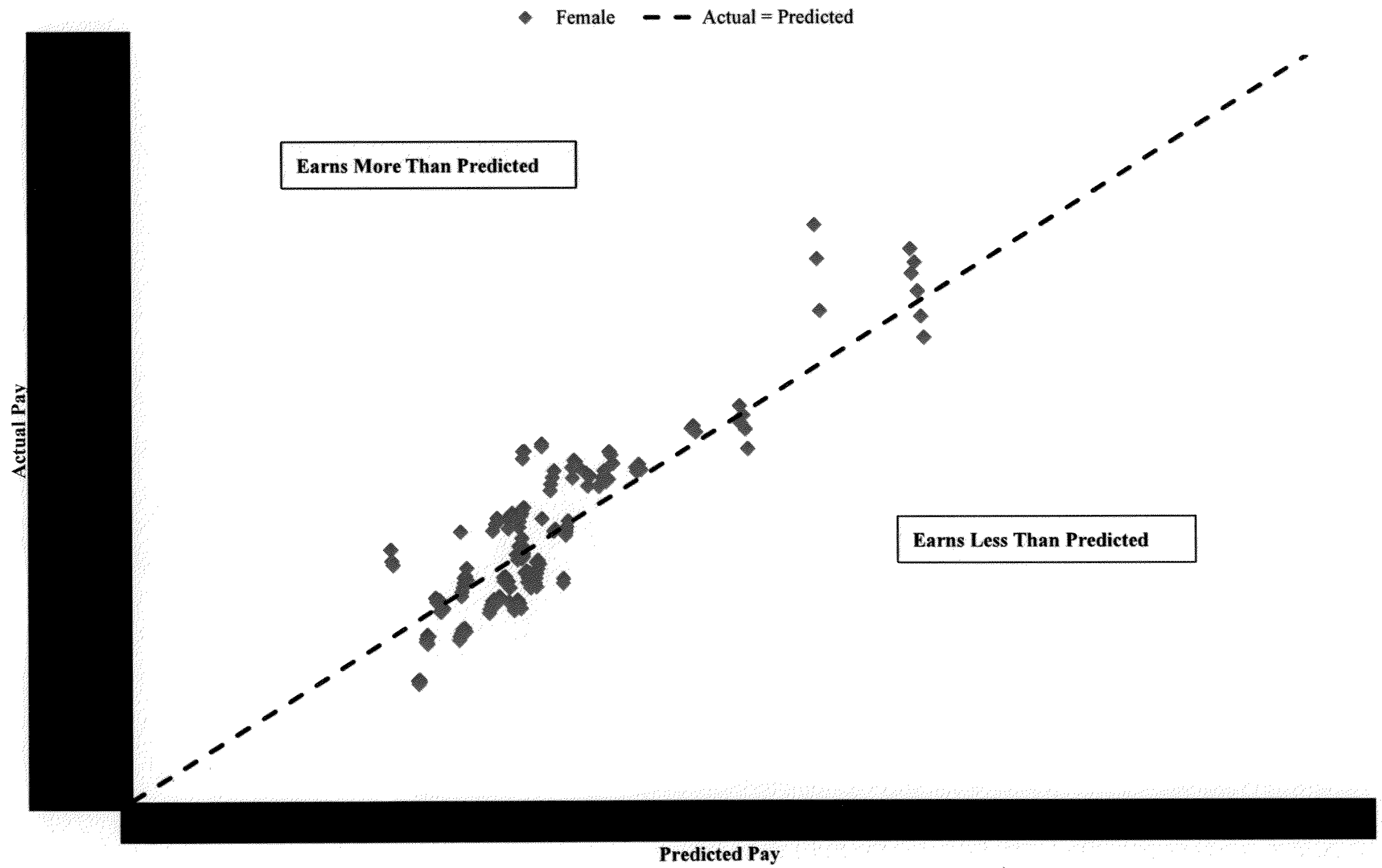
C29

Exhibit C30

RS70 - Enterprise IT - ORCL USA: Actual Base Pay vs. Predicted Base Pay

- Prediction Based on Dr. Neumark's Data and Model, Without a Gender Control -

- Female Incumbents in Neumark's Dataset, 2013-2018 -



C30

Exhibit C31

SI26 - Oracle Plan R&D - ORCL USA: Actual Base Pay vs. Predicted Base Pay

- Prediction Based on Dr. Neumark's Data and Model, Without a Gender Control -

- Female Incumbents in Neumark's Dataset, 2013-2018 -

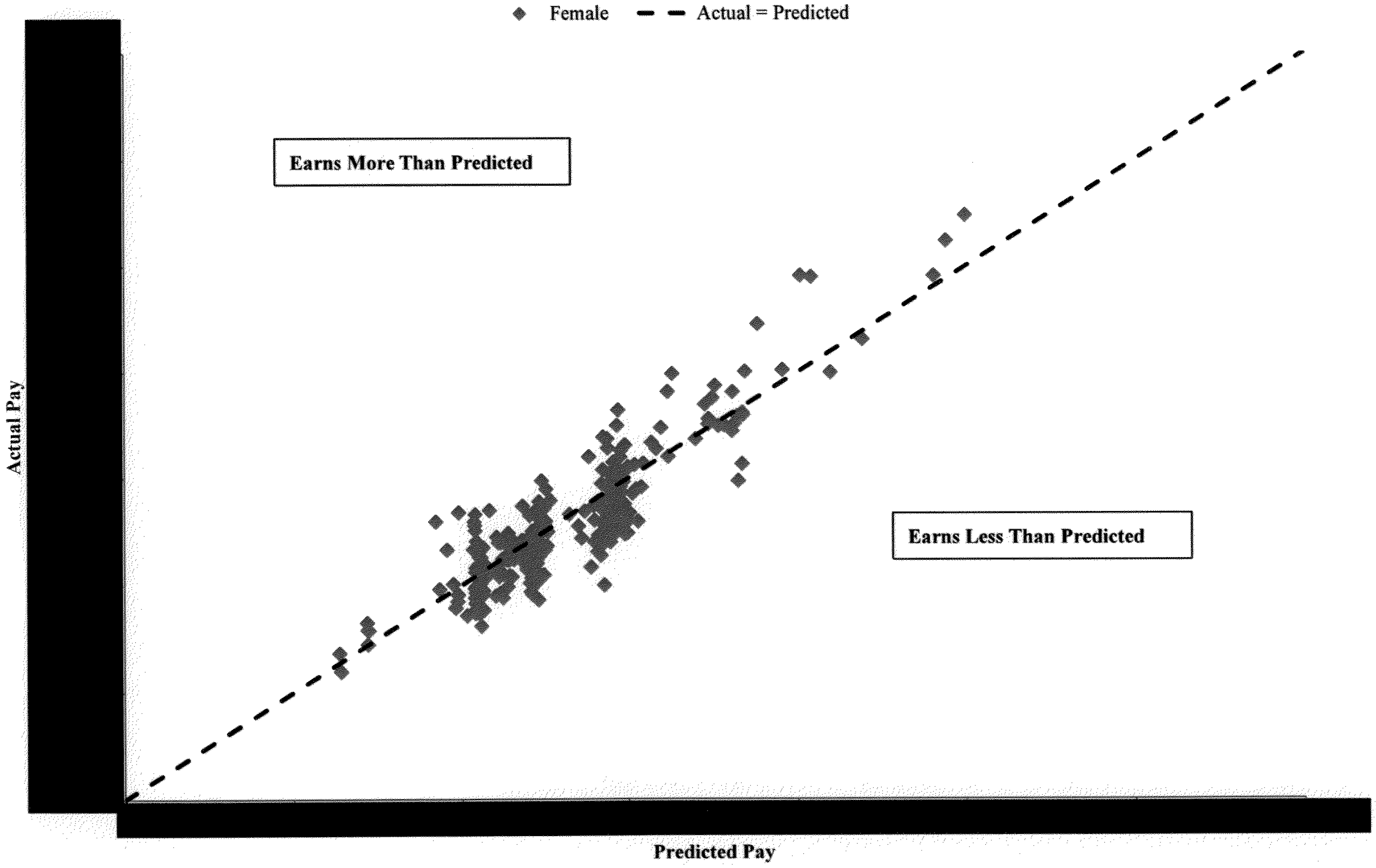
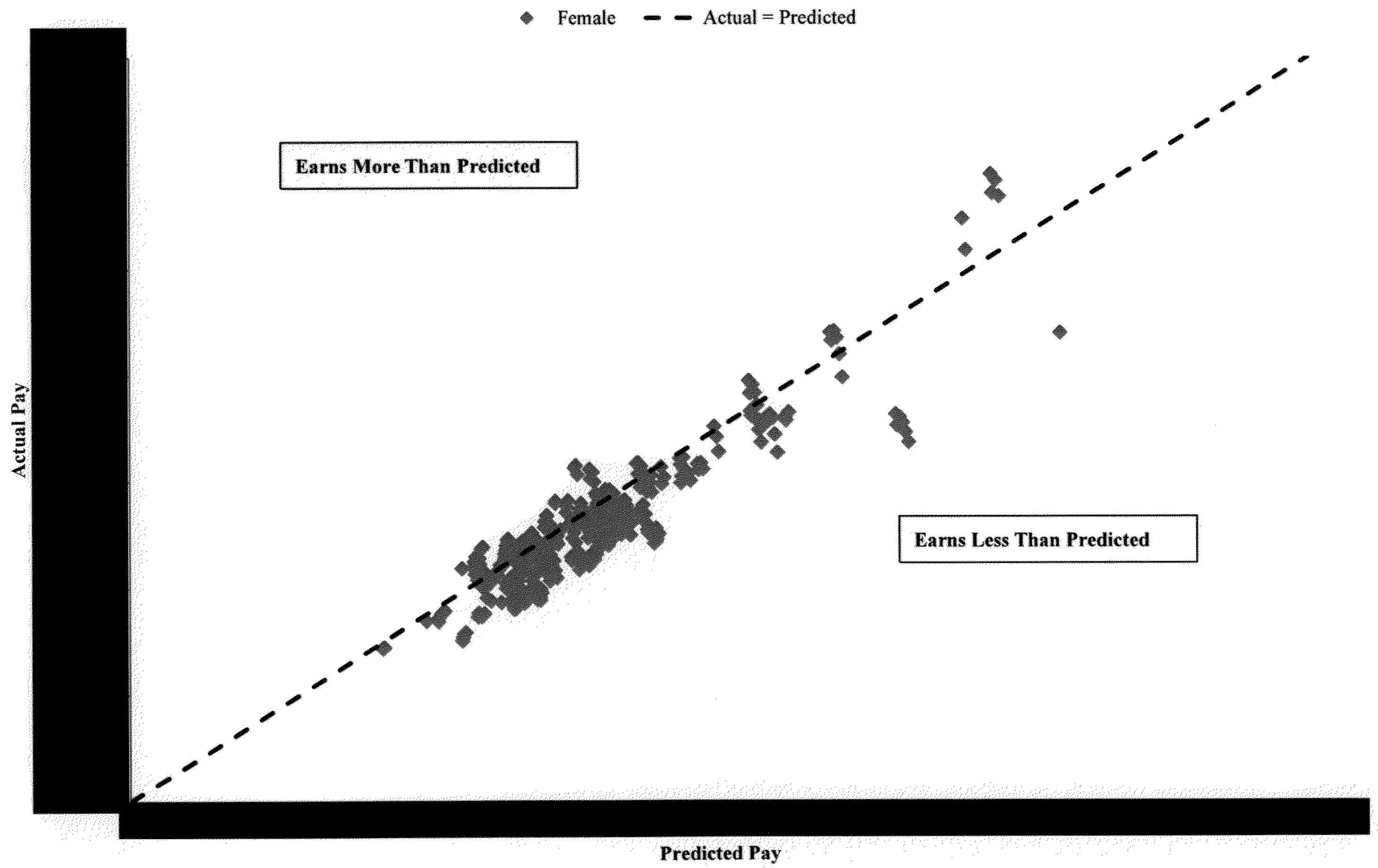


