JANET M. HEROLD
Regional Solicitor
JEREMIAH MILLER
Acting Counsel for Civil Rights
LAURA C. BREMER
Senior Trial Attorney
Office of the Solicitor
UNITED STATES DEPARTMENT OF LABOR
90 7th Street, Suite 3-700
San Francisco, California 94103
Tel: (415) 625-7757
Fax: (415) 625-7772
Email: bremer.laura@dol.gov

UNITED STATES DEPARTMENT OF LABOR
OFFICE OF ADMINISTRATIVE LAW JUDGES

OFFICE OF FEDERAL CONTRACT COMPLIANCE PROGRAMS, UNITED STATES DEPARTMENT OF LABOR,	
Plaintiff,	Case No. 2017-OFC-00006
v.	
ORACLE AMERICA, INC.	
Defendant.	

RECEIVED
APR 10 2019
Office of Administrative Law Judges
San Francisco, Ca

OFCCP'S MEMORANDUM OF POINTS AND AUTHORITIES IN SUPPORT OF MOTION TO COMPEL HISTORIC DATA OF COMPARATOR EMPLOYEES

I. INTRODUCTION

OFFCP hereby moves to compel data necessary to analyze compensation claims for more than 70% of employees covered by this lawsuit. OFCCP cannot analyze the claims of protected class members absent historical compensation data for *all* employees, i.e class comparators, going back in time. Oracle has provided a biased subset of the full sample affected, namely employees who have stayed at Oracle. As detailed in the Second Amended Complaint ("SAC"), OFCCP alleges that Oracle discriminated against its employees by paying female, Asian, and Black employees[1] less than comparable male or White employees. SAC at ¶¶ 12-32. As part of this discrimination claim, OFCCP sought to identify and analyze the mechanisms through which discrimination is occurring, including the impact of prior pay on starting salaries (SAC ¶ 32) and the assignment of those employees to lower-paying positions on hire (SAC ¶¶ 19-21). Specifically, OFCCP alleged that the starting pay of female, Asian, and Black employees is depressed relative to their male and White peers (SAC ¶¶ 22-24), and that the pay of these groups is suppressed over their tenure at Oracle. (SAC ¶¶ 25-31).

OFCCP's expert requires historical compensation data for the full affected population—rather than the partial population sample Oracle provided—in order to create an unbiased statistical analysis of the impact of starting pay, and importantly here, the ongoing suppression of pay on wage disparities over time. This is central to OFCCP's claim that Oracle pays women Asians, and Blacks less on hire and the pay disparity worsens the longer the employee stays at Oracle. SAC ¶¶ 19-32. While Oracle provided the necessary historical data dating back to the

[1] OFCCP grouped Oracle employees based on the data provided, which OFCCP understands to be the result of employees completing self-identification forms at Oracle's request. OFCCP acknowledges that categories like "Asian" and "Black" are broad, and the gender binary category is quite narrow, and they do not fully reflect any individual's experience.

early 1980s for a particular subset of the affected group—employees hired any time before OFCCP's investigation who have *chosen to stay* with the company—Oracle refused to provide pay data for those employees' counterparts who left Oracle prior to 2013. The undesirable outcome of using the self-selecting subset of data in an analysis examining compensation at hire and over time is a statistically biased result that does not, for example, properly account for employees with lower initial salaries who quit at different rates than those with higher initial salaries. Accordingly, to fully analyze compensation data and the impact of starting pay, data for *all* employees who received compensation during the same period as the class members is critical.

Despite OFCCP's requests for production of this information, Oracle refuses to produce comparator data for *any* period of time for employees who quit prior to 2013, i.e. the data necessary to assess the impact of starting pay for 76.14% of covered employees. Failing to understand the consequences of "selection bias" in a proper statistical analysis, Oracle insists that its provision of starting pay for all employees on the payroll as of January 1, 2013 is sufficient.

