

2018 WL 2124019

Only the Westlaw citation is currently available.
United States District Court, W.D. Washington,
at Seattle.

Katherine MOUSSOURIS, et al., Plaintiffs,
v.
MICROSOFT CORPORATION, Defendant.

CASE NO. C15-1483JLR

Signed 04/25/2018

Attorneys and Law Firms

Adam T. Klein, Pro Hac Vice, Elizabeth V. Stork, Pro Hac Vice, Daniel Stromberg, Pro Hac Vice, Michael C. Danna, Pro Hac Vice, Ossai Miazad, Pro Hac Vice, Rachel Bien, Pro Hac Vice, Outten & Golden LLP, New York, NY, Rachel J. Geman, Pro Hac Vice, Lieff Cabraser Heimann & Bernstein, New York, NY, Anne B. Shaver, Pro Hac Vice, Kelly M. Dermody, Pro Hac Vice, Michael Levin-Gesundheit, Pro Hac Vice, Michelle Lamy, Pro Hac Vice, Tiseme Zegeye, Pro Hac Vice, Lieff Cabraser Heimann & Bernstein, San Francisco, CA, Pamela Disney, Pro Hac Vice, Outten & Golden LLP, Washington, DC, Paul W. Mollica, Pro Hac Vice, Outten & Golden LLP, Chicago, IL, Michael C. Subit, Frank Freed Subit & Thomas, Sharon M. Lee, Lieff, Cabraser, Heimann & Bernstein, LLP, Seattle, WA, for Plaintiffs.

Allison Riechert Giese, Pro Hac Vice, Jessica R. Perry, Pro Hac Vice, Joseph Liburt, Pro Hac Vice, Lynne C. Hermle, Pro Hac Vice, Megan M. Lawson, Pro Hac Vice, Orrick Herrington & Sutcliffe, Menlo Park, CA, Jill L. Rosenberg, Pro Hac Vice, Marc R. Shapiro, Pro Hac Vice, Orrick Herrington & Sutcliffe, New York, NY, Lauri A. Damrell, Pro Hac Vice, Orrick, Herrington & Sutcliffe LLP, Sacramento, CA, Wendy Butler Curtis, Pro Hac Vice, Orrick Herrington & Sutcliffe, Washington, DC, Mark Steven Parris, Orrick Herrington & Sutcliffe LLP, Seattle, WA, for Defendant.

ORDER ON MOTIONS TO EXCLUDE

JAMES L. ROBERT, United States District Judge

I. INTRODUCTION

*1 Before the court are three motions to exclude filed by the parties: (1) Defendant Microsoft Corporation's ("Microsoft") motion to exclude Dr. Henry S. Farber's expert opinions (Farber Mot. (Dkt. # 362)); (2) Plaintiffs Katherine Moussouris, Holly Muenchow, and Dana Piermarini's (collective, "Plaintiffs") motion to exclude certain expert opinions of Dr. Ali Saad (Saad Mot. (Dkt. # 364)); and (3) Plaintiffs' motion to exclude Ms. Rhoma Young's expert opinions (Young Mot. (Dkt. ## 367 (sealed), 368 (redacted))). The court has reviewed the parties' filings in support of and in opposition to the motions, the relevant portions of the record, and the applicable law. Being fully advised,¹ the court DENIES Microsoft's motion to exclude Dr. Farber's opinions, GRANTS in part and DENIES in part Plaintiffs' motion to exclude Dr. Saad's opinions, and GRANTS Plaintiffs' motion to exclude Ms. Young's opinions.

II. BACKGROUND

Plaintiffs filed this putative class action to challenge Microsoft's "continuing policy, pattern, and practice of sex discrimination against female employees in technical and engineering roles ... with respect to performance evaluations, pay, promotions, and other terms and conditions of employment." (SAC (Dkt. # 55) ¶ 1.) As a result of these alleged policies and practices, Plaintiffs claim that female technical employees "receive less compensation and are promoted less frequently than their male counterparts." (*Id.* ¶ 3; *see also id.* ¶ 25 ("Microsoft discriminates against female technical employees in (1) performance evaluations; (2) compensation; and (3) promotions.")) Plaintiffs additionally allege that Microsoft "retaliates against female technical employees who complain about this discrimination." (*Id.* ¶ 1.)

On October 27, 2017, Plaintiffs filed a motion to certify a proposed class of women employees in Stock Levels 59-67² who work in the Engineering and/or the I/T Operations Professions from September 16, 2012, to the present. (Mot. for Class Cert. (Dkt. ## 228 (sealed), 232 (redacted)) at 1.) Specifically, Plaintiffs argue that Microsoft maintains a "common, discriminatory pay and promotions process"—the "Calibration Process" or "People Discussion Process"—that results in lower pay and fewer promotions for women. (*Id.*) To support their claim that gender-based differentials in pay and promotions result from Microsoft's Calibration Process,

Plaintiffs rely upon the statistical analysis performed by Dr. Farber. (*See id.* at 2, 5–10.)

*2 Microsoft opposes class certification. (*See* Class Cert. Resp. (Dkt. ## 286 (sealed), 285 (redacted).) In its opposition, Microsoft relies on the statistical analysis performed by Dr. Saad to challenge Dr. Farber's conclusions and to establish that no significant gender-based disparity exists in either pay or promotion. (*See id.* at 21, 23–28.) Microsoft also relies on Ms. Young's evaluation of Microsoft's Employment Relations Investigation Team (“ERIT”) to bolster the efficacy of ERIT as a tool Microsoft employs against discrimination. (*Id.* at 35.)

Subsequently, both parties filed motions to exclude. Microsoft challenges the admissibility of Dr. Farber's opinions, and Plaintiffs challenge the admissibility of some of Dr. Saad's opinions and the entirety of Ms. Young's opinions. (*See* Farber Mot.; Saad Mot.; Young Mot.) The court summarizes the relevant portions of each expert's opinions in turn.

A. Dr. Farber

Dr. Farber is the Hughes–Rogers Professor of Economics at Princeton University, where he has served on the faculty since 1991. (Farber Rep. ¶ 1.) He received a Ph.D. in economics from Princeton University, a Master of Science in Industrial and Labor Relations from Cornell University, and a B.S. in economics from Rensselaer Polytechnic Institute. (*Id.*) Dr. Farber teaches the analysis of wages, hours, and other issues in labor economics, as well as econometrics, which is the application of statistics to economics problems. (*Id.*)

Dr. Farber analyzed whether there is statistical evidence of discrimination in compensation or advancement rates between male and female technical employees in the relevant Stock Levels. (*Id.* ¶ 4.) After analyzing various data on Microsoft employees from January 1, 2010, through May 31, 2016 (*see id.* ¶¶ 12–13), Dr. Farber concludes that female employees in the putative class “are paid less than otherwise similar men, on average, and the average difference in pay is statistically significant.” (*Id.* ¶ 5.) Dr. Farber further concludes that “women in the class lag behind men in their rate of advancement at Microsoft.” (*Id.* ¶ 7.)

To reach these conclusions, Dr. Farber utilized three main statistical techniques. (*Id.* ¶ 29.) First, Dr. Farber analyzed pay differentials using a multiple regression analysis. (*Id.* ¶ 34.) A multiple regression analysis produces a numerical estimate, called a “coefficient,” which measures the relative impact various factors have on pay. (*Id.*) In other words, the multiple regression analysis can isolate the “estimate of the difference in pay between women and men” after controlling for other differences, such as work experience, type of work performed, geographic location, age, and performance reviews. (*Id.* ¶¶ 34, 38.) The analysis also measures the likelihood that the difference occurred by chance, as measured by the t-statistic and the p-value. (*Id.* ¶ 41.) Larger absolute values of the t-statistic indicate that the estimated pay difference is less likely to have occurred by chance; in the same vein, lower p-values indicate a lower probability that the observed difference arose by chance.³ (*Id.*)

Dr. Farber's multiple regression analysis revealed that female technical employees earn 8.6% less than male technical employees, with a statistically significant t-statistic of –25.42. (*Id.* ¶ 52.) After controlling for work experience, age, compensation year, and geographic location, the gender pay gap is reduced slightly to 7.4%, with a t-statistic of –25.62. (*Id.* ¶ 53.) Additionally controlling for performance review and Discipline—“job families within a Profession ... that produce similar business results”—narrows the gap to 6.3%. (*Id.* ¶ 54–55.) And finally, controlling for each worker's Standard Title, or their job title, shrinks the gender pay gap to 2.8% but remains statistically significant with a t-statistic of –21.73. (*Id.* ¶ 56.) However, Dr. Farber cautions that including a worker's Standard Title may understate the true gender pay gap because women are “systematically under-leveled relative to men.” (*Id.*)

*3 In fact, Dr. Farber characterizes Standard Title, as well as Career Stage and Stock Level, as examples of “tainted variables”—variables that appear to reduce the pay gap but only because the factors themselves are correlated with gender and pay through potentially discriminatory employer decisions. (*Id.* ¶ 46, 58.) To demonstrate this, Dr. Farber performs an ordered probit analysis, which returned a statistically significant, negative coefficient to the female indicator variable. (*See id.* ¶¶ 44, 60–61, 64.) This negative coefficient indicates that women are systematically assigned to lower Stock Levels and Career Stages, and thus, women are overrepresented

in the lower levels but underrepresented in the higher levels. (*Id.* ¶¶ 58–64.)