Exhibit C32

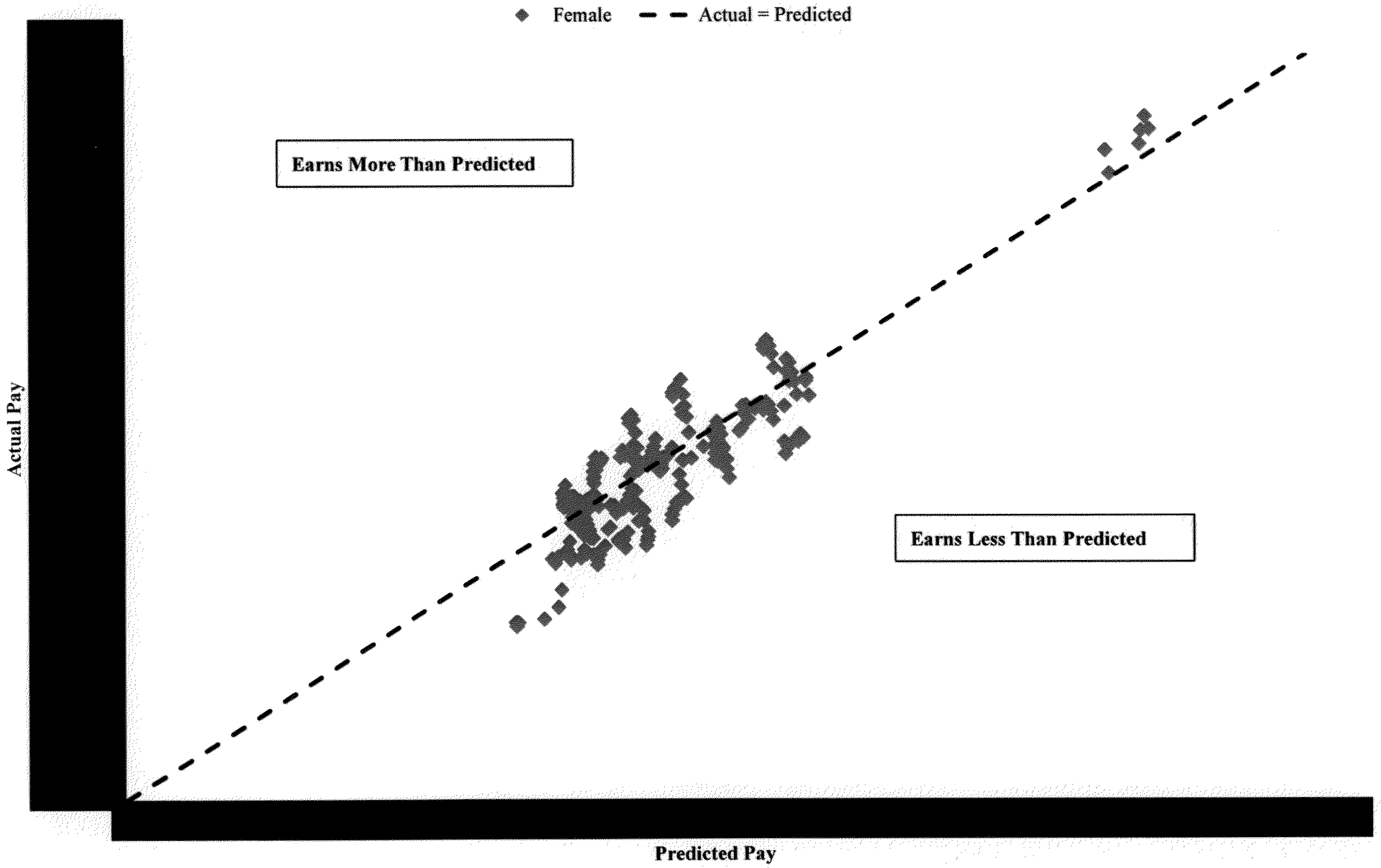
SL64 - Fusion SCM Development - ORCL USA: Actual Base Pay vs. Predicted Base Pay
- Prediction Based on Dr. Neumark's Data and Model, Without a Gender Control -
- Female Incumbents in Neumark's Dataset, 2013-2018 -



C32

Exhibit C33

SL65 - Fusion CRM Development - ORCL USA: Actual Base Pay vs. Predicted Base Pay
- Prediction Based on Dr. Neumark's Data and Model, Without a Gender Control -
- Female Incumbents in Neumark's Dataset, 2013-2018 -



C33

Exhibit C34

2013: Actual Base Pay vs. Predicted Base Pay
- Prediction Based on Dr. Neumark's Data and Model, Without a Gender Control -
- Female Incumbents in Neumark's Dataset, 2013-2018 -

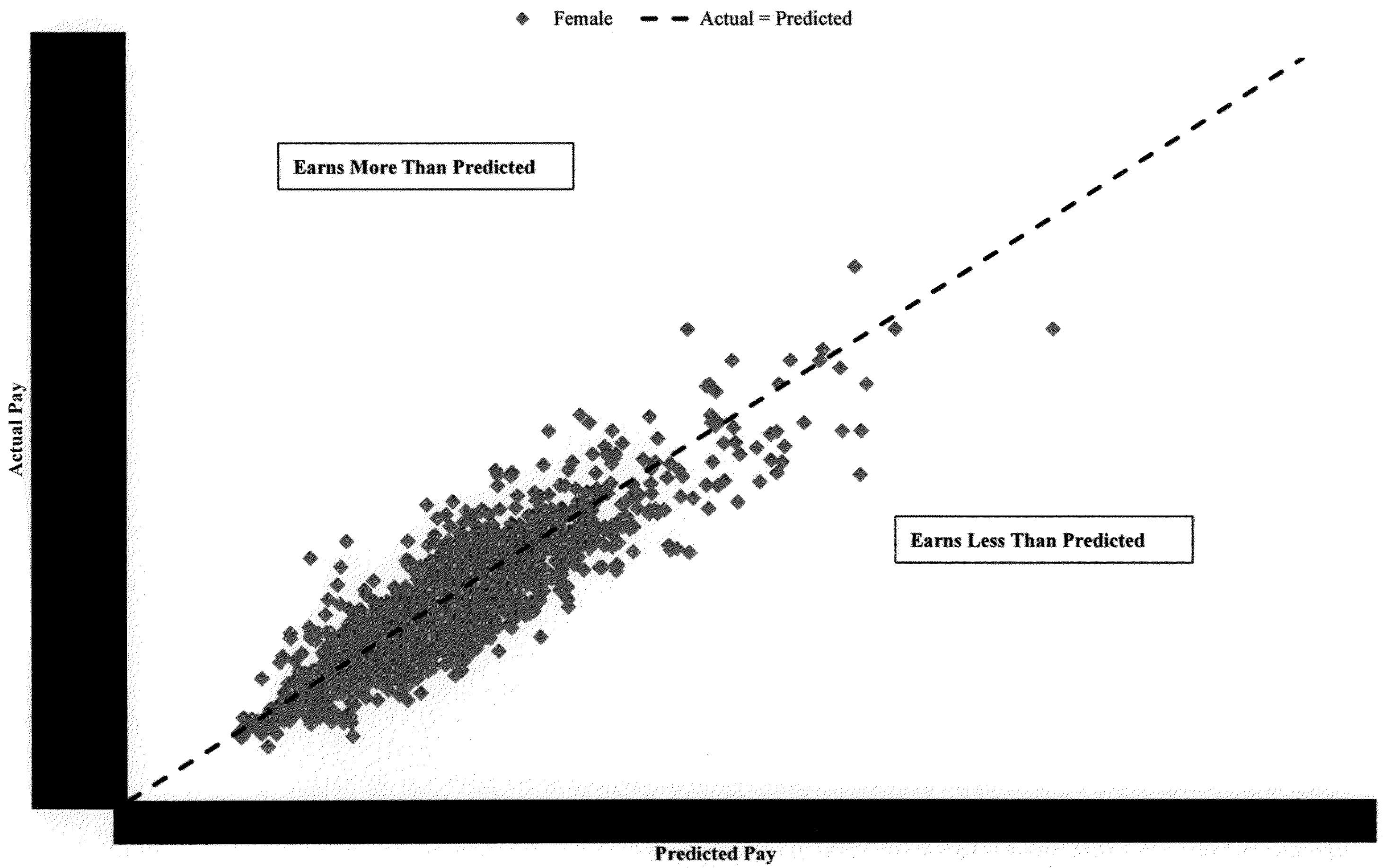


Exhibit C35

2014: Actual Base Pay vs. Predicted Base Pay
- Prediction Based on Dr. Neumark's Data and Model, Without a Gender Control -
- Female Incumbents in Neumark's Dataset, 2013-2018 -