To understand the problem and prejudice created by Oracle's position, consider the problem posed by trying to analyze the role of starting pay for any employee hired before 2013 – which is the situation for 5,705 of the 7,493 employees Oracle had in the Product Development, Support and Information Technology Job Functions between 2013 and very early 2017. To analyze the role, for example, of starting pay, the analysis would study a person's starting pay in relation to *the starting pay of all other employees hired in that year*.[2] However, because Oracle

[2] The reasoning described below is equally applicable to the analysis of starting global career level, wage growth, and compensation over time, as alleged in the SAC. To be clear, OFCCP seeks to analyze changes in compensation over time, not just starting salary.

refuses to provide the data regarding starting pay for the protected class' coworkers for any hire data prior to 2013, OFCCP lacks the necessary comparator data to run the analysis. Attempting to run the analysis without the full data means that OFCCP's analysis for a full 76.14% of the protected class as to starting pay would be impacted by what is known in statistical analysis as "selection bias."

Selection bias is the distortion or skewing of analysis caused by analyzing a subset of data which was "selected" rather than chosen randomly. As the federal courts often must understand complex statistical analysis, the Federal Judicial Center's Reference Guide on Statistics provides specific instruction on "selection bias," providing the following famous example of such bias:

> A well-known example of selection bias is the 1936 *Literary Digest* poll. After successfully predicting the winner of every U.S. presidential election since 1916, the *Digest* used the replies from 2.4 million respondents to predict that Alf Landon would win the popular vote, 57% to 43%. In fact, Franklin Roosevelt won by a landslide vote of 62% to 38%. [Citation omitted]. The *Digest* was so far off, in part, because it chose names from telephone books, rosters of clubs and associations, city directories, lists of registered voters, and mail order listings. [Citation omitted]. In 1936, when only one household in four had a telephone, the people whose names appeared on such lists tended to be more affluent. Lists that overrepresented the affluent had worked well in earlier elections, when rich and poor voted along similar lines, but the bias in the sampling frame proved fatal when the Great Depression made economics a salient consideration for voters.

Here, the selection bias infecting the data set Oracle provided thus far is the "selection" of the data subset as only those employees still on the payroll in 2013. According to Oracle, OFCCP can assess the impact of starting pay but it can only study the impact of starting pay in relation to other employees who also were still employed as of 2013. This elimination from the data set of the starting pay for all of these workers' colleagues save those who were also still on payroll as

of 2013 creates a classic example of selection bias, as there is nothing random about confining the comparator data to this group of people.

To understand concretely the problems created by this selection bias, consider two different possible scenarios involving two hypothetical women still on Oracle's payroll in 2013, but who had hire dates in 1995. In the data set provided thus far by Oracle, OFCCP could only compare the starting pay of these two women against the starting pay of each other and any other women and men also hired in Oracle in 1995 but who were also still on the payroll in 2013. OFCCP cannot analyze how these two women's starting pay compares to the full universe of their colleagues' starting pay because Oracle refuses to give the starting pay data for any worker no longer still employed by Oracle as of 2013. Thus, on one end of the extreme, it could be that the two remaining women hired in 1995 and still at Oracle in 2013 boasted some of the highest starting pay among their colleagues in general and among women, and thus decided to stay long-term while their lesser paid colleagues left for better-paying work. Conversely, it also could be that the two women hired in 1995 who remained at Oracle until 2013 had some of the lowest starting pay among their colleagues or for their gender group and, thus, those who left Oracle were those most highly-compensated or valued by Oracle. This broad range of possibilities in the full comparator data set obviously could skew any analysis of the starting pay or compensation over time for the class of women employed at the time (such as in 1995 or at any point from 1995 until 2013).