Lastly, Dr. Farber utilizes a probit analysis to assess the differences in the probability of advancement between men and women. (*See id.* ¶¶ 65–77.) A probit analysis is used when the outcome can take on one of two discrete values—for instance, being promoted or not being promoted. (*Id.* ¶ 42.) A probit analysis, like a multiple regression analysis, reveals the influence a factor has on the outcome and thus can estimate the difference between women and men in the probability of promotion. (*See id.* ¶ 43.) Dr. Farber concludes that women are 2.1 percentage points less likely to advance from one Stock Level to the next and 2.6 percentage points less likely to advance from one Career Stage to the next. (*Id.* ¶ 71.) Both differences are statistically significant. (*Id.*) Dr. Farber also calculates the differential in absolute terms by comparing the expected number of Stock Level advancements for women to the actual number of advancements. (*Id.* ¶ 72.) He finds that in total, female technical employees in Stock Levels 60–64 were denied 518 advancements that should have been expected. (*Id.* ¶ 77, 82.) Put another way, Dr. Farber found a shortfall of advancements among women in Stock Levels 60 to 64, all of which were statistically significant. (*Id.*)

In his deposition, Dr. Farber made clear that he did not analyze the relationship between pay and promotion and any particular supervisors or managers. (Farber Dep. (Dkt. # 363–1) at 241:11–19.) Thus, he has no opinion on whether “women working for different supervisors have higher or lower pay” or were “more or less likely to be promoted.” (*Id.* at 242:4–15.) Nonetheless, in his rebuttal report, Dr. Farber calculated the pay differentials with regard to each Level 1 supervisor at Microsoft⁴ and found that “virtually all woman years (more than 99%) are worked under Level 1 supervisors under whom women earn less than men, on average, after controlling for the factors in the model.” (Farber Rebuttal ¶ 16.) Dr. Farber did not further disaggregate the data. (*See* Farber Rep., Farber Rebuttal.)

B. Dr. Saad

Dr. Saad is the managing partner of Resolution Economics Group LLC, a firm that performs economic and statistical analyses in connection with litigation and consulting matters. (Saad Rep. (Dkt. ## 354–1 (sealed),

355–1 (redacted)) ¶ 2.) He holds a Ph.D. in economics from the University of Chicago and a B.A. in History and Economics from the University of Pennsylvania. (*Id.*) Prior to his consulting career, Dr. Saad was on the faculty of the economics and finance department at Baruch College of the City University of New York, where he taught labor economics, micro and macroeconomics, econometrics, and economic history. (*Id.*) Dr. Saad's work focuses on statistical analyses of systemic gender discrimination claims. (*Id.*)

*4 On February 16, 2018, Microsoft filed a second corrected version of Dr. Saad's report (the “revised report”) pursuant to [Federal Rule of Civil Procedure 26\(e\)](#). (*See* Not. of Saad Revised Rep. (Dkt. # 355).) Microsoft asserts that the revised report rectifies several mathematical errors but makes no substantive changes. (*Id.* at 1.) In this revised report, as before, Dr. Saad criticizes Dr. Farber for “aggregating all women together to gauge statistical significance overall” rather than taking into account variation among supervisors. (*Id.* ¶¶ 29, 30.) Dr. Saad analyzed how supervisors impact the inquiry by determining the number of supervisors under whom women earn what they are expected to earn under to Dr. Farber's model. (*Id.* ¶ 35.) In other words, Dr. Saad looked at what women should have been paid under Dr. Farber's model and analyzed how many supervisors oversaw women who were over or underpaid. (*Id.*) Dr. Saad concludes that “supervisors under whom more women earn significantly less as opposed to earning significantly more ... are in the minority.” (*Id.* ¶ 36.)

Regarding promotions, Dr. Saad compared Dr. Farber's predicted number of promotions to the actual number of promotions for the Plaintiffs and their declarants. (*See id.* ¶¶ 47–49.) In doing so, Dr. Saad assumes that “a probability above 50% indicates a promotion is more likely than not to occur.” (*Id.* ¶ 49.) Dr. Saad concludes that Plaintiffs and their declarants received more promotions than expected under Dr. Farber's model: “Only once does [Dr. Farber's] model predict that a promotion should have occurred when it did not.” (*Id.*)

As a part of his promotions analysis, Dr. Saad draws a distinction between promotions that occurred at the annual end-of-the-year review (“annual review”) and those that occurred at other times in the year, such as at a mid-year review (“mid-year/other”). (*Id.* ¶¶ 102–07.) Dr. Saad finds that annual review promotions showed no

gender difference, whereas the mid-year/other promotions show a statistically significant shortfall of promotions. (*Id.* ¶ 105.) To explain this difference, Dr. Saad opines that mid-year/other promotions are “different in character” than annual review promotions because “one sees business reasons cited more frequently” as justifications for mid-year/other promotions. (*Id.* ¶ 107.)

To confirm this hypothesis, Dr. Saad read a sample of 1,000 promotion justifications⁵ and identified 229 “different phrases that would indicate a promotion made for business reasons,” such as “there is a need,” “need:,” and “key to keeping Bing Ads Private Lab infrastructure [sic] running.” (*Id.* ¶¶ 108–09; *see also* Klein Decl. (Dkt. # 365) ¶ 6, Ex. 2 (“Business Need Index”).) He then searched for those 229 terms in all of the promotion justifications and concludes that business reasons were provided for 10.6% of the mid-year/other promotions as compared to only 7.2% of the annual review promotions. (Saad Rep. ¶ 109.) Dr. Saad did not speak with anybody at Microsoft regarding what qualifies as a business need. (Klein Decl. ¶ 5, Ex. 1 (“Saad Dep.”) at 22:23–23:10.)

Also as a part of his promotions analysis, Dr. Saad utilized a Z-model rather than a probit model. (*Id.* ¶ 116.) The Z-model analysis “construct[s] presumed homogeneous pools with respect to the variables available for study, where employees within these pools are considered similar, except for their gender.” (*Id.*) For each pool, or “strata,” the proportion of female promotions should be roughly equal to the proportion of women in the group. (*Id.*) Thus, the Z-model allows examination of the underlying strata while also allowing for aggregation across pools to obtain an overall result. (*Id.*) For women working in the Engineering Profession, Dr. Saad's aggregated Z-model analysis shows a 2.21% shortfall in promotions that is statistically significant. (*Id.* at 77.)⁶ For women in the I/T Operations Profession, Dr. Saad finds a 0.65% surplus in promotions that is not statistically significant. (*Id.* at 78.)

*5 In conducting his Z-model analysis, Dr. Saad created close to 59,000 selection pools for women in Engineering and 5,000 selection pools for those in I/T Operations. (Farber Rebuttal ¶ 47.) Nearly 60% of these pools in each profession had no female employees, and an additional 8% of the pools in Engineering and 12% of the pools in I/T Operations had only female employees. (*Id.*) Thus, almost 70% of the pools in each profession comprised of only a

single gender. (*Id.*) Thirty-nine percent of the Engineering pools and 43% of the I/T Operations pools contained only one individual. (*Id.*) Dr. Saad acknowledges in his deposition that pools with only men or only women “do not add to the analysis.” (Saad Dep. at 204:16–205:3.)

C. Ms. Young

Ms. Young is a Human Resources (“HR”) consultant with over 35 years of HR management experience, dealing directly with employee and employer issues. (Young Rep. (Dkt. ## 395 (sealed), 394 (redacted)) ¶ 7.) Her consulting work focuses on designing and implementing policies and procedures to minimize possible discrimination, harassment, and retaliation. (*Id.* ¶ 8.) Ms. Young has conducted over 300 HR audits, for a wide variety of organizations. (*Id.* ¶ 9.) She holds a master's degree in management from Pepperdine University and an undergraduate degree from University of California, Los Angeles. (*Id.* ¶ 12.) She has also taught HR and operations management professionals on how to gather and analyze information, conduct investigations, and resolve employee complaints. (*Id.* ¶ 14.)