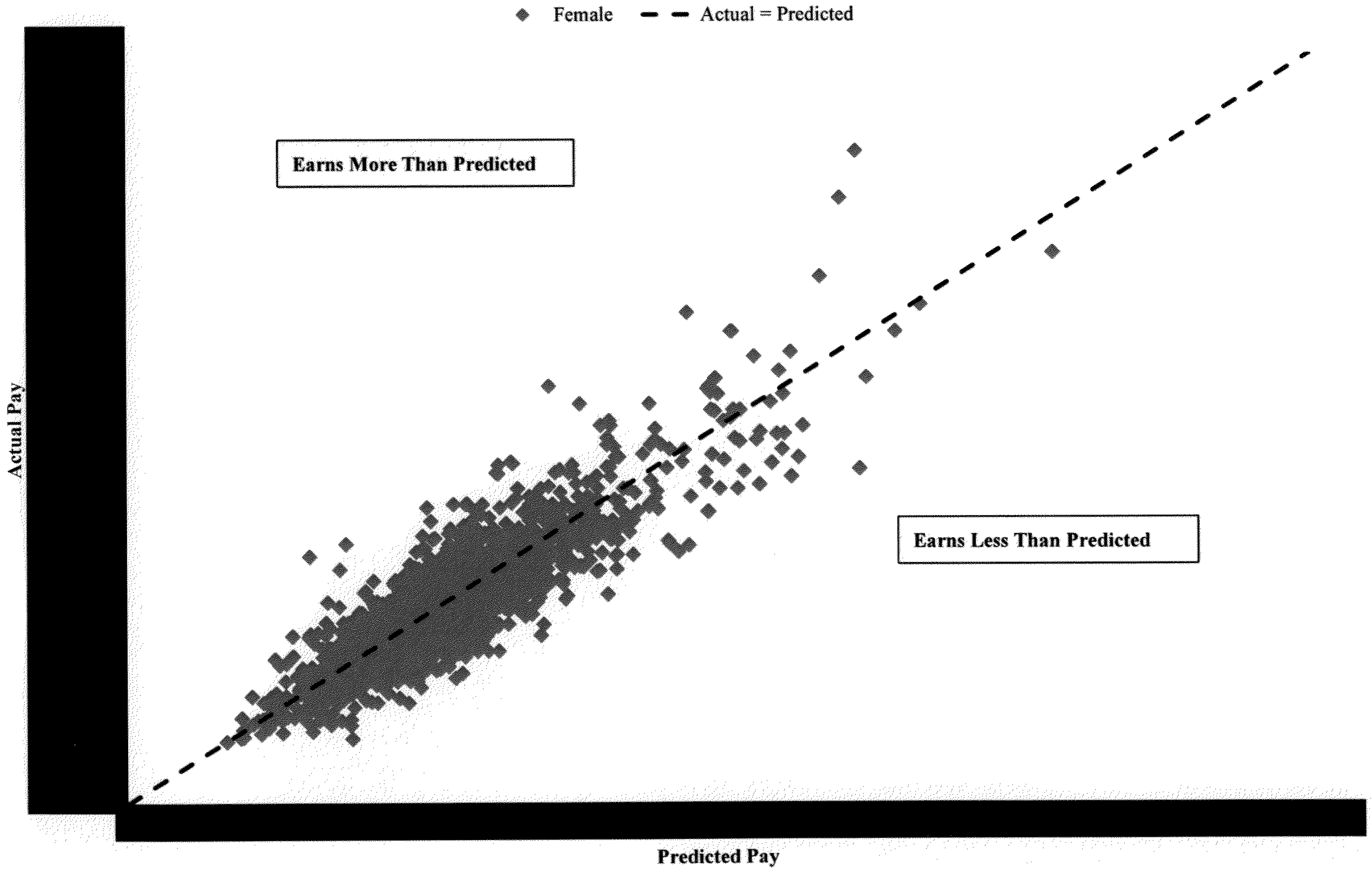
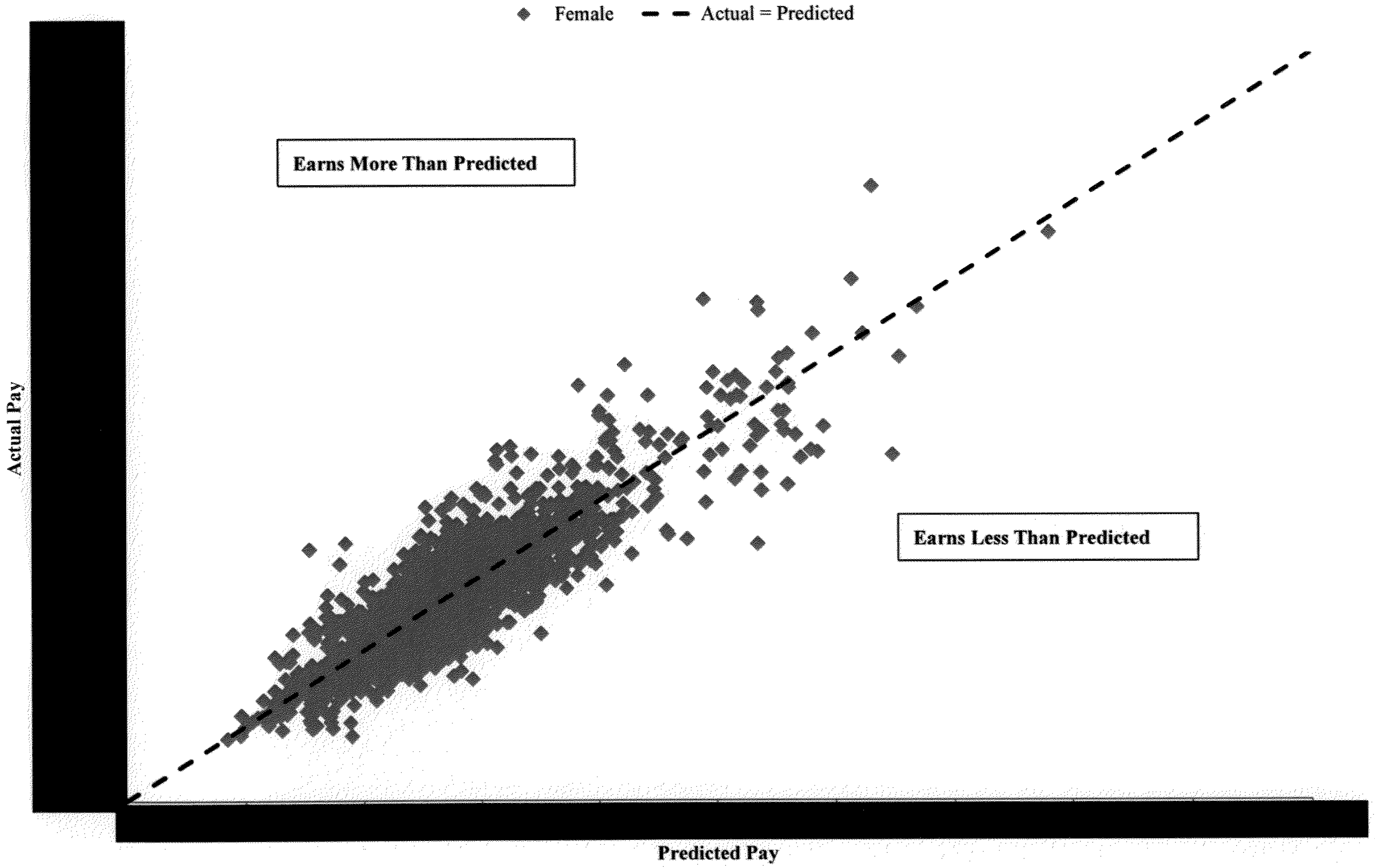