Given the wide range of possibilities as to what the analysis of starting pay would show with the complete data set for each member of the protected class' hiring class, OFCCP cannot conduct its analysis regarding starting pay for a full 76.14% of the protected class. What this means is that the only analysis OFCCP can complete currently without such pollution by

selection bias concerns only the 23.86% of the Oracle's workforce that was hired since 2013. In other words, OFCCP's analysis of the impact of starting pay is confined to analysis of a group which has a job tenure of 5 years or less. Significantly here, such a short timeframe drastically limits OFCCP's ability to evaluate how compensation changes over time as it seeks to do. There is no basis for restricting OFCCP's analysis or discovery to this small subset of Oracle's workforce.[3]

OFCCP is fully entitled to the production of this highly relevant and necessary information. As OFCCP has repeatedly explained to Oracle, OFCCP simply cannot analyze the impact of starting pay for 76.14% of the protected class without the data regarding the starting pay for their colleagues – Oracle cannot just "select" this subset of data without creating an unacceptable skewing of the statistical analysis.

OFCCP requires the immediate production of this data to permit OFCCP's expert to timely complete its analysis and expert report. OFCCP fully supports and is prepared and able to meet all current discovery and hearing deadlines, including deadlines for submission of the expert report. However, substantial work and analysis is required upon receipt of the class comparator data and OFCCP cannot afford Oracle to continue its delay in the production of this critical data. OFCCP asks the Court to order the immediate production of this essential

---

[3] Even more troublingly, as noted in the SAC, despite the obvious distortion created by selection bias (the scope of which is impossible to assess given the restricted data Oracle provided), OFCCP analyzed the effect of starting pay and job tenure using only the limited universe Oracle provided and the results of that analysis tell a story that is very unfavorable to Oracle. According to that analysis, comparing the data of protected class members only to the data of other class members as opposed to the full universe of their colleagues employed at the same time in the same positions, the amount of wage discrimination suffered by women and minorities at Oracle increases with every year worked. Thus, in addition to the fact that OFCCP is entitled to the complete data it seeks as all of the data requested is relevant and fully necessary for OFCCP to complete its analysis, it is simply not in Oracle's interest to continue to deny OFCCP this data. If OFCCP is confined to this limited data set, the story it tells about Oracle's pay practices is a story of ever-increasing discrimination. Oracle cannot later complain regarding inadequacies in OFCCP's statistical analysis arising from selection bias when Oracle refuses to provide the data necessary to eliminate that bias.

information. Alternatively, if the Court does not compel production of the full data set for class member comparators, OFCCP requests the Court prohibit any challenge to OFCCP's statistical analysis on the basis that it includes only a subset of the pertinent data.[4]

II. BACKGROUND

In October 2017, Oracle produced a subset of the data at issue, specifically historical data going back to the early 1980s for all employees still at Oracle during the period of 2013 to 2016, acknowledging that such historical data is relevant to an analysis of compensation. Decl. of Laura Bremer Decl. ("Bremer Decl."), at ¶ 3. Oracle produced this historical data for class members in response to OFCCP's Second Set of Requests for Production, which OFCCP served in February 2017. *Id.* at ¶¶ 2-3. The data produced, however, distinctly lacks the necessary comparator information to analyze the claims of the protected class: data for employees who worked at Oracle at the same time as class members but who left prior to 2013. *Id.* at ¶ 4. To ensure production of the critical comparator data for everyone employed during the same period as the class members, OFCCP specifically requested this same data for "every employee working in the Product Development, Information Technology AND Support Job Functions from January 1, 1985, through AND including December 31, 2012, whether OR not they were employees in one of these Job Functions after January 1, 2013." *Id.* at ¶5, Exh. 4. During meet and confer on this request (RFP 178), OFCCP asked that Oracle specifically update the following files with the historic information sought by OFCCP's request: [5]

---

[4] In *Moussouris v. Microsoft Corp.*, Microsoft, while represented by the same counsel that is handling the data requests here, refused to produce the entire universe of complaints plaintiffs requested then criticized plaintiffs for relying on the smaller set of complaints. 2018 WL 3584701 at *9 (W.D. Wash. July 11, 2018). The court stated it may evaluate Microsoft's position in the context of the withheld information. *Id.*

[5] In order to address any future argument construing OFCCP's request for historical compensation data more narrowly than OFCCP intended, OFCCP included additional requests in its Sixth Set of Requests for Production naming the specific files OFCCP needs Oracle to supplement. *Id.* at ¶ 8, Exh. 8.