Ms. Young was asked to evaluate the HR complaint investigation process at Microsoft and how the ERIT investigation processes compare with the “suggested management practices” (“SMPs”). (*Id.* ¶¶ 1–3.) To do so, Ms. Young reviewed the deposition testimony of ERIT investigators Melinda De Lanoy and Judy Mims; interviewed Ms. De Lanoy and Kimberly Meyers, the Manager of ERIT; looked over 18 out of 231 ERIT investigation case files; and analyzed the relevant Microsoft HR policies. (*Id.* ¶ 4; *see also id.*, Ex. 3; Shapiro Decl. (Dkt. ## 360–8 (sealed), 359–8 (redacted)) ¶ 5(h), Ex. 8 (“Young Dep.”) at 135:12–15.) Ms. Young did not interview any complainants who underwent an ERIT investigation. (*Id.* at 177:24–178:3.)

Ms. Young reviewed the case files to “assess [ERIT's] communication style and investigation process.” (Young Rep. ¶ 5.) Plaintiffs' counsel identified some of the case files during the deposition of Ms. De Lanoy. (Young Dep. at 136:6–8.) Defense counsel selected the remainder of the cases. (*Id.* at 136:24–137:18.) Although different ERIT investigators oversaw these cases and conducted the investigations in different years, Ms. Young acknowledged in her deposition that the sample was not meant to be representative in the following ways:

Q: So is it fair to say that your sample was not an attempt to be representative of the population in terms of the percentages of times that ERIT investigations complained about gender versus sexual harassment.

THE WITNESS: No.

Q: And it wasn't intended to be representative of the percent of complaints that come from any particular part of the company versus another part of the company?

THE WITNESS: That's correct.

Q: And it wasn't intended to be representative of the number of complaints that come from a certain year versus a different year?

THE WITNESS: That's correct.

Q: And it wasn't meant to be representative of the percentage of complaints that are investigated by one particular ERIT investigator versus another ERIT investigator?

THE WITNESS: I was not trying to single out one or two ERIT investigators to focus on in cases that they had done.

(*Id.* at 139:8–140:10 (objections omitted).)

Ms. Young concludes that “the investigation steps and sequence at Microsoft are consistent with [SMPs].” (Young Rep. ¶ 21.) She finds that “ERIT investigators are skilled, objective, and highly experienced.” (*Id.* ¶ 22.) She further finds that Microsoft has clear anti-discrimination and anti-harassment policies that are communicated effectively through internal communications, trainings, and websites. (*Id.* ¶¶ 23, 29–36.) These policies, Ms. Young remarks, are implemented in a “neutral, logical, and skillful manner.” (*Id.* ¶ 24.) Consistent with SMPs, Ms. Young concludes that ERIT investigations were thorough, timely, accurate, objective, well-documented, and based on the unique facts of each case. (*Id.* ¶¶ 37–55.)

*6 Ms. Young also observes that Microsoft’s “number of complaints ... is not unusual in a company of Microsoft’s size.” (*Id.* ¶ 24.) In fact, Ms. Young remarks that the “number of complaints may be the result of heightened employee awareness from increased training,

communication about ... policies, and/or increasing employee comfort and trust in using the complaint procedure.” (*Id.*)

III. ANALYSIS

Rule 702 of the Federal Rules of Evidence governs the admission of expert testimony in federal court:

A witness who is qualified as an expert by knowledge, skill, experience, training, or education may testify in the form of an opinion or otherwise if:

- (a) the expert's scientific, technical, or other specialized knowledge will help the trier of fact to understand the evidence or to determine a fact in issue;
- (b) the testimony is based on sufficient facts or data;
- (c) the testimony is the product of reliable principles and methods; and
- (d) the expert has reliably applied the principles and methods to the facts of the case.

Fed. R. Evid. 702. Rule 702 requires that the expert be qualified and that the “[e]xpert testimony ... be both relevant and reliable.” *Estate of Barabin v. AstenJohnson, Inc.*, 740 F.3d 457, 463 (9th Cir. 2014) (en banc) (quoting *United States v. Vallejo*, 237 F.3d 1008, 1019 (9th Cir. 2001)); Fed. R. Evid. 702. Relevancy “simply requires that [t]he evidence ... logically advance a material aspect of the party's case.” *Estate of Barabin*, 740 F.3d at 463 (quoting *Cooper v. Brown*, 510 F.3d 870, 942 (9th Cir. 2007)).

Reliability requires the court to assess “whether an expert's testimony has a ‘reliable basis in the knowledge and experience of the relevant discipline.’” *Id.* (quoting *Kumho Tire Co. v. Carmichael*, 526 U.S. 137, 149, 119 S.Ct. 1167, 143 L.Ed.2d 238 (1999)) (internal citations and alterations omitted). The Supreme Court has suggested several factors that courts can use in determining reliability: (1) whether a theory or technique can be tested; (2) whether it has been subjected to peer review and publication; (3) the known or potential error rate of the theory or technique; and (4) whether the theory or technique enjoys general acceptance within the relevant scientific community. See *Daubert v. Merrell Dow Pharm., Inc.*, 509 U.S. 579, 592–94, 113 S.Ct. 2786, 125 L.Ed.2d 469 (1993). The reliability inquiry is flexible, however,

and trial judges have broad latitude to focus on the considerations relevant to a particular case. *Kumho Tire*, 526 U.S. at 150, 119 S.Ct. 1167.

In determining reliability, the court must rule not on the correctness of the expert's conclusions but on the soundness of the methodology, *Estate of Barabin*, 740 F.3d at 463 (citing *Primiano v. Cook*, 598 F.3d 558, 564 (9th Cir. 2010)), and the analytical connection between the data, the methodology, and the expert's conclusions, *Gen. Elec. Co. v. Joiner*, 522 U.S. 136, 146, 118 S.Ct. 512, 139 L.Ed.2d 508 (1997); see also *Cooper*, 510 F.3d at 942 (“Rule 702 demands that expert testimony relate to scientific, technical or other specialized knowledge, which does not include unsubstantiated speculation and subjective beliefs.”); *Fed. R. Evid. 702* Advisory Committee's Notes to 2000 Amendments (“[T]he testimony must be the product of reliable principles and methods that are reliably applied to the facts of the case.”). Moreover, “the proponent of the expert ... has the burden of proving admissibility.” *Cooper*, 510 F.3d at 942 (citing *Lust v. Merrell Dow Pharms., Inc.*, 89 F.3d 594, 598 (9th Cir. 1996)).

*7 The exact application of *Daubert* at the class certification stage remains unclear. See *Fosmire v. Progressive Max Ins. Co.*, 277 F.R.D. 625, 628–29 (W.D. Wash. 2011) (“The Ninth Circuit, however, has not yet resolved whether a full analysis under *Federal Rule of Evidence 702* and *Daubert* is required at the class certification stage.”); see also *Wal-Mart Stores, Inc. v. Dukes*, 564 U.S. 338, 354, 131 S.Ct. 2541, 180 L.Ed.2d 374 (2011) (expressing in dictum that it is doubtful *Daubert* does not apply at the class certification stage). However, courts have struck a balance by applying *Daubert* at the class certification stage, but doing so in a manner that “recognizes the specific criteria under consideration, as well as the differing stage of discovery and state of the evidence.” *Fosmire*, 277 F.R.D. at 629; see also *Hovenkotter v. SAFECO Ins. Co. of Ill.*, No. C09-0218JLR, 2010 WL 3984828, at *4 (W.D. Wash. Oct. 11, 2010) (“[T]he court's consideration of the [experts'] opinions requires it to determine whether their opinions tend to show commonality of claims and damages among the class members; the court need not conduct a full *Daubert* analysis as to the admissibility for trial of the expert's opinions.”). Thus, on a motion for class certification, “it is not necessary that expert testimony resolve factual disputes going to the merits of plaintiff's

claims”; the testimony simply must be relevant in assessing whether there was a common pattern or practice that could affect the class as a whole. *Cholakyan v. Mercedes-Benz, USA, LLC*, 281 F.R.D. 534, 543 (C.D. Cal. 2012) (citing *Ellis v. Costco Wholesale Corp.*, 657 F.3d 970, 983 (9th Cir. 2011)).

The court addresses Microsoft's motion to exclude Dr. Farber's opinions before turning to Plaintiffs' motions to exclude the opinions of Dr. Saad and Ms. Young.

A. Dr. Farber

Microsoft makes two main arguments to exclude Dr. Farber's opinions. First, Microsoft contends that Dr. Farber's aggregated analysis is irrelevant to showing commonality because (1) aggregate disparities cannot show a common pattern of disparities across the relevant decision-making units; and (2) Dr. Farber does not analyze what role, if any, the challenged Calibration Process plays in the disparities. (Farber Mot. at 4–9.) Second, Microsoft argues that Dr. Farber's methodology is unreliable because he failed to take into account several “key determinations of [the] employment decisions at issue.” (*Id.* at 10; see *id.* at 10–12.) The court disagrees and takes each argument in turn.