Exhibit C36

2015: Actual Base Pay vs. Predicted Base Pay

- Prediction Based on Dr. Neumark's Data and Model, Without a Gender Control -

- Female Incumbents in Neumark's Dataset, 2013-2018 -



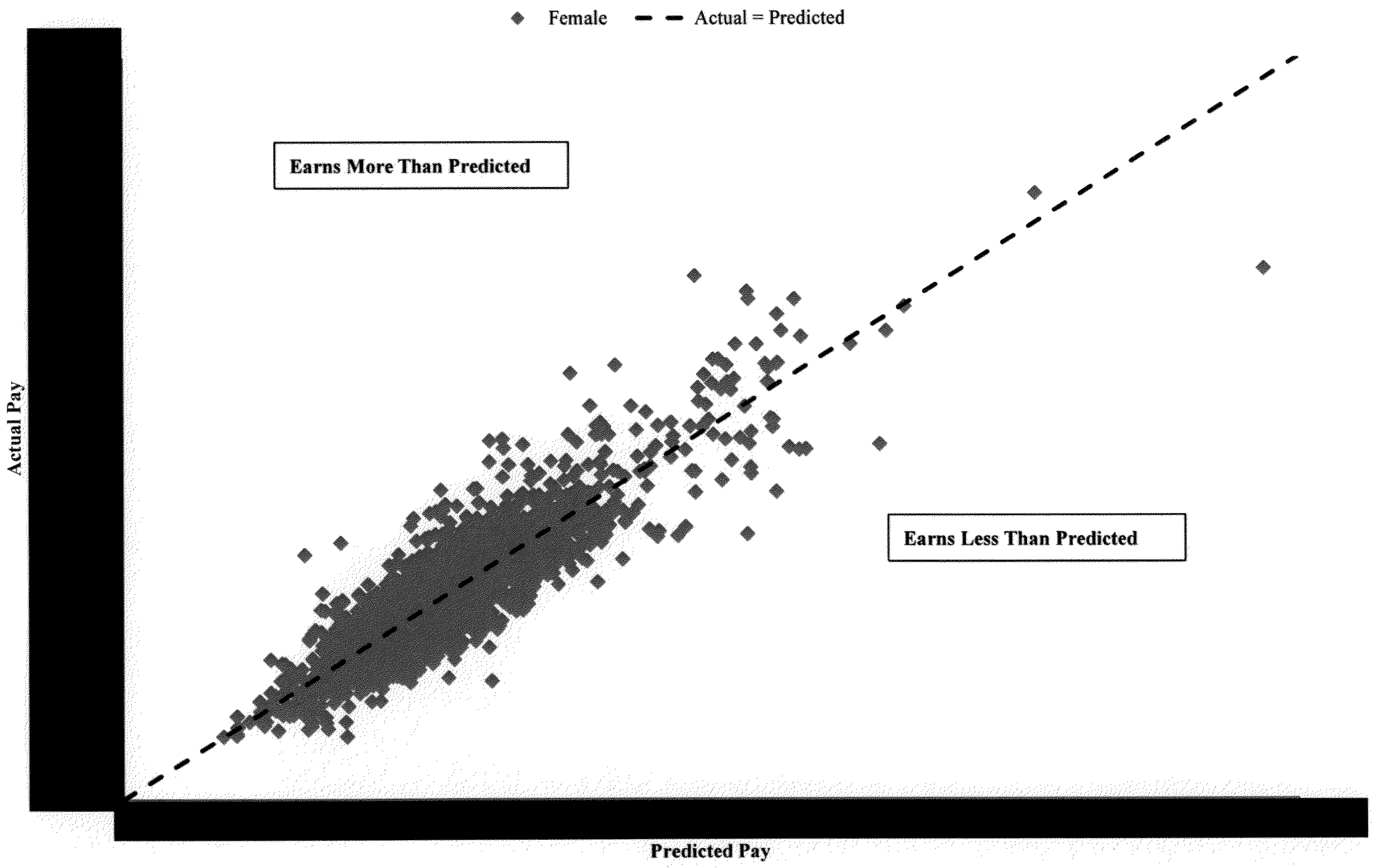
C36

Exhibit C37

2016: Actual Base Pay vs. Predicted Base Pay

- Prediction Based on Dr. Neumark's Data and Model, Without a Gender Control -

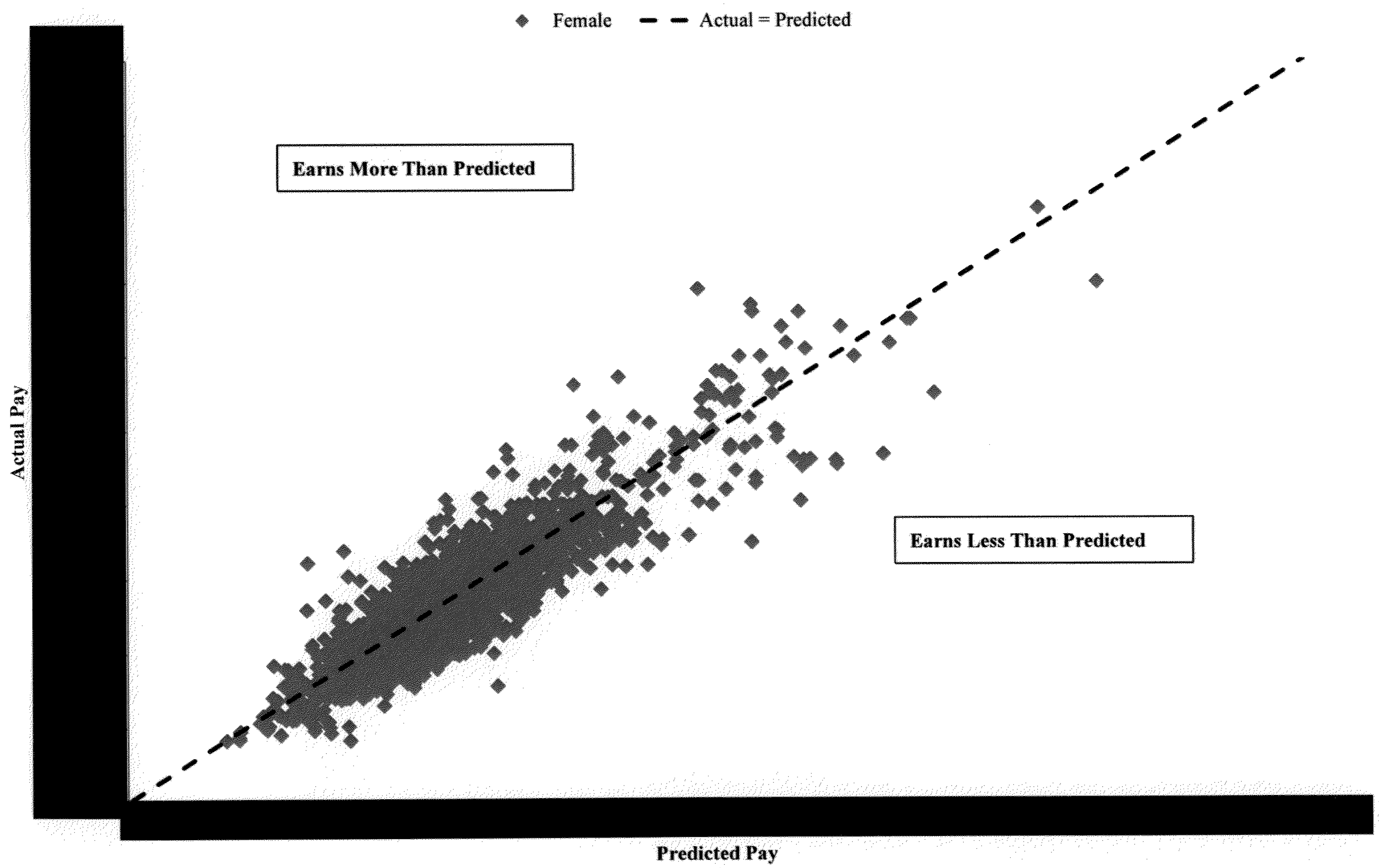
- Female Incumbents in Neumark's Dataset, 2013-2018 -



C37

Exhibit C38

2017: Actual Base Pay vs. Predicted Base Pay
- Prediction Based on Dr. Neumark's Data and Model, Without a Gender Control -
- Female Incumbents in Neumark's Dataset, 2013-2018 -