- Emp_Personal_Experience_Qualification_Assign_Details.xlsx
- Merged Assignment History, Medicare and Sal Admin.xlsx
- gsi_cwb_detail.xlsx
- gsi_comp_history.xlsx
- AllEarnings.xlsx
- AllEarnings2.xlsx
- Appraisal_Audit_All_Data.xlsx
- HQCA_IREC_DATA.xlsx
- Any files that Oracle produces that include additional data regarding education and experience

Id. at ¶ 7, Exh. 7.

The Parties have fully met and conferred on the issue of the historical comparator data that OFCCP seeks and have reached an impasse. *Id.* at ¶ 6, Exh. 6. [6]

III. ARGUMENT

A. Historical compensation data for all class comparators is necessary to analyze claims of the protected class.

Historical compensation data for all class comparators, which includes employees outside the class period, is central to OFCCP's core allegations. Oracle produced a biased subset of the relevant historical data and OFCCP cannot currently analyze the claims of the *protected class* with the information Oracle has provided. Oracle's data shows that it had 7,493 employees in the Product Development, Support and Information Technology Job Functions between 2013 and very early 2017. Decl. Jeremiah Miller ("Miller Decl.") at ¶ 4.[7] Oracle has produced starting

[6] In a sincere attempt to come to an agreement on this issue and avoid filing this motion, OFCCP proposed a compromise during the prolonged meet and confer process by which it would agree to limit the data production for employees who left Oracle prior to 2013 to a period of 2005 to 2013. *Id.* Exh. 6 (Bremer Ltr. March 20, 2019). The smaller data set OFCCP proposed for purposes of coming to an agreement, however, does not allow for the full analysis of compensation over time that OFCCP seeks to conduct to support its claims. As such, OFCCP stated in its March 20, 2019 meet and confer letter that it would seek to compel the historical data from 1985 to 2013 required for a complete analysis. *Id.*

[7] Oracle committed to providing additional compensation data to January 19, 2019. This supplementation should increase the total number of employees at issue in this matter. However, it will not increase the number of employees for whom OFCCP requests historic information, since Oracle's supplementation for 2017 to 2019 only includes employees who were not employed by Oracle prior to 2017.

compensation information and compensation data over time information only for 23.86% (1,788) of those individuals. For the remaining 76.14% (5,705) of those employees, Oracle has refused to provide starting pay information or compensation tenure information for *all* of their comparators. This data is critical to full understanding of Oracle's discriminatory compensation practices.

1. OFCCP seeks all comparator data for starting compensation and compensation over time for employees covered by this action.

OFCCP requested compensation data for all employees employed in Product Development, Support and Information Technology Job Functions from 1985 to the present. Importantly, OFCCP is not merely interested in starting pay; rather it seeks to analyze how tenure at Oracle impacts pay disparities. OFCCP requests the Court compel the full data set OFCCP requested in order to conduct a complete statistical analysis.

2. Oracle's selective production of comparator data undermines analysis of starting compensation, and compensation over time.

To this point, Oracle has provided compensation information for all employees who were employed in the three job functions at issue in this case in the 2013 to 2016 period, including 7,493 individuals. Decl. of Jeremiah Miller. at ¶ 4. An analysis of this data shows that these individuals have hire dates beginning in 1980, see Table 1:

Year	Hires	Cumulative %	% With Complete Comparator Data
1980	1	0.01	100
1981	1	0.03	99.99
1982	0	0.03	99.97
1983	1	0.04	99.97
1984	4	0.09	99.96
1985	3	0.13	99.91
1986	5	0.20	99.87
1987	13	0.37	99.80
1988	19	0.63	99.63
1989	49	1.28	99.37
1990	41	1.83	98.72
1991	21	2.11	98.17
1992	50	2.78	97.89
1993	80	3.84	97.22
1994	126	5.53	96.16
1995	156	7.61	94.47
1996	268	11.18	92.39
1997	257	14.61	88.82
1998	311	18.76	85.39
1999	308	22.87	81.24
2000	330	27.28	77.13
2001	316	31.50	72.72
2002	84	32.62	68.50
2003	84	33.74	67.38
2004	144	35.66	66.26
2005	285	39.46	64.34
2006	601	47.48	60.54
2007	238	50.66	52.52
2008	270	54.26	49.34
2009	146	56.21	45.74
2010	449	62.20	43.79
2011	416	67.76	37.80
2012	628	76.14	32.24
2013	442	82.04	23.86
2014	453	88.08	17.96
2015	450	94.09	11.92
2016	433	99.87	5.91
2017	10	100.00	0.13