1. Relevance

Microsoft argues first that Dr. Farber's aggregation of the data renders his opinions irrelevant because “his analysis is aggregated beyond the level where pay and promotion decisions are made.” (Farber Mot. at 9.) However, “it is a generally accepted principle that aggregated statistical data may be used where it is more probative than subdivided data.” *Paige v. California*, 291 F.3d 1141, 1148 (9th Cir. 2002). Aggregation is “particularly appropriate where small sample size may distort the statistical analysis and may render any findings not statistically probative.” *Id.* Thus, a plaintiff “should not be required to disaggregate the data into subgroups which are smaller than the groups which may be presumed to have been similarly situated and affected by common policies.” *Id.* (internal quotation marks omitted) (quoting *Eldredge v. Carpenters 46 N. Cal. Counties Joint Apprenticeship and Training Comm'n*, 833 F.2d 1334, 1340 n.8 (9th Cir. 1987)). Accordingly, whether aggregation is appropriate necessarily depends on “the structure of the entity being studied in light of the questions sought to be answered.”

Chen–Oster v. Goldman, Sachs & Co., 114 F.Supp.3d 110, 120 (S.D.N.Y. 2015).

In *Chen–Oster v. Goldman, Sachs & Co.*, the company, like Microsoft here, challenged the aggregation of data across a myriad of business units when the decisions regarding compensation and promotion were made at the business unit level. *Id.* at 120. The court rejected that argument. *Id.* at 120–21. In that case, the putative class challenged the company's use of “360 review” and “quartiling” processes, which were utilized in every division and every business unit. *Id.* at 120. Thus, “even if the effects of those policies may vary in different business units,” the court determined it was “appropriate to examine these policies across the entire population.” *Id.* at 120. Indeed, the court observed, disaggregating data to the business unit level produced such small sample sizes that it “tend[ed] to mask common mechanisms.” *Id.* at 120–21. Accordingly, the court allowed the aggregated statistical analysis “in light of the evidence that [the company] applies common performance measures” and “in light of the statistical pitfalls of disaggregation.” *Id.* at 121.

*8 Both of those considerations in *Chen–Oster* are also present here. Plaintiffs challenge Microsoft's use of the Calibration Process—a companywide system—that was used across levels and managers in determining pay and promotion. (SAC ¶¶ 26–52.) Moreover, disaggregating data as Microsoft suggests seems to produce pools with varying sizes, some of which may “mask common mechanisms” because of their small sample size. *See supra* § II.B (discussing the disaggregated pools created by Dr. Saad). Thus, as in *Chen–Oster*, even if Microsoft is correct that the effects of the Calibration Process vary across decision-maker, that variety does not render an aggregate statistical analysis wholly irrelevant. *See* 114 F.Supp.3d at 120–21; *see also Ellis v. Costco Wholesale Corp.*, 285 F.R.D. 492, 522 (N.D. Cal. 2012) (finding “good reason” to rely on aggregated statistics because it yields “more reliable and more meaningful statistical results” when the company's promotion practices are “uniform across the company”) (emphasis removed).

In construing “aggregated statistics as irrelevant to commonality” (Farber Mot. at 6), Microsoft relies on an overreading of *Dukes*. *Dukes* did not, as Microsoft implies (*see* Farber Mot. at 4–6), determine that the aggregation of data rendered the statistical analysis irrelevant; nor did *Dukes* preclude the use of aggregate statistics altogether

for future considerations of commonality. *See* 564 U.S. at 356, 131 S.Ct. 2541. Indeed, *Dukes* did not concern a motion to exclude at all. *See id.* at 342, 131 S.Ct. 2541. *Dukes* instead held that the aggregated statistical evidence presented by plaintiffs in that case was insufficient to satisfy the commonality requirement when the “only corporate policy ... is [one] of *allowing discretion* by local supervisors.” 564 U.S. at 355, 131 S.Ct. 2541. Thus, *Dukes* addressed commonality on the merits: when can an aggregate analysis establish “significant proof” that the company “operated under a general policy of discrimination” such that there is “a common answer to the crucial question *why was I disfavored*.” *Id.* at 353, 131 S.Ct. 2541. And, as discussed above, whether aggregation is appropriate is a case-specific determination, dependent on the company at issue and the claims before the court. *See Paige*, 291 F.3d at 1148. Thus, at best, *Dukes* commented on the relevancy of the aggregated statistical analysis that was before it. The court declines, as Microsoft urges, to read *Dukes* as holding that every expert opinion that aggregates data is irrelevant.⁷

Microsoft additionally faults Dr. Farber for not specifically analyzing the Calibration Process as a cause for any disparities. (Farber Mot. at 8–9.) It is certainly true that Dr. Farber did not draw any specific causal links between the Calibration Process and the disparities he found. (*See* Farber Dep. at 264:13–24.) But that does not render Dr. Farber's analysis of the alleged disparity irrelevant to the question of commonality. *See Chen–Oster*, 114 F.Supp.3d at 125 (“[I]t is not necessary for each expert to provide evidence establishing every element of a party's case.”). The disparity that Dr. Farber finds between women and men in pay and promotion “logically advance[s] a material aspect of [Plaintiffs'] case”—namely, that a gender-based disparity, in fact, exists that impacts all putative class members. *See Estate of Barabin*, 740 F.3d at 463.

*9 In summary, Dr. Farber's opinions are relevant, despite the fact that Dr. Farber's statistical analysis aggregates data across the entire population and does not reach any conclusion regarding the Calibration Process. Accordingly, the court declines to exclude Dr. Farber's opinions on relevance grounds.

2. Reliability

Microsoft additionally argues for exclusion on reliability grounds. (Farber Mot. at 10–12.) Microsoft claims that Dr. Farber “did not account for obvious, non-discriminatory indicators designating difference in type of work” and thus is “so incomplete as to be inadmissible.” (*Id.* at 10 (quoting *Bazemore v. Friday*, 478 U.S. 385, 400 n.10, 106 S.Ct. 3000, 92 L.Ed.2d 315 (1986).) Specifically, Microsoft faults Dr. Farber for omitting Stock Level in the pay analysis and Standard Title in the promotion analysis. (Farber Mot. at 11–12.)

Both parties correctly rely on *Bazemore v. Friday* as setting out the relevant standard for evaluating the inclusion of variables in regression analyses. (See Farber Mot. at 10; Farber Resp. at 8–9.) *Bazemore* explains:

While the omission of variables from a regression analysis may render the analysis less probative than it otherwise might be, it can hardly be said, absent some other infirmity, that an analysis which accounts for the major factors “must be considered unacceptable as evidence of discrimination.” Normally, failure to include variables will affect the analysis' probativeness, not its admissibility.

478 U.S. at 400, 106 S.Ct. 3000 (quoting *Bazemore v. Friday*, 751 F.2d 662, 672 (4th Cir. 1984)). Even a regression analysis “that includes less than ‘all measurable variables’ may serve to prove a plaintiff's case.” *Id.* *Bazemore* recognized, however, that there may be “some regressions so incomplete as to be inadmissible as irrelevant.” *Id.* at 400 n.10, 106 S.Ct. 3000.

Dr. Farber's regression analysis is far from one that is “so incomplete as to be inadmissible.” See *id.* As a preliminary matter, Dr. Farber took into account several different factors that might affect pay and promotion. He considered a worker's tenure, age, location, compensation year, performance review outcomes, job category, and Standard Title, also known as a job title, in his pay analysis. (See Farber Rep. ¶¶ 53–56.) For his promotion analysis, he factored in Discipline, Stock Level, age, experience, location, and performance review outcomes. (*Id.* ¶¶ 70–72.) On this basis alone, Dr. Farber's report is distinguishable from the cases Microsoft relies upon, in which the reports relied on factors that had “nothing to do with actual job performance or job requirements.” See *Anderson v. Westinghouse Savannah River Co.*, 406 F.3d 248, 263 (4th Cir. 2005); see also *Raskin v. Wyatt Co.*, 125 F.3d 55, 68 (2d Cir. 1997) (excluding a statistical

report that made “no attempt to account for other possible causes”).

Microsoft is correct that Dr. Farber did not include Stock Level in his pay analysis, or Standard Title in his promotion analysis. (See Farber Rep.) But Dr. Farber provides reasoned explanations for excluding these two factors. First, Dr. Farber explains that because Stock Level is a pay band, regressing compensation on Stock Level would “simply be regressing pay on a proxy for pay, which is inappropriate.” (*Id.* ¶ 47; see also Farber Rebuttal ¶ 59 (describing controlling for Stock Level as asking the “meaningless” question, “after controlling for how much each worker earns, do women earn less than men?”).)