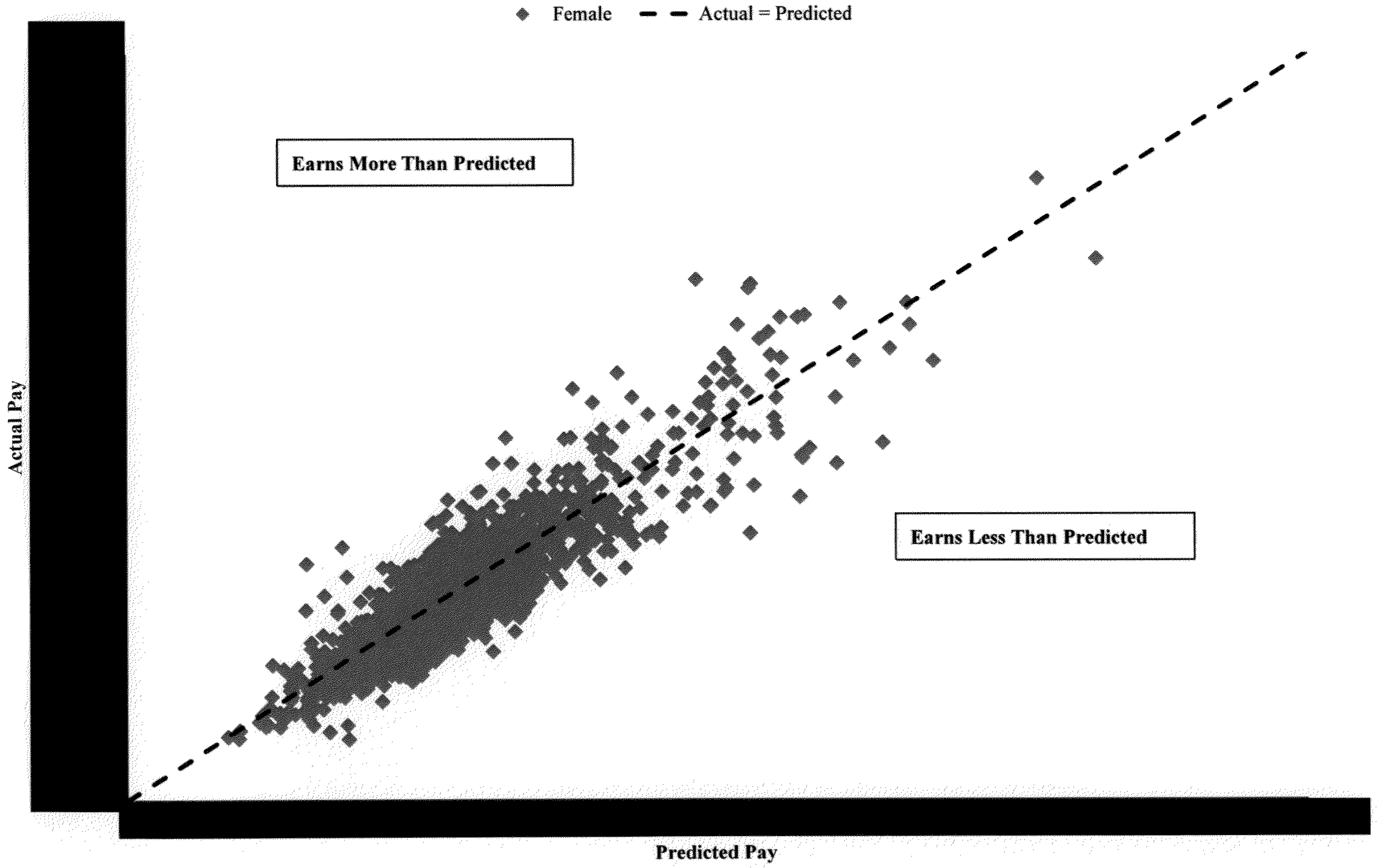
C38

Exhibit C39

2018: Actual Base Pay vs. Predicted Base Pay

- Prediction Based on Dr. Neumark's Data and Model, Without a Gender Control -

- Female Incumbents in Neumark's Dataset, 2013-2018 -



C39

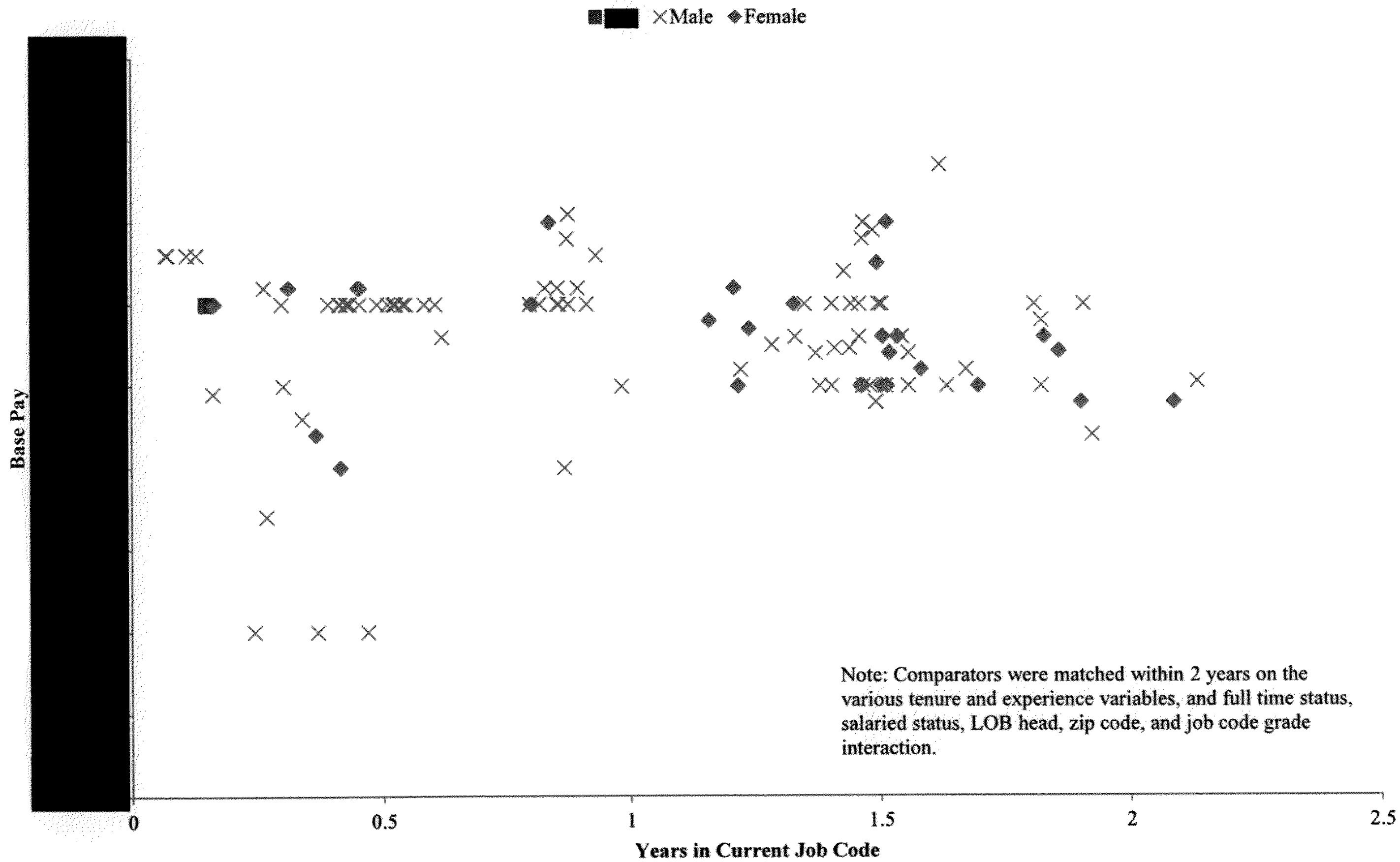
Attachment D: Matching Comparators

Exhibit D1

[redacted] and Matched Comparators Using Dr. Neumark's Regression Variables to Create the Match, 2016

- Software Developer 2, Grade E.6 -

- Average Age: 25.8 Years -



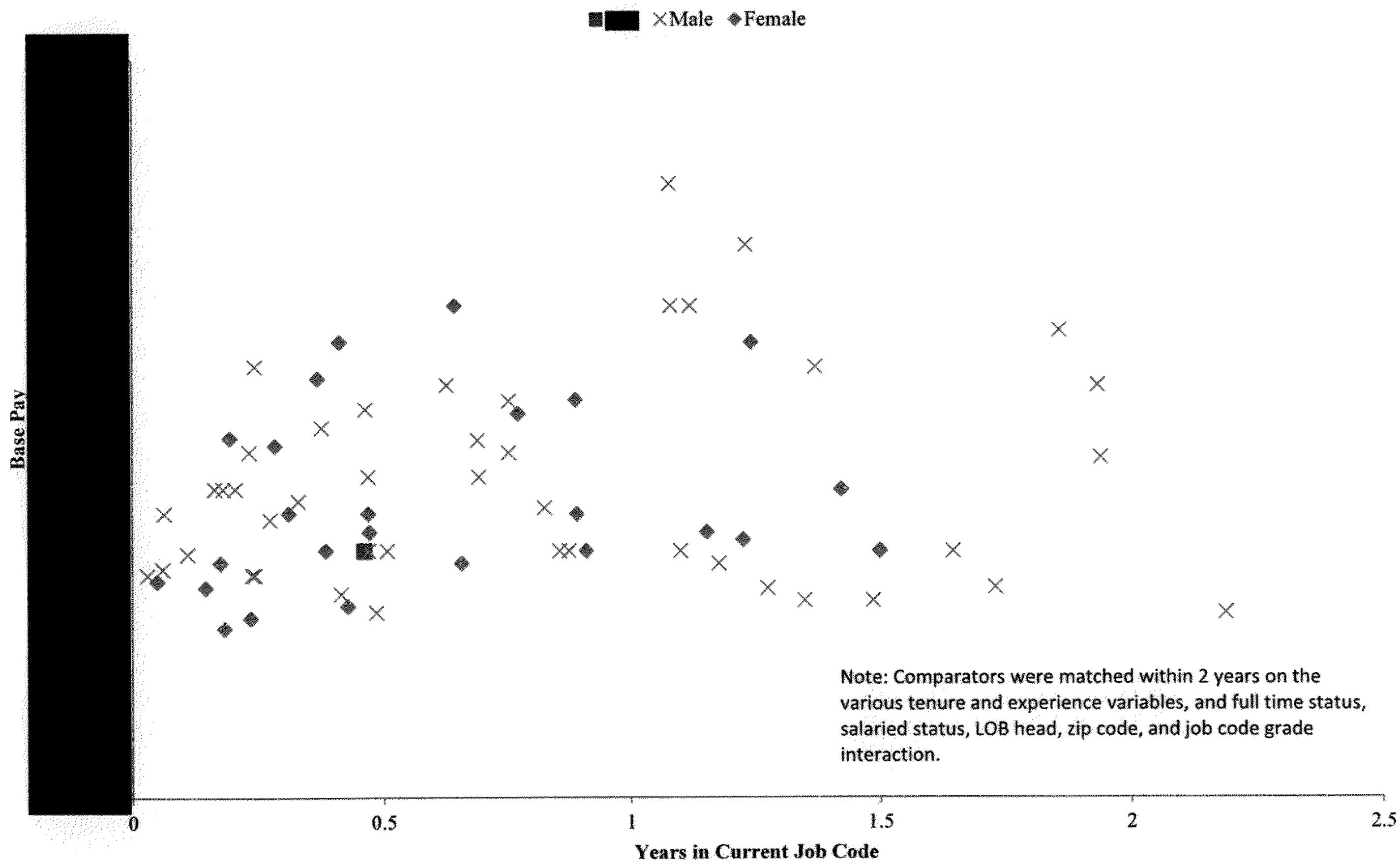
D1

Exhibit D2

[redacted] and Matched Comparators Using Dr. Neumark's Regression Variables to Create the Match, 2016

- Software Developer 3, Grade E.08 -

- Average Age: 26.0 Years -



D2

Exhibit D3

[redacted] and Matched Comparators Using Dr. Neumark's Regression Variables to Create the Match, 2016