Table 1: Hires by year for employees employed in Product Development, Support and IT functions between 2013 and early 2017 at Oracle. Green cells represent current complete comparator data.

For any employee appearing on Oracle's payroll in 2013, Oracle has provided full compensation data back to the early 1980s. Thus, for all employees hired from 2013 to early 2017 (23.86% of those covered by this action), this data permits comparison of starting compensation and compensation tenure information for the comparable employees in that time period.

However, for employees hired prior to 2013, Oracle refuses to provide the comparator data. Oracle refuses to provide starting and compensation data tenure for any employees hired into the same functions at the same time as a member of the protected class, but who left employment with Oracle *before* 2013. Thus, OFCCP cannot compare the starting pay and tenure compensation data for 76.14% of the protected class to the full universe of their comparators, their colleagues.

Oracle's refusal to provide all comparators for employees covered by this suit introduces a selection bias into the analysis, leaving only employees who happened to remain employed to be analyzed. In contrast to the data set Oracle has produced from 2013-2017, OFCCP cannot include data for hires-to-departure employees if that departure occurred before 2013.

Oracle's refusal to produce data for all employees hired at the same time as employees still employed between 2013 and early 2017 introduces significant selection bias.[8] Analyzing only employees who remained at Oracle past 2013 is the equivalent of assessing the viability of businesses by only examining the ones that do not fail, or concluding that college dropouts are wildly successful by focusing on billionaires and ignoring the millions of other impoverished

---

[8] *See generally* David H. Kaye and David A. Freedman, *Reference Guide on Statistics,* in REFERENCE MANUAL ON SCIENTIFIC EVIDENCE 211 (3d ed. 2011) ("Federal Judicial Center Reference Guide", available at https://www.fjc.gov/content/reference-guide-statistics-2 (last accessed April 6, 2019)); *see also* H. Gregg Lewis, *Comments on Selectivity Biases in Wage Comparisons*, 82 Journal of Political Economy 1145 (1974).

dropouts.[9] In light of OFCCP's allegations of widespread discrimination in these job functions it would be unsurprising to find that the selection bias at play here artificially reduced the number of female, Black and Asian employees among those long-serving employees, impacting OFCCP's analysis. The loss of protected class members from the data sets could have a disastrous impact on the analysis of all employees employed between 2013 and early 2017; removing their compensation data would under-represent the effect of belonging to those protected classes on starting pay and the suppression of pay over time.

Oracle's refusal to supply the historical data OFCCP requested impairs the analysis of more than 70% of the employees covered by this lawsuit. This Court should order Oracle to produce the data it possesses that would permit an appropriate analysis of all employees covered by this action.

B. Oracle's objections to producing the requested historical data of class comparators are meritless and cannot outweigh OFCCP's need for the information.

In refusing to produce comparator data, Oracle contends that: (1) the comparator employees are outside the population at issue; and (2) producing the historical data is burdensome.[10] Bremer Decl., Exh. 6. Neither argument has merit and cannot justify denying OFCCP the information it needs to prove its case.

First, contrary to Oracle's assertion, employees who left Oracle prior to 2013 are not entirely separate from the population at issue. Rather, as fully detailed above, these employees

---

[9] *See, e.g.*, https://www.scientificamerican.com/article/how-the-survivor-bias-distorts-reality/ (last accessed on April 5, 2019); https://www.theatlantic.com/business/archive/2013/03/the-myth-of-the-successful-college-dropout-why-it-could-make-millions-of-young-americans-poorer/273628/ (last accessed April 5, 2019).