*10 Second, and most importantly, Dr. Farber identifies both Stock Level and Standard Title as “tainted” variables, that is, both of these variables are themselves affected by gender bias.⁸ (*Id.* ¶¶ 48, 58–64.) “[I]llegitimate reasons—reasons themselves representative of the unlawful discrimination at issue—should be excluded from the regression (or otherwise dealt with) to avoid underestimating the significance of a disparity.” *Morgan v. United Parcel Serv. of Am., Inc.*, 380 F.3d 459, 469–70 (8th Cir. 2004); see D. James Greiner, *Causal Inference in Civil Rights Litig.*, 122 Harv. L. Rev. 533, 546–49 (2008). Dr. Farber shows how there is gender disparity in both Standard Title and Stock Level by illustrating how women are overrepresented in the lower stages and underrepresented in the higher stages.⁹ (See Farber Rep. ¶¶ 60–62, 64.) Microsoft may disagree with Dr. Farber's conclusions, but the correctness of the expert's conclusions does not bear on the *Daubert* determination. See *Estate of Barabin*, 740 F.3d at 463. Because these variables were themselves subject to bias and would therefore mask any discrimination, there were legitimate bases for Dr. Farber to have omitted both Stock Level and Standard Level.

In sum, the court finds that Dr. Farber's expert opinions are both relevant and reliable. Accordingly, the court denies Microsoft's motion to exclude Dr. Farber's expert reports and conclusions.

B. Dr. Saad

As a threshold matter, the parties disagree on which version of Dr. Saad's report should be considered. On February 16, 2018, Microsoft submitted, for the second time, a corrected version of Dr. Saad's report, which

Microsoft asserts rectifies several mathematical errors but makes no substantive changes. (*See* Not. of Saad Revised Rep.) Plaintiffs disagree. They filed a surreply asking the court to strike this newest version as untimely (Surreply (Dkt. # 357)) and also argued in their motion to exclude that the new report cannot be considered (*see* Saad Mot. at 5 n.1, 10–11).

As to the motion itself, Plaintiffs do not challenge Dr. Saad's conclusions as a whole. (Saad Mot. at 1.) Instead, Plaintiffs seek to exclude four specific portions of Dr. Saad's opinions because Plaintiffs assert they fail to meet the standards of reliability under [Rule 702](#) and [Daubert](#): (1) Dr. Saad's identification and subsequent search for “business reason” terms; (2) his use of 50% as a probability threshold for determining whether a predicted promotion would occur; (3) the number of pools Dr. Saad constructed and their usefulness in a Z-model analysis when many of those pools contained employees of only one gender; and (4) the mathematical errors contained within his predicated pay analysis. (*Id.*) The court determines which of Dr. Saad's reports to consider first before addressing each of Plaintiffs' challenges.

1. Revised Report

Plaintiffs assert that Dr. Saad's February 16, 2018, revised report “constitute[s] an impermissible supplemental expert opinion” and is thus untimely pursuant to the court's scheduling order. (Surreply at 1; *see* Sched. Order (Dkt. # 226) at 1.) Microsoft maintains that Dr. Saad's revised report “simply corrected the programming error and re-ran the same analyses” and thus is permitted—and even required—by [Federal Rule of Civil Procedure 26\(e\)](#). (Saad Resp. at 11.)

A party must submit its expert witness disclosures “at the times and in the sequence that the court orders.” [Fed. R. Civ. P. 26\(a\)\(2\)\(C\)](#). However, “if the party learns that in some material respect the disclosure or response is incomplete or incorrect, and if the additional or corrective information has not otherwise been made known to the other parties during the discovery process or in writing,” the party “must supplement or correct its disclosure or response.” [Id.](#) 26(e)(1). [Rule 26\(e\)](#)'s duty to supplement is not “a loophole through which a party who ... wishes to revise her disclosures in light of her opponent's challenges to the analysis and conclusions therein, can add to them to her advantage after the court's deadline.” [Luke v. Family Care and Urgent Med. Clinics](#), 323 Fed.Appx. 496, 500

(9th Cir. 2009). Instead, [Rule 26\(e\)](#) should only apply when the party “correct[s] an inaccuracy” or “fill[s] in a gap based on information previously unavailable.” *Id.*

*11 Plaintiffs aptly do not contest what [Rule 26\(e\)](#) allows.¹⁰ Instead, they contend that Dr. Saad's revised report falls outside the ambits of [Rule 26\(e\)](#) because it “offer[s] modified original opinions as well as unauthorized rebuttal” of Dr. Farber's report. (Surreply at 1.) But Plaintiffs offer no specific citation as to where these “new opinions” are in the revised report. (*See* Surreply at 1 (citing only to Dr. Saad's declaration); Saad Mot. at 11 (citing to Plaintiffs' surreply); Saad Reply (Dkt. # 413) at 6 (same).) In its review of the red-lined version filed by Microsoft, the court finds that the changes are corrections of mathematical errors that, as Dr. Saad represents, “ha[ve] only a minor impact on the results and do[] not change the substantive conclusions.” (*See* Saad Revisions (Dkt. # 354–2) at 2.) These changes are merely “updated figures using the same methodology” that are not “significantly different from the original reports.” *See Holiday Resales, Inc. v. Hartford Cas. Ins. Co.*, No. 07–1321JLR, 2008 WL 11343449, at *1–2 (W.D. Wash. Oct. 2, 2008). Accordingly, the court concludes that the revised report is properly before the court pursuant to [Rule 26\(e\)](#).¹¹

2. “Business Reason” Analysis

Plaintiffs seek to exclude the portions of Dr. Saad's report in which he details how he assessed that mid-year/other promotions “were more likely to be related to business reasons than are promotions that occur during the annual review.” (Saad Mot. at 7–9.) Plaintiffs criticize Dr. Saad's word-search analysis—in which Dr. Saad personally identified words he believed to connote a “business need” and subsequently searched for those words in all of Microsoft's promotion justifications—as unscientific and unreliable. (*Id.* at 7.) The court agrees.

First, Microsoft has not proved that Dr. Saad possesses a “reliable basis in the knowledge and experience of” identifying what business needs are in the context of Microsoft. *See Estate of Barabin*, 740 F.3d at 463. Dr. Saad admits that, out of the hundreds of cases in which he served as an expert, this is the first time he has conducted a “word search business need study” like the one conducted here. (Saad Dep. at 25:6–13.) Moreover, Dr. Saad did not discuss the issue with the Microsoft HR department or

anyone responsible for making promotion decisions at the company; indeed, Dr. Saad did not speak with anyone at Microsoft at all regarding what the company regards as a business need. (*Id.* at 22:23–23:10.)

When asked about how he identified the “business need” terms, Dr. Saad offers nothing more than the conclusory sentence that the identified phrases “all occurred in context where it was clear a promotion justification included a business reason.” (*See id.* at 22:2–4; *see also id.* at 24:8–11 (“The process was to read many, many promotion justifications and to identify those that clearly had a business justification.”).) Thus, Dr. Saad's methodology seemingly boils down to reading the justifications and pulling out phrases that struck him or his staff as “clearly” having a business justification. (*See id.*)

*12 Given the vague methodology offered, it is no surprise that Microsoft offers no evidence that this methodology has been subjected to peer review or publication, or that the technique enjoys general acceptance within the relevant scientific community. *See Daubert*, 509 U.S. at 592–94, 113 S.Ct. 2786. Nor does Microsoft provide this technique's known or potential error rate; in fact, Dr. Saad disclosed that whether his identification was over or under-exclusive “didn't concern [him].” (Saad Dep. at 51:23–25.) Nonetheless, Dr. Saad maintains “with certainty” that his conclusions “would not be changed by any other sort of search [Plaintiffs] would do.” (*Id.* at 52:4–8.) Dr. Saad supports this belief not with an indication of accuracy, but instead with his “knowledge of how statistical processes and distributional properties work.” (*Id.* at 52:13–17.)

The lack of reliability in Dr. Saad's methodology is highlighted in his final list of 229 identified terms. (*See Business Need Index.*) Some terms are so specific that they are unlikely to appear in any other comments. (*See id.* at 4–5 (identifying phrases such as “azure notification hub is expanding,” “the need to have a strong senior ic band grows with it,” and “key to keeping Bing Ads Private Lab infrastructure [sic] running”); *see also* Farber Rebuttal ¶ 40 (stating that nearly half of Dr. Saad's terms returns only one promotion justification—presumably the original justification from which the term was pulled—as relating to business need). Other terms are extremely general. (*See Business Need Index* at 1 (identifying phrases such as “we have asked,” “I could use,” and “new

branch”).) The court focuses on these business-need terms not to rule on the correctness of Dr. Saad's identifications, but instead to illustrate how his identifications resemble the “unsubstantiated speculation and subjective beliefs” that do not pass muster under [Rule 702](#) and *Daubert*. *See Cooper*, 510 F.3d at 942.

Microsoft clarifies that Dr. Saad “did not offer absolute opinions about which promotions definitively were made for business need.” (Saad Resp. at 7.) But “a witness must either have first-hand knowledge of the matter about which he testifies ... or he must utilize expertise in order to aid the finder of fact in understanding esoteric or complex evidence.” *Chen–Oster*, 114 F.Supp.3d at 124. Dr. Saad's “business needs” analysis falls into neither category. As discussed above, he does not have the expertise to determine what qualifies as a “business need,” and it is the employees or managers who have first-hand knowledge of the reason behind the promotions. Thus, it is inappropriate to present the “business needs” evidence with the imprimatur of an expert witness.