- Software Developer 4, Grade E.9 -

- Average Age: 39.6 Years -

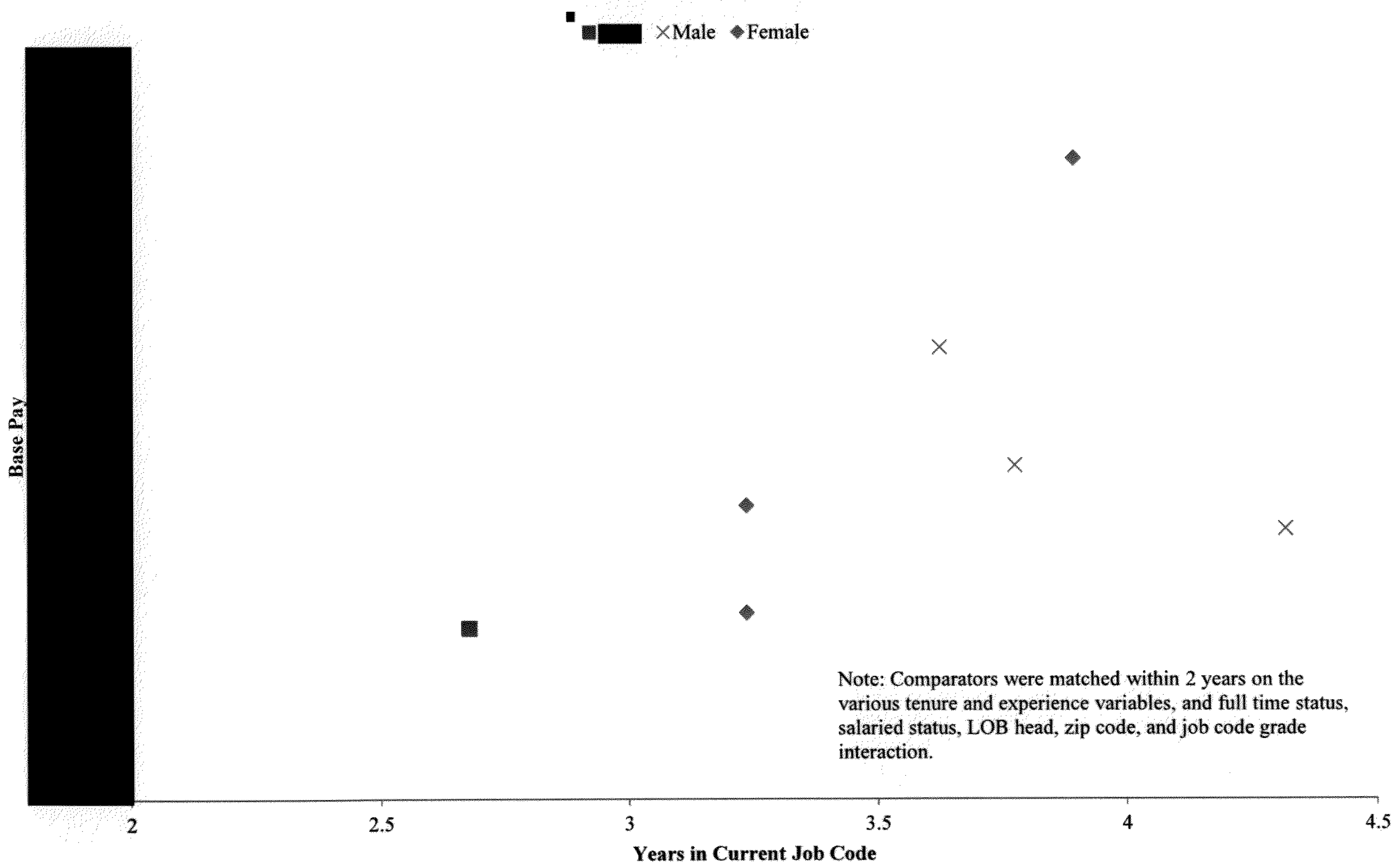


Exhibit D4

[redacted] and Matched Comparators Using Dr. Neumark's Regression Variables to Create the Match, 2016

- Software Developer 5, Grade E.11 -

- Average Age: 47.2 Years -

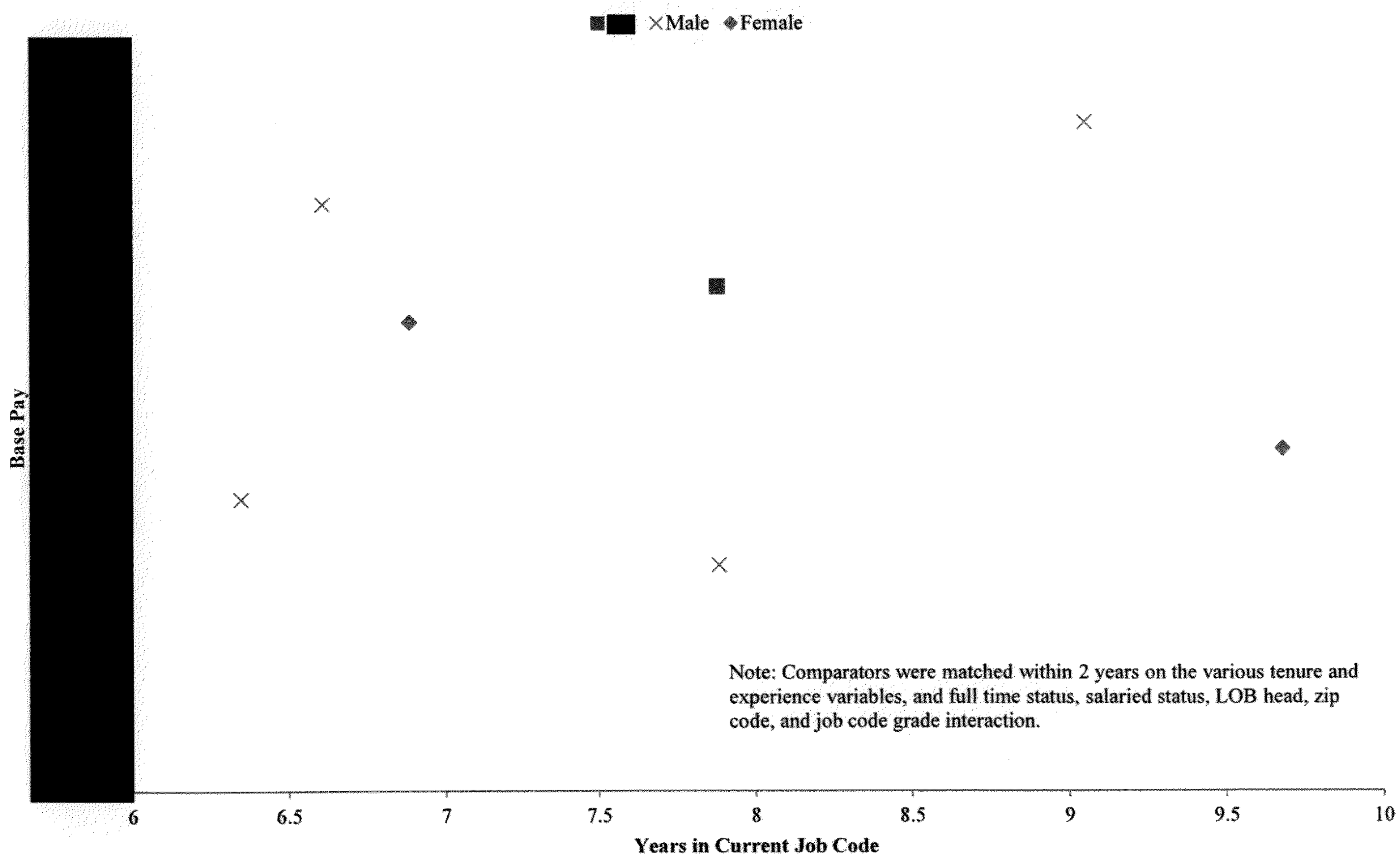
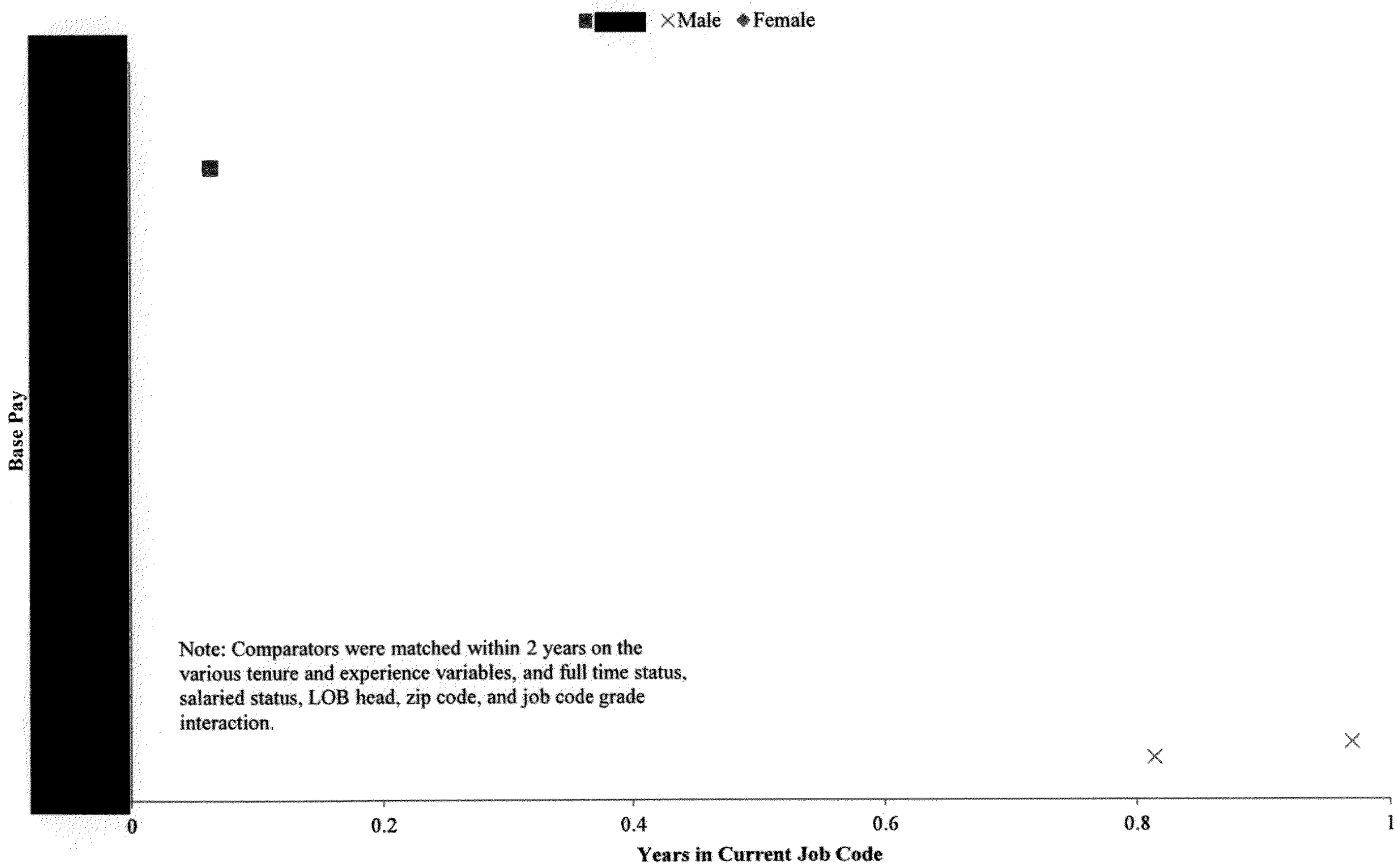


Exhibit D5

[redacted] and Matched Comparators Using Dr. Neumark's Regression Variables to Create the Match, 2016

