BINGHAM MCCUTCHEN LLP
DONN P. PICKETT (SBN 72257)
GEOFFREY M. HOWARD (SBN 157468)
HOLLY A. HOUSE (SBN 136045)
ZACHARY J. ALINDER (SBN 215695)
BREE HANN (SBN 215695)
Three Embarcadero Center
San Francisco, CA 94111-4067
Telephone: 415.393.2000
Facsimile: 415.393.2286
donn.pickett@bingham.com
geoff.howard@bingham.com
holly.house@bingham.com
zachary.alinder@bingham.com
bree.hann@bingham.com

BOIES, SCHILLER & FLEXNER LLP
DAVID BOIES (Admitted *Pro Hac Vice*)
333 Main Street
Armonk, NY 10504
Telephone: 914.749.8200
dboies@bsfllp.com
STEVEN C. HOLTZMAN (SBN 144177)
1999 Harrison St., Suite 900
Oakland, CA 94612
Telephone: 510.874.1000
sholtzman@bsfllp.com

DORIAN DALEY (SBN 129049)
JENNIFER GLOSS (SBN 154227)
500 Oracle Parkway, M/S 5op7
Redwood City, CA 94070
Telephone: 650.506.4846
Facsimile: 650.506.7114
dorian.daley@oracle.com
jennifer.gloss@oracle.com

Attorneys for Plaintiffs
Oracle USA, Inc., *et al.*

UNITED STATES DISTRICT COURT
NORTHERN DISTRICT OF CALIFORNIA
OAKLAND DIVISION

ORACLE USA, INC., *et al.*,

Plaintiffs,

v.

SAP AG, *et al*,

Defendants.

No. 07-CV-01658 PJH (EDL)

DECLARATION OF DANIEL S. LEVY IN SUPPORT OF ORACLE'S MOTION NO. 6: TO EXCLUDE TESTIMONY OF DEFENDANTS' EXPERT BRUCE SPENCER

Date: September 30, 2010
Time: 9 a.m.
Place: Courtroom 3
Judge: Hon. Phyllis J. Hamilton

I, Daniel S. Levy, declare as follows:

1. I am the National Managing Director and a founder of Advanced Analytical Consulting Group, Inc. ("AACG"). I have personal knowledge of the facts stated within this Declaration and could testify competently to them if required.

2. I have been retained by counsel for the Plaintiffs in the matter of Oracle USA, Inc, *et al.* v. SAP AG, *et al.* (Case No. 07-CV-01658 PJH (EDL)) to provide an expert report, deposition testimony and this declaration in support of Oracle's motion to exclude Dr. Spencer's testimony. My billing rate for this case is $627 per hour. The rates of my staff assigned to this project range from $250 to $507. Compensation for AACG is not contingent on the outcome of the proceedings.

3. I have a Ph.D. in Economics from The University of Chicago. I have testified in a range of matters over a number of years, including regression analysis, statistical methods and damages analysis. I, and my company, are currently engaged in consulting projects for Fortune 500 companies in the US and internationally in which the main purpose of our work is the construction of advanced econometric models, regression analyses, statistical analyses, large-scale sample designs and data collections to help major corporations understand their revenues, costs, liabilities and risks. I have taught classes in statistical methods, including regression analysis, to corporate economists, accountants and statisticians. I have served as a computer advisor at The University of Chicago Computation Center, where I advised researchers on the implementation of statistical and econometric methods, including regression analysis.

Significance Levels and Their Use in Hypothesis Testing for Random Chance

4. Scientific hypothesis tests are often used to detect when events differ significantly from what would occur by random chance.

5. Scientists often declare events that would only occur 1 time out of 100 or 5 times out of 100 by random chance as significantly different from random. These .01 and .05 probabilities are often called statistical significance, or probability, levels by scientists (denoted α-levels).

6. The .01 and .05 significance levels are the most commonly used scientific

standards and are listed in numerous scholarly texts for determining whether some characteristic in the data was consistent with being a random sample. (*See* Federal Judicial Center, *Reference Manual on Scientific Evidence* ("FJC Manual"), 124 (2d ed, 2000)[1]; Agresti A. and B. Finlay, Statistical Methods for the Social Sciences, Upper Saddle River, New Jersey: Prentice Hall, 2009, p. 154 ("In practice, the most common α-levels are 0.05 and 0.01"); Pindyck R. S. and D.L. Rubinfeld, Econometric Models and Econometric Forecasts, Fourth Edition, Boston, MA: Irwin McGraw-Hill, 1998, p. 43 ("Most statistical analyses report tests of statistical significance by pointing out which coefficients are significant at the 1 percent, 5 percent, or other appropriate significance level.") Significance levels higher than .05 are generally not used because they are too likely to declare something non-random when it has a high chance of being caused by a random process. For example, in scientific testing, if researchers see a set of events and know that these events happen 10% (.1) of the time by random chance, they would conclude that this set of events is not very rare, by common statistical significance levels, because they would see such an outcome in 10 out of every 100 tests they ran simply by random chance.

7. In his report, Dr. Spencer says that he found that my critical support random sample possessed a characteristic that would occur in repeated samples approximately 927 times out of 10,000 by random chance, that is with a probability of .0927(9.27%) in random samples.[2] This means the event is not significantly different from a random sample based on the most commonly used scientific standard of .05. (*See* FJC Manual at 124). Based on the most common scientific statistical standards, Dr. Spencer's *statistical test* does not reject the hypothesis that my sampling process was random.

8. Whenever two samples are drawn (as here with retrofit and critical support), the probability of observing some given characteristic by random chance in one or the

---

[1] The confidence intervals I produce in my report are related to the .05 "one-tailed" test for the lower bound as discussed in *See* Federal Judicial Center, *Reference Manual on