Microsoft also argues that Plaintiffs' criticisms regarding Dr. Saad's judgment “at most go to weight and do not warrant exclusion.” (*Id.* at 8.) In doing so, Microsoft mischaracterizes Plaintiffs' objections. Plaintiffs are not simply challenging the accuracy of the final list of business-need terms at which Dr. Saad arrived; they question the methodology—or lack thereof—that Dr. Saad relied upon to arrive at that list. Such a criticism, when founded, warrants exclusion. *See Estate of Barabin*, 740 F.3d at 463.

In sum, Dr. Saad's analysis regarding the business-need distinction between mid-year/other promotions and annual review promotions was not the product of reliable principles and methods. *See Fed. R. Evid. 702*. Accordingly, the court grants Plaintiffs' motion to exclude the portions of Dr. Saad's opinion that rely on his business-need analysis. Thus, the court excludes paragraphs 106 to 110, as well as the graph on page 72, that speak of how promotions may differ based upon business need.

3. 50% Probability Threshold

Plaintiffs challenge Dr. Saad's use of an assumption that “if the predicted probability of a promotion ... is over 50% (meaning more likely to occur than not) then the promotion ‘should have occurred’ ” and conversely, that

“if the same figure was under 50%, the promotion should not have occurred.” (Saad Mot. at 4.) Microsoft responds that Dr. Saad's use of a 50% threshold is an acceptable statistical method, and thus, any dispute goes to weight, not admissibility. (Saad Resp. at 8–9.) The court agrees with Microsoft.

*13 The court's *Daubert* duty is to judge the reasoning used in forming an expert conclusion and whether that reasoning is scientific. *Kennedy v. Collagen Corp.*, 161 F.3d 1226, 1230–31 (9th Cir. 1998). The presence of opposing scientific tests or methods “should not preclude the admission of the expert's testimony—they go to the weight, not the admissibility.” *Id.*; accord *McCulloch v. H.B. Fuller Co.*, 61 F.3d 1038, 1044 (2d Cir. 1995) (stating that disputes as to “faults in [an expert's] use of [a particular] methodology” go to weight, not admissibility). Thus, so long as the methods employed by the expert are scientifically valid, disagreement with the assumptions behind the methods or the methodology employed does not warrant exclusion. *S.E.C. v. Das*, 723 F.3d 943, 950 (8th Cir. 2013).

Microsoft has provided evidence that utilization of a 50% threshold is an acceptable method for measuring the “goodness of fit” in a probit analysis. (See Saad Decl. (Dkt. # 374) ¶ 3, Ex. A (“Woodridge Textbook”) at 465 (utilizing a 0.5 threshold when computing predicted probability); *Id.* ¶ 4, Ex. B (“Maddala Textbook”) at 334 (utilizing a 0.5 threshold to “count the number of correct predictions”).) Plaintiffs are, of course, free to dispute whether utilization of that statistical method, or whether Dr. Saad's prediction model overall, is appropriate. (See Saad Reply at 3 (arguing that a 50% threshold is only applicable to evaluate an aggregated regression analysis, not to predict individual promotion outcomes).) But that dispute does not warrant the exclusion of Dr. Saad's opinion at the *Daubert* stage. See *Kennedy*, 161 F.3d at 1230–31. Accordingly, the court denies Plaintiffs' motion to exclude this portion of Dr. Saad's opinion.

4. Pools in Z-Model Analysis

The court reaches a similar conclusion regarding Plaintiffs' challenge of Dr. Saad's disaggregation of the data while performing his Z-model analysis. Both parties recognize that a Z-model analysis, or the selection pool method, is a well-recognized statistical modeling technique that is often utilized to analyze disparities, especially those present in employment litigation cases.

(Saad Resp. at 9–10; Saad Reply at 4.) Thus, the court finds that Dr. Saad utilized scientifically sound and widely accepted methodology while conducting his study.

Plaintiffs, instead, take issue with how Dr. Saad applied the selection pool method—specifically the number of pools into which Dr. Saad split the data and how many of those pools “provide no useful information.” (See Saad Mot. at 4–5.) Plaintiffs are correct that Dr. Saad split the data into numerous pools and that a proportion of those pools contain no gender diversity and thus, as Dr. Saad admits, “do not added to the analysis.” (See Saad Dep. 204:16–205:3.) But Microsoft emphasizes that a Z-model analysis must compare similarly situated employees, and thus, Dr. Saad's small pool sizes are merely “a function of the highly differentiated work that employees ... perform.” (Saad Resp. at 10.) Indeed, Dr. Saad observes that having some strata offer “no useful information” is “a very common outcome in any aggregated selection model [where] you are using a number of factors to create the strata.” (Saad Dep. at 206:15–21.)

Plaintiffs' challenges qualify as “objections to the inadequacies of [Dr. Saad's] study”—the exact kind of concerns that “go[] to the weight of the evidence rather than its admissibility.” See *Hemmings v. Tidyman's Inc.*, 285 F.3d 1174, 1188 (9th Cir. 2002). In *In re Phenylpropanolamine (PPA) Prod. Liab. Litig.*, 289 F.Supp.2d 1230 (W.D. Wash. 2003), the court considered a similar challenge regarding the small numbers upon which a study was based. *Id.* at 1241. Yet, despite those inadequacies, the court found “the methodology scientifically sound,” and thus, concluded that “any flaws that might exist go to the weight afforded the [study at issue], not its admissibility.” *Id.* at 1240–41. The court reaches the same conclusion here. Despite the “inadequacies” Plaintiffs indentify in Dr. Saad's analysis, Dr. Saad utilizes a scientifically sound method of analysis—one that requires him to disaggregate the data into pools containing individuals who are similar with respect to variables that influence promotions. Thus, any flaws that Plaintiffs identify may affect the weight afforded to Dr. Saad's study but do not warrant the study's exclusion. Accordingly, the court denies Plaintiffs' motion to exclude this portion of Dr. Saad's opinions.

5. Mathematical Errors

*14 Plaintiffs lastly seek to exclude Dr. Saad's predictions of individual employees' pay because Plaintiffs

allege that Dr. Saad made “numerous mathematical and statistical errors.” (Saad Mot. at 5.) Plaintiffs point out three errors: (1) assuming that all employees, regardless of Standard Title, receive the same average compensation; (2) incorrectly computing the standardized residual; and (3) omitting a mathematical component, called the “smearing factor,” when transforming a natural logarithm (“log”) of total compensation into dollar terms. (*Id.*) The first two of these errors, as Plaintiffs concede, were corrected in Dr. Saad's revised report.¹² (*See* Saad Mot. at 11; *see also* Saad Rep.) The court has determined that this revised report is properly before the court, *see supra* § III.B.1, and thus, the only error that remains in dispute is the omission of the “smearing factor.”

Plaintiffs' concern regarding the “smearing factor” does not render Dr. Saad's analysis unreliable. As Microsoft argues, there is evidence that the “smearing factor” is not required in all circumstances. (*See* Saad Decl. ¶ Ex. C (“Manning Article”) at 464 n.3 (noting other ways to “get a consistent estimate of the smearing factors”).) Thus, Dr. Saad's decision to not utilize the “smearing factor” in this scenario—a choice within the bounds of statistical analysis—is another purported “inadequacy” that Plaintiffs may argue affords the opinion less weight. *See Hemmings*, 285 F.3d at 1188. It does not, however, warrant exclusion of the entire opinion. Because Dr. Saad's revised report corrected the first two errors, and the third, even if erroneous, does not warrant exclusion, the court denies this portion of Plaintiffs' motion to exclude Dr. Saad's opinions.¹³

In summary, the court grants Plaintiffs' motion as it pertains to Dr. Saad's “business needs” analysis and excludes paragraphs 106 to 110 as well as the graph on page 72 of Dr. Saad's revised report. However, the court denies Plaintiffs' motion in all other respects.

C. Ms. Young

Plaintiffs seek to exclude Ms. Young's report for four reasons: (1) Ms. Young “lacks a sufficient understanding of the HR field” and thus is not qualified; (2) her report is not based on sufficient facts or data; (3) her report is not based on any scientific method and is thus unreliable; and (4) her report is inconsistent with the accepted principles and methods in the HR field. (Young Mot. at 1–2.) Ms. Young seems to be relying on her personal experience to evaluate Microsoft's ERIT practices (*see* Young Mot.

at 4–5; Young Resp. at 5), but the court “must ensure that expert testimony, whether it is based on ‘professional studies or personal experience,’ employs ... the same level of intellectual rigor that characterizes the practice of an expert in the relevant field.” *Fortune Dynamic, Inc. v. Victoria's Secret Stores Brand Mgmt., Inc.*, 618 F.3d 1025, 1035–36 (9th Cir. 2010) (quoting *Kumho*, 526 U.S. at 152, 119 S.Ct. 1167). Because the court finds that Ms. Young's study is not based on sufficient facts or data, the court does not reach the remainder of Plaintiffs' objections.