- Software Development Snr Manager, Grade E.11 -

- Average Age: 36.8 Years -



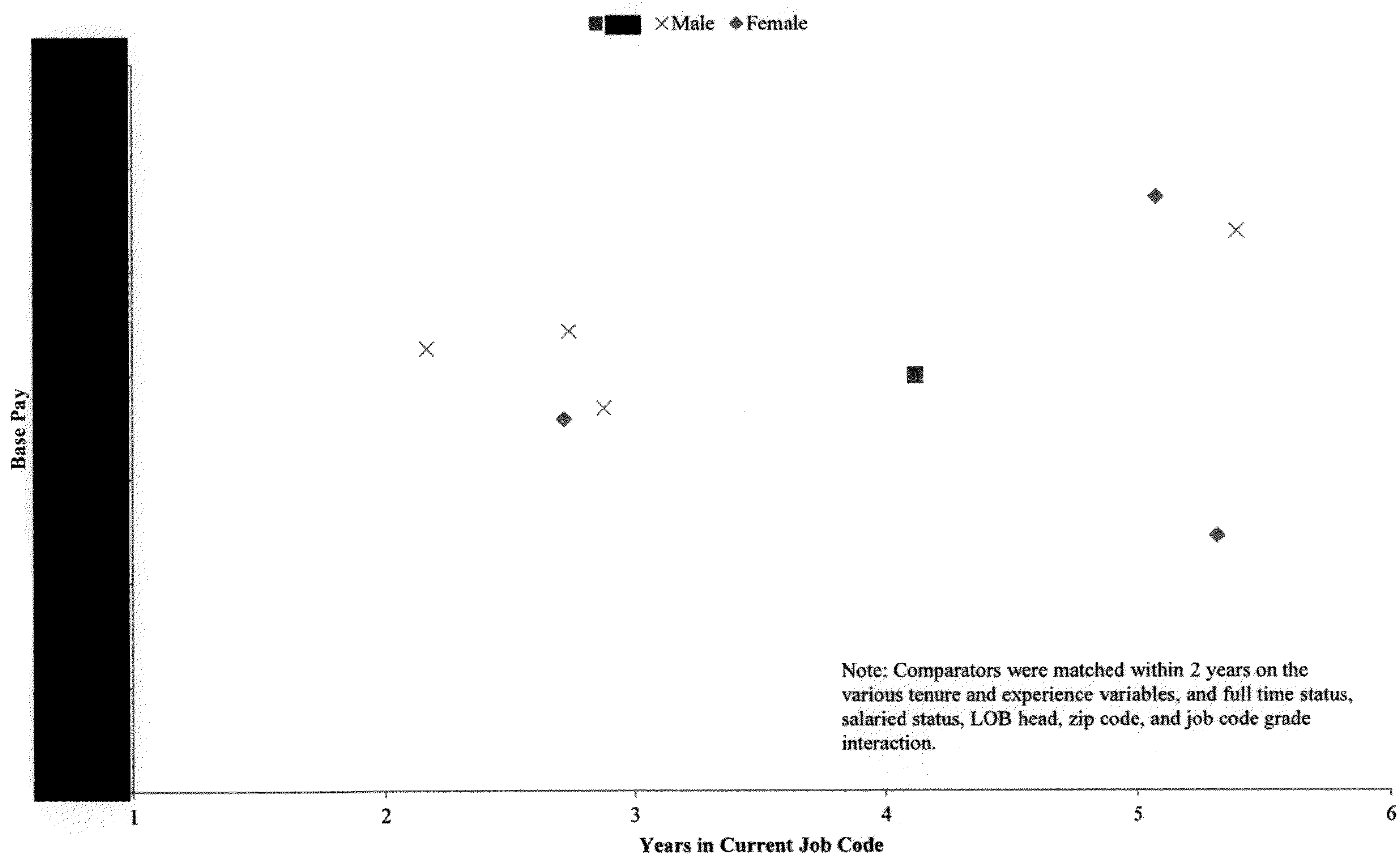
D5

Exhibit D6

[redacted] and Matched Comparators Using Dr. Neumark's Regression Variables to Create the Match, 2016

- Software Development Director, Grade E.12 -

- Average Age: 47.8 Years -



D6

Exhibit D7

[redacted] and Matched Comparators Using Dr. Neumark's Regression Variables to Create the Match, 2016

- Product Manager/Strategy 4-ProdDev, Grade E.09 -

- Average Age: 32.7 Years -

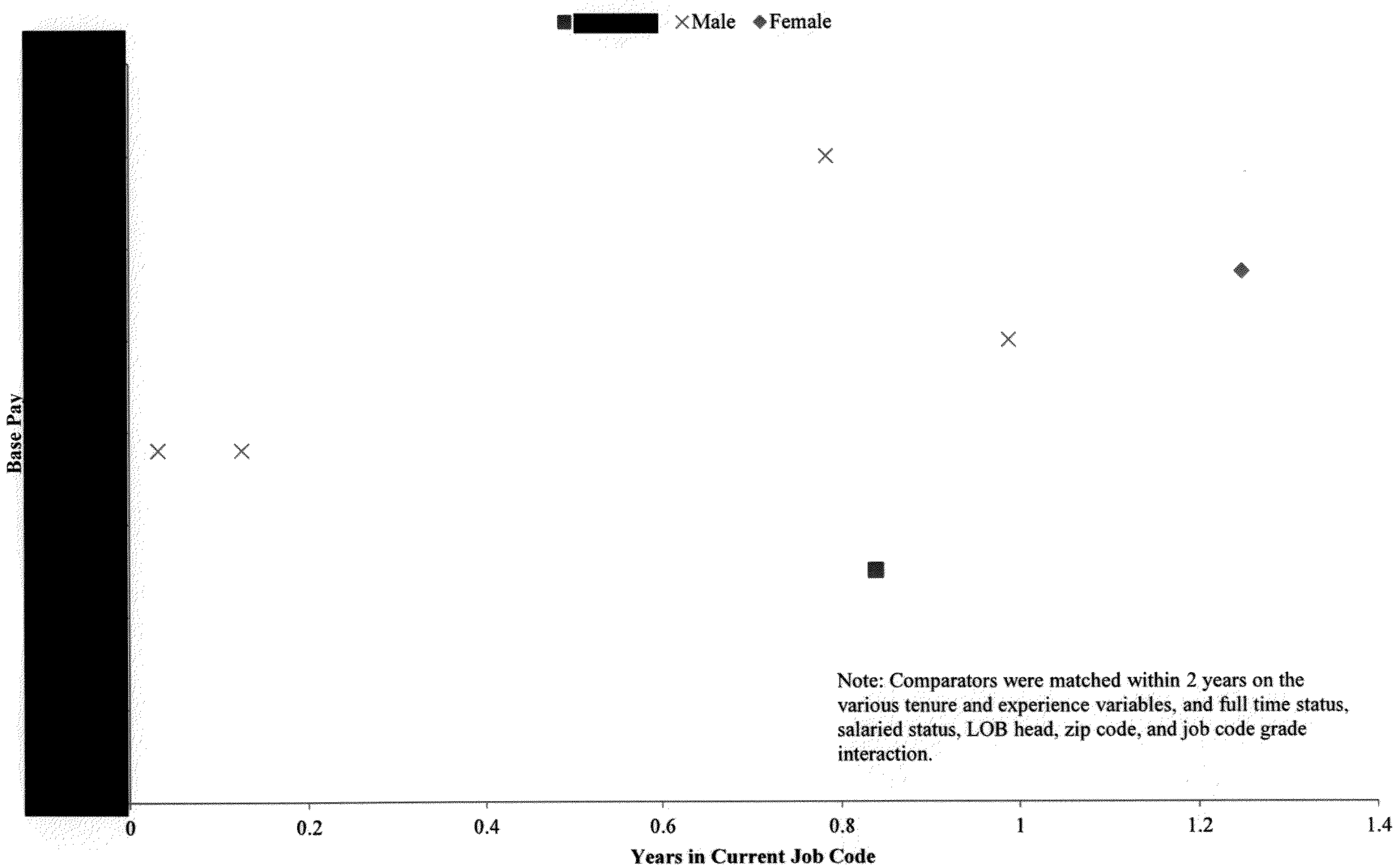


Exhibit D8

[redacted] and Matched Comparators Using Dr. Neumark's Regression Variables to Create the Match, 2016