[10] Oracle also raised the privacy rights of employees unconnected to the claims as a basis to withhold the information. Bremer Decl., Exh. 6 (Mantoan Ltr March 29, 2019). However, not only will OFCCP use the data in aggregated form such that it cannot be tied to any individual, the Protective Order in place in this litigation addresses this concern.

are an important subset of the full affected population and, while not included in the class or OFCCP's compensation claim, they are necessary to *analyze* the claims of the *protected class*. OFCCP seeks the same data fields for former employees (located in the same databases) as Oracle already identified and produced for the employees at Oracle during the 2013 to 2016 period. Further, it is of no moment that the information requested is from a period prior to OFCCP's compliance review as it is highly relevant—indeed essential—to claims at issue in this litigation. *See OFCCP v. Uniroyal, Inc.*, No. OFCCP 1977-1, at 9 (Sec'y June 28, 1979) ("discovery is not limited to the issues raised by the pleadings and that the correct test for the scope of discovery is relevancy to subject matter of the suit."); *see also, U.S. Dep't of Labor v. Prudential Ins. Co.*, 1980 WL 275523, *8 (July 27, 1980) (allowing OFCCP to obtain historical data prior to review period finding "[t]he right to obtain information concerning past practices to determine present violations is not inconsistent with immunity to…liability for violations found to have existed [prior to review period]."); *U.S. Dep't of Labor v. Harris Trust & Savings Bank*, 78 OFC-2 (ALJ Decision Dec. 22, 1986) (OFCCP obtained data including the initial salaries and promotions of employees outside to the review period).

Second, while it is unclear clear why production of the requested data would be unduly burdensome for a company with Oracle's resources specializing in data management, OFCCP is entitled to this information despite any burden Oracle might incur. OFCCP does not dispute that Oracle has produced a significant amount of data in this litigation. However, this is largely a function of Oracle's sheer size and number of employees. Oracle cannot use its size as global multibillion-dollar company to shield itself from its obligation to produce information as a government contractor, or as a party to this litigation.

There can be no credible assertion here that OFCCP's request for production of the class comparator data is in any way "disproportionate" to the scope of the case or damages at issues. In cases such as this one, which seek to "vindicate vitally important . . . public values," proportionality is assessed in light of the important public policy issues at stake. Fed. R. Civ. P. 26 (Cmte Notes, 2015 Amend). Indeed, courts recognize that given the complexity of issues involved and the employer's control over key evidence, plaintiffs should generally be afforded "liberal discovery" in cases to redress employment discrimination. *Finch v. Hercules, Inc.*, 149 F.R.D. 60, 62 (D. Del. 1993) (citations omitted); *see also, Chen-Oster v. Goldman, Sachs & Co.*, 285 F.R.D. 294, 308 (S.D.N.Y 2012) (compelling production of computerized data despite defendant's purported burden in producing information from multiple databases).

This case not only implicates important public policy interests, including whether more than $100 million each year in taxpayer monies is being used to subsidize a government contractor engaging in systemic race and gender discrimination, but also directly affects the interests of thousands of employees at Oracle's headquarters. OFCCP has already identified that its analysis of wages due through 2016 alone exceeds $400 million. SAC ¶¶ 14-16. OFCCP's request for data that goes to the very heart of its claims on behalf of these employees is in no way disproportionate. OFCCP should not be denied the information it needs to support its compensation claims regardless of any purported burden in producing the data.

IV. CONCLUSION

For the reasons stated above, the Court should grant OFCCP's motion and compel production of historical data for comparator employees for the period of 1985 through 2012. In the event the Court does not compel production of the full data set for class member

comparators, the Court should prohibit any challenge to OFCCP's statistical analysis based on any subset of data Oracle provides.

Dated: April 10, 2019

Respectfully submitted,

KATE S. O'SCANNLAIN
Solicitor of Labor

JANET M. HEROLD
Regional Solicitor

JEREMIAH MILLER
Counsel for Civil Rights

LAURA C. BREMER
Senior Trial Attorney