*15 “Relevant expert testimony is admissible only if an expert knows of facts which enable him to express a reasonably accurate conclusion.” *Smith v. Pac. Bell Tel. Co., Inc.*, 649 F.Supp.2d 1073, 1096 (E.D. Cal. 2009); *see also* Fed. R. Evid. 702 (requiring the expert opinion to be based on sufficient facts or data to be reliable). Opinions that are derived from erroneous or incomplete data are appropriately excluded. *Id.*; *see Arjangrad v. JPMorgan Chase Bank, N.A.*, No. 3:10-CV-01157-PK, 2012 WL 1890372, at *6 (D. Or. May 23, 2012). In *Powell v. Anheuser–Busch Inc.*, 2012 WL 12953439 (C.D. Cal. Sept. 24, 2012), the court considered an expert report that relied on evidence that did not “provide a complete picture of the relevant events.” *Id.* at *6. The court thus reasoned that the expert “offer[ed] opinions without a full understanding or knowledge of the facts of this case.” *Id.* Because the expert “failed to sufficiently consider the relevant underlying facts necessary to support his opinions and conclusions,” the court excluded the opinion as unreliable. *Id.* at *7.

Here, the court finds several issues with the facts and data relied upon by Ms. Young. First, Ms. Young reviewed only 18 out of 231 gender complaints—or 7.79% of the gender complaints reviewed by ERIT. Plaintiffs' rebuttal expert, Dr. Caren Goldberg, observes that any sample “that comprises less than 10% of the total population of cases raises questions about representativeness.” (Goldberg Rep. (Dkt. ## 346 (sealed), 359–12 (redacted) ¶ 15.) This is because “[s]maller samples are more likely to be different from the population than are larger ones, so smaller samples have more sampling error.” (*Id.*) Indeed, Ms. Young concedes that she reviewed only “a small portion” of the total number of complaints. (Young Dep. at 135:8–10.) Given the small sample size, the sampling error of the complaints is likely substantial. (*See* Goldberg Rep. ¶ 15.)

Second, the complaints reviewed by Ms. Young are not representative of all complaints. Representativeness is essential to the reliability of a study because to the extent that a sample “systematically differs from the population, inferences about the population from the sample are misleading.” (Goldberg Rep. ¶ 16.) In other words, if the 18 complaints reviewed by Ms. Young is not representative of all of the complaints processed by ERIT, Ms. Young's opinion regarding ERIT's efficacy would fail to meet the standards of Rule 702 and *Daubert*. Ms. Young did not randomly select the files to be considered. Instead, Ms. Young chose ten files that were attached to Plaintiffs' deposition of Ms. De Lanoy and asked Microsoft's counsel to “add[] a few extra cases.” (See Young Dep. at 136:6–10.) The fact that some of the files were selected by defense counsel could suggest that “the selection of which [files] were analyzed was subject to bias.” See *Chen–Oster*, 114 F.Supp.3d at 124. However, standing alone, selection of some files by defense counsel would likely not undermine the representativeness of the selected files.¹⁴ See *Arjangrad*, 2012 WL 1890372, at *6 (recognizing “courts generally permit” the “common practice for counsel to select a subset of documents to give to a potential expert”).

The bigger problem is that Ms. Young admits the sample was not representative—and indeed, was not intended to be representative—of the larger universe of complaints reviewed by ERIT. (See Young Dep. 139:8–140:10.) Ms. Young repeatedly conceded during her deposition that the sample she reviewed was not an attempt to be representative of the claims presented, the source of the complaints, or the years in which claims arose. (See *id.*) She did not give defense counsel any specific instructions on how to select cases that would be representative. (See *id.* at 137:11–18.) Nor did Ms. Young, after she received the sample, take any additional steps to check or ensure that it is representative of how ERIT handles all of its complaints. (See generally Young Rep.; Young Dep.) This lack of representativeness—combined with the sampling error resulting from the small sample size and potential selection bias stemming from choice of the files by defense counsel—suggests that the case files Ms. Young relied upon did not “provide a complete picture of the relevant events.” See *Powell*, 2012 WL 12953439, at *6. Thus, Ms. Young's methodology, based upon inferences drawn from the incomplete sample, is unreliable.¹⁵

*16 Microsoft offers no evidence that the sample reviewed by Ms. Young met any measure of scientific representativeness. (See Young Resp.) Instead, Microsoft baldly asserts—without any legal citation—that “Ms. Young's approach was biased in Plaintiffs' favor” because ten of the files were identified by Plaintiffs in their deposition of Ms. De Lanoy. (Young Resp. at 7; see also Young Dep. at 136:6–8.) But, as Plaintiffs aptly argue, Ms. Young's reliance on a deposition only “underscores the problem” rather than resolves the issue. (See Young Reply (Dkt. # 397) at 4.) Documents selected for use in a deposition are not designed to be representative of the overall universe of complaints. Tellingly, Microsoft offers no support for the notion that a sample is adequate simply because some of the files originated from Plaintiffs' deposition.

Nor was the other data relied upon by Ms. Young sufficient. Ms. Young chose only to interview arguably the two most experienced members of the ERIT investigations team—Ms. De Lanoy, the investigator with the most years of experience, and Ms. Meyers, the manager of ERIT. (See Young Dep. at 241:9–19, Ex. 6; Goldberg Rep. at ¶ 12.) Moreover, Ms. Young chose to interview Ms. De Lanoy after already reviewing Ms. De Lanoy's deposition, which “is not only superfluous, [but] results in an overweighting of that individual's perspective.” (Goldberg Rep. ¶ 12.) Ms. Young did not interview the least experienced member of the ERIT team. Indeed, when evaluating the experience of the ERIT team, Ms. Young omitted an inexperienced member from the list with no explanation. (See Young Rep. Ex. 6; Goldberg Rep. ¶ 12.) Additionally, Ms. Young's choice to interview ERIT investigators suffers from “a significant danger of reporting bias.” See *Chen–Oster*, 114 F.Supp.3d at 124. Ms. Young essentially sought information to evaluate a system from the very investigators whose work are being evaluated. Although the investigator's perspective may be a relevant piece of information, Ms. Young never sought to supplement that information by interviewing any complainant who initiated an ERIT investigation or any of the declarants in the present case.¹⁶ (See Young Rep. ¶ 4.) The court therefore finds that Ms. Young's other data are likewise unrepresentative and thus insufficient.

If only some of Ms. Young's underlying data was suspect, the court may well find that the problem did not undermine the study's overall reliability. However, here, the confluence of factors discussed above compels the

court to conclude that Ms. Young's opinion is not based on sufficient facts or data, as is required under [Rule 702](#). See [Fed. R. Evid. 702](#). Accordingly, the court grants Plaintiffs' motion to exclude Ms. Young's opinion in its entirety.

Even if Ms. Young relied on sufficient facts or data, the court is doubtful that she adequately specified the method by which she arrived at her conclusions. (See [Young Mot.](#) at 9–11.) Experts that rely primarily on experience must explain “how that experience leads to the conclusion reached, why that experience is a sufficient basis for the opinion, and how that experience is reliably applied to the facts.” [Fed. R. Civ. P. 702](#), Advisory Committee Notes, 2000 Amendments. As applicable here, an expert cannot “rel[y] on the mere fact of his experience with respect to human resources matters to support [his] conclusion.” [Parton v. United Parcel Serv.](#), No. 1:02-cv-2008-WSD, 2005 WL 5974445, at *5 (N.D. Ga. Aug. 2, 2005).

*17 That seems to be exactly what Ms. Young does here. For example, Ms. Young comments that, based upon her experience, the number of complaints at Microsoft is not unusual for a company of Microsoft's size. ([Young Rep.](#) ¶ 56.) But she proffers nothing more. She does not explain how her experience leads her to that conclusion or how her experience is applied specifically to the size of Microsoft or the number of complaints here. (See *id.*) Nor does she provide any additional detail, such as the size of other comparable companies or the volume of complaints at those companies. (See *id.*) In similar circumstances,

courts have excluded such expert opinions. See [Easton v. Asplundh Tree Experts, Co.](#), No. C16-1694RSM, 2017 WL 4005833, at *4–5 (W.D. Wash. Sept. 12, 2017) (excluding HR expert's opinion when he “never explains how his experience led him to” his conclusions).

In sum, the court grants Plaintiffs' motion to exclude Ms. Young's opinions for failing to meet the standard as set out in [Rule 702](#) and [Daubert](#).