- Product Manager/Strategy 5-ProdDev, Grade E.011 -

- Average Age: 43.2 Years -

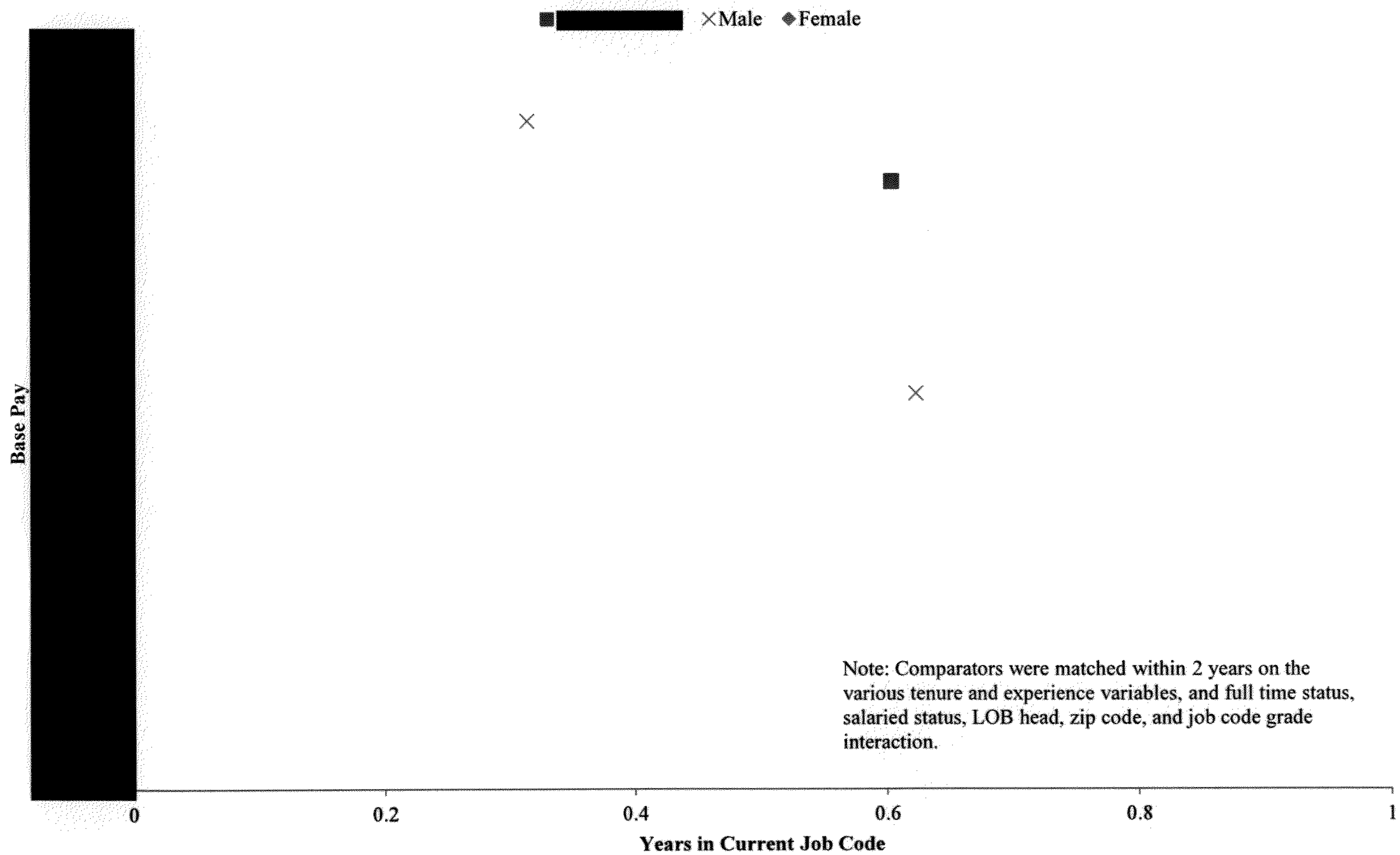


Exhibit D9

[redacted] and Matched Comparators Using Dr. Neumark's Regression Variables to Create the Match, 2016

- Technical Analyst 4-Support, Grade E.12 -

- Average Age: 52.1 Years -

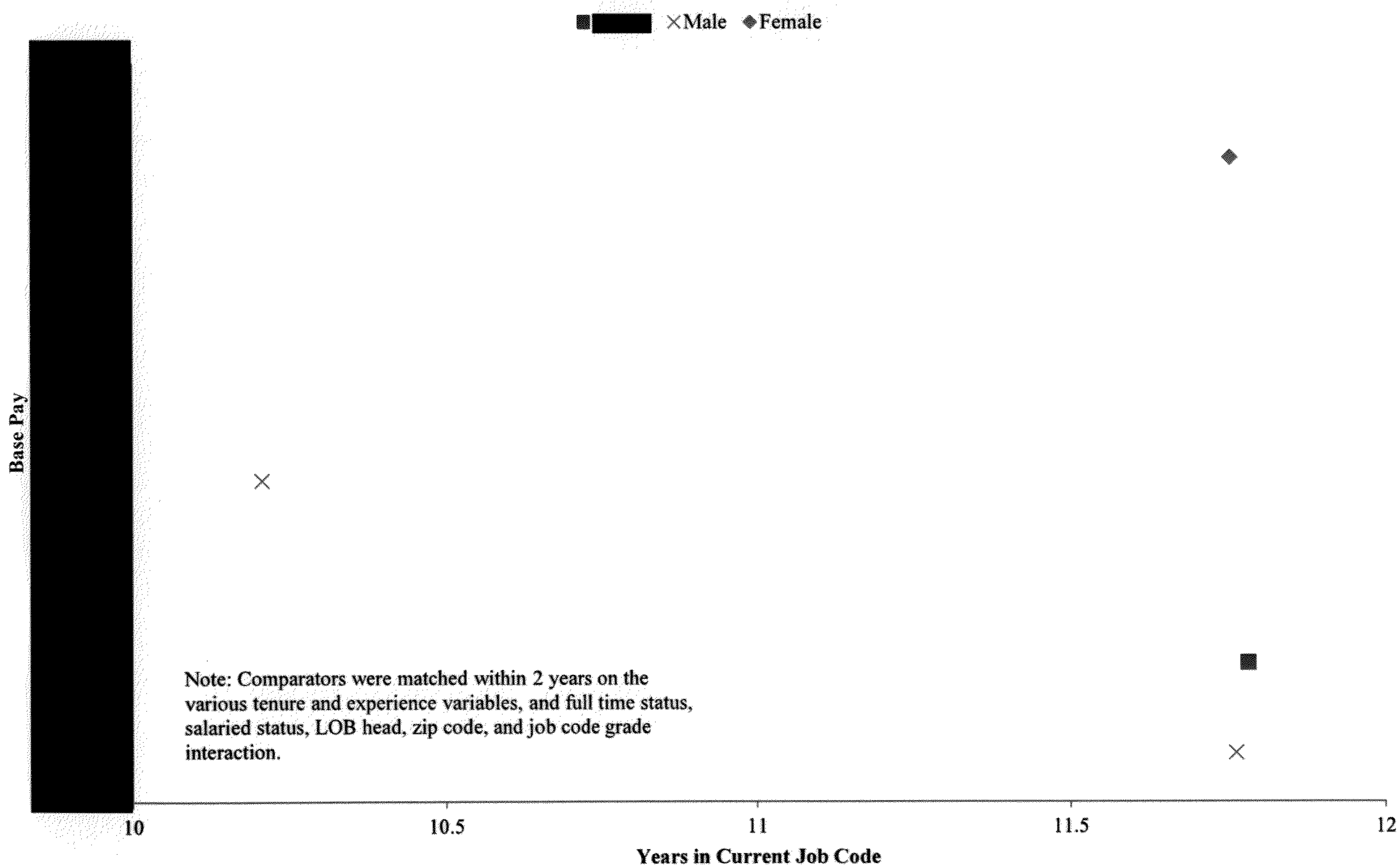
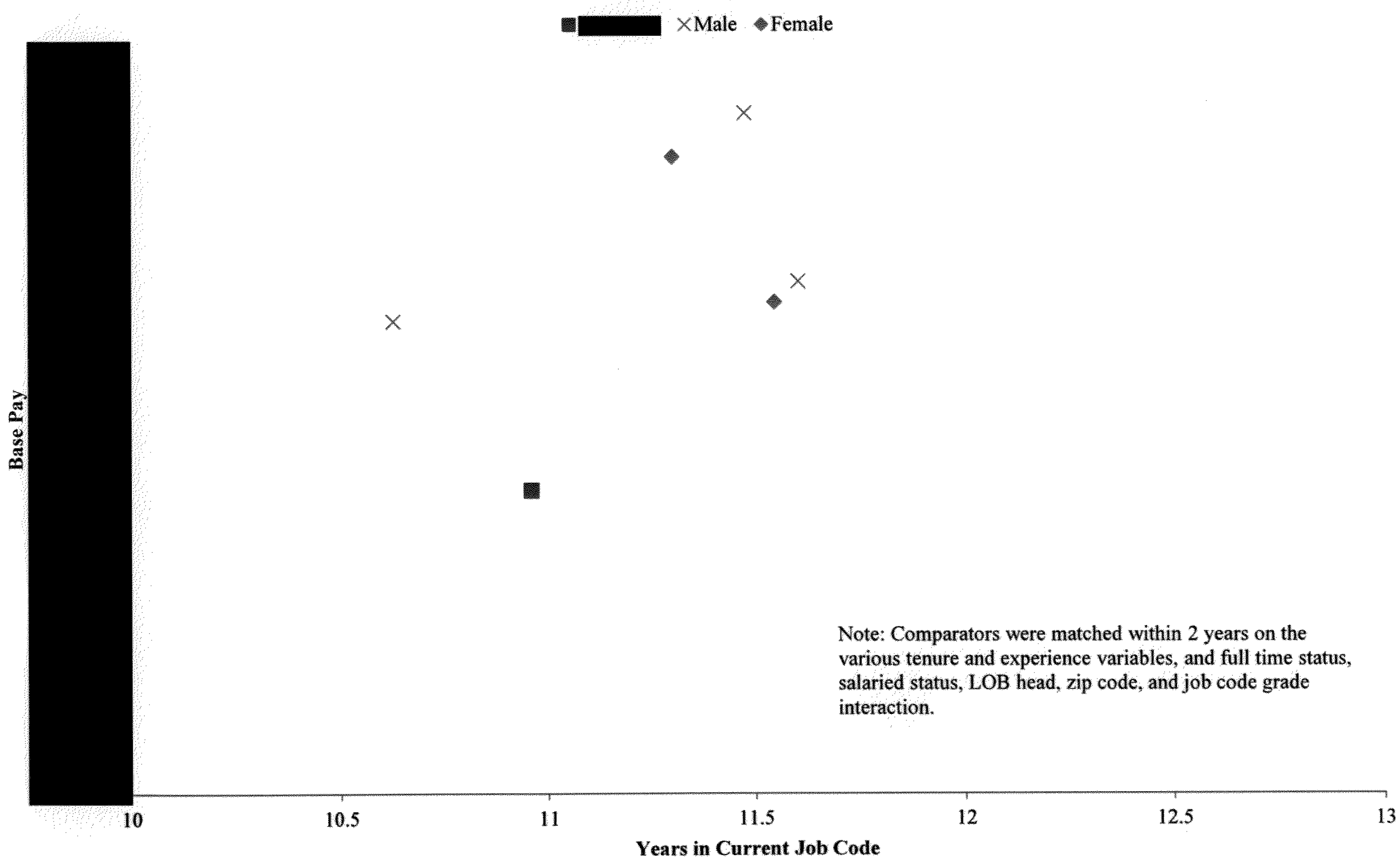


Exhibit D10

[redacted] and Matched Comparators Using Dr. Neumark's Regression Variables to Create the Match, 2016

-Applications Developer 4, Grade E.9 -

- Average Age: 44.9 Years -



D10

Attachment E: Clusters

Exhibit E1

Cluster 1

E1

Exhibit E2

Cluster 2

Exhibit E3

Cluster 3

E3

Exhibit E4

Cluster 4

Exhibit E5

Cluster 5

E5

Exhibit E6

Cluster 6

E6

Exhibit E7

Cluster 7

E7

Exhibit E8

Cluster 8

E8

Exhibit E9

Cluster 9

E9

Exhibit E10

Cluster 10

E10

Exhibit E11

Cluster 11

E11

Exhibit E12

Cluster 12

E12

Exhibit E13

Cluster 13

E13

Exhibit E14

Cluster 14

Exhibit E15

Cluster 15

E15