IV. CONCLUSION

Based on the foregoing analysis, the court DENIES Microsoft's motion to exclude Dr. Farber's expert opinions (Dkt. # 362), GRANTS in part and DENIES in part Plaintiffs' motion to exclude Dr. Saad's expert opinions (Dkt. # 364), and GRANTS Plaintiffs' motion to exclude Ms. Young's expert opinions (Dkt. ## 367 (sealed), 368 (redacted)).

Additionally, the court DIRECTS the clerk to provisionally file this order under seal. The court ORDERS counsel to meet and confer regarding the need for redaction and to jointly file a statement on the docket within ten (10) days of the date of this order to indicate any such need.

All Citations

--- F.Supp.3d ----, 2018 WL 2124019

Footnotes

- 1 Microsoft requests oral argument (see generally [Farber Mot.](#); [Saad Resp.](#) (Dkt. # 376); [Young Resp.](#) (Dkt. # 377)), but the court finds that oral argument would not be helpful to its disposition of the motions, see [Local Rules W.D. Wash. LCR 7\(b\)\(4\)](#). Moreover, district courts are not required to hold a separate [Daubert](#) hearing before ruling on motions to exclude. See [Millenkamp v. Davisco Foods Int'l, Inc.](#), 562 F.3d 971, 979 (2009). The court finds that the briefing, expert reports, and deposition testimony of the various experts presented here comprise an adequate record from which the court can make its ruling. See *id.*
- 2 “Stock levels,” also known as “pay levels,” represent the level associated with an employee and delineate the various salary ranges that Microsoft employees receive. ([Farber Rep.](#) (Dkt. # 332) (sealed) ¶ 20.)
- 3 Generally, a t-statistic with an absolute value that is greater than 1.96 indicates that the observed difference would have occurred by chance less than five percent of the time. (*Id.* ¶ 41.) Similarly, a p-value of 0.0005 indicates that there is only a 0.05% chance that the observed difference arose randomly. (*Id.*)
- 4 Dr. Farber indicates that 98% of the putative class members report to one of four Level 1 supervisors. ([Farber Rebuttal](#) (Dkt. # 344 (sealed)) ¶ 16.) Level 1 supervisors approve and oversee all promotion and pay decisions but do not usually make promotion decisions for lower-level employees. ([DeCaprio Decl.](#) (Dkt. # 295) ¶ 12; [Jarvis Decl.](#) (Dkt. # 306) ¶¶ 17–18.)

- 5 Microsoft managers provide comments that detail their reasons for recommending a certain employee for promotion. (Saad Rep. at 70 n.80.) These comments are referred to as “promotion justifications” or “promotion justification comments.” (See *id.*)
- 6 Dr. Saad does not number the tables that he embeds in the body of his report. Thus, when citing to these tables, the court cites to the page number of the report located in the bottom right-hand corner.
- 7 Tellingly, the majority of the case law cited by Microsoft involves whether the specific statistical evidence presented was sufficient to establish commonality, and not the question currently before the court—whether the expert opinion was relevant to the commonality question. See *Bolden v. Walsh Const. Co.*, 688 F.3d 893, 897 (7th Cir. 2012) (“We need not determine whether [the] study should have been excluded under Fed. R. Evid. 702.”); *Abram v. United Parcel Serv. of Am., Inc.*, 200 F.R.D. 424, 431 (W.D. Wis. 2001) (precluding finding of commonality after determining that the aggregate statistical evidence was insufficient); *Artis v. Yellen*, 307 F.R.D. 13, 26 (D.D.C. 2014) (citing multiple reasons why the statistical evidence presented fell short of proving commonality).
- 8 Dr. Farber shows how Career Stage is a tainted variable. (See Farber Rep. at ¶¶ 59–62.) Because Standard Level is dependent on Career Stage, Dr. Farber concludes that Standard Level is likewise a tainted variable. (*Id.* ¶ 58.)
- 9 Indeed, Microsoft’s expert witness, Dr. Saad, reaches the same conclusion: Dr. Saad found a gender-based disparity in Stock Level with a t-statistic of –17.55; a t-statistic of just 1.96 is considered statistically significant. (Saad Rep. ¶ 129; *id.* at 88.)
- 10 Indeed, Plaintiffs recently filed a motion seeking leave to file a supplemental rebuttal report from Dr. Farber. (See Mot. to Supp. Farber Rep. (Dkt. # 458).) In that motion, Plaintiffs make the same argument regarding Rule 26(e) that they chastise Microsoft for making here. (*Compare id.* at 2 (“The Rules not only permit a party to supplement expert opinions in these circumstances, but require it.”), with Saad Resp. at 11 (“Rule 26(e), however, not only permits but requires an expert witness to correct his report.”).)
- 11 Plaintiffs also assert that Rule 26(e) “explicitly states that [any] changes must be disclosed prior to court-ordered discovery deadlines.” (Saad Reply at 6.) Not so. Rule 26(e)(2) specifies that any changes must adhere to Rule 26(a)(3), which governs pretrial disclosures. Fed. R. Civ. P. 26(e)(2). Rule 26(a)(3), in turn, specifies that disclosures must be made “at least 30 days before trial.” *Id.* 26(a)(3)(B); see also *Holiday Resales*, 2008 WL 11343449, at *2 (“The party must supplement the expert report by the time the party’s pretrial disclosures are due under Federal Rule of Civil Procedure 26(a)(3)(B), *i.e.*, at least 30 days prior to trial.”). The pretrial disclosure deadline is plainly not the same as any court-ordered discovery deadline.
- 12 Even if the first two mathematical errors were not corrected, it is doubtful that these technical errors would have been adequate grounds to exclude an entire opinion. See *Todd v. Tempur-Sealy Int’l, Inc.*, No. 13-cv-04984-JST, 2016 WL 5462428, at *6 n.5 (N.D. Cal. Sept. 28, 2016) (refusing to exclude expert opinion because of a mathematical error).
- 13 Based on the alleged mathematical errors, Plaintiffs also challenge the reliability of three pie charts Dr. Saad utilized to illustrate the number of supervisors who over or under-pay female employees. (Saad Mot. at 11.) Because the court finds that the mathematical errors do not warrant exclusion, the court likewise denies Plaintiffs’ request to exclude the pie charts on this ground. Moreover, the court does not agree that the pie charts are “fundamentally misleading,” as Plaintiffs allege. (See *id.*) Although Plaintiffs are correct that the pie charts represent the number of supervisors, and not the number of female employees those supervisors oversee, that distinction does not render Dr. Saad’s methodology unreliable. Thus, the court denies Plaintiffs’ motion to exclude the three pie charts.
- 14 Contrary to Plaintiffs’ assertion, *Chen–Oster* does not stand for the proposition that counsel selection of data alone renders an expert analysis unreliable. (See Young Mot. at 7 n.7.) Instead, *Chen–Oster* found the expert opinion unreliable for “a number of reasons,” only one of which was related to who selected the sample to be analyzed. See 114 F.Supp.3d at 124.
- 15 Curiously, Ms. Young emphasizes that the case files were only “complements to [her] opinion” and that her aim was “not to rehash ... how well [the investigation] was conducted.” (Young Dep. 132:23–25; see also *id.* at 267:6–9 (stating that she did not review the case files to determine “whether or not [she] think[s] [the ERIT investigators] did a good or bad job.”).) But Ms. Young’s assignment was to “evaluate[] how Microsoft’s investigation processes compare with usual and customary HR,” including an evaluation of the “quality of the investigation process.” (Young Rep. ¶¶ 3.) Such an evaluation necessarily involves, at some level, a review of how well the investigation was conducted in each case file. As Dr. Goldberg opines, “[b]ecause the investigations are the unit of analysis, they necessarily need to be the primary source of data.” (Goldberg Rep. ¶ 11.) Ms. Young’s choice to relegate what should be the primary source of data to a “complementary” role further augments the unreliability of the study.
- 16 Microsoft suggests that because Dr. Goldberg “rarely” conducts interviews, Ms. Young is “engaging in a more thorough investigation that Plaintiffs’ expert would have undertaken.” (Young Resp. at 9.) This argument misses the mark. Dr.

Goldberg's comment regarding interviews was made in the context of explaining how she does not rely on interviews because for her, the case files serve as the primary source of data. (See Goldberg Rep. ¶ 11.) Additionally, Dr. Goldberg did not take issue with the fact that Ms. Young conducted interviews; rather, she objected to how Ms. Young's interviews were conducted. Specifically, Dr. Goldberg criticized Ms. Young for the lack of representativeness in her interviewees and the interviewee's inability to "shed light on issues that are not otherwise illuminated. (*Id.* ("[I]t is unlikely that limiting [Ms. Young's] scope to two investigators provided a balanced understanding of the fairness and adequacy of the [ERIT] process."))

End of Document

© 2018 Thomson Reuters. No claim to original U.S. Government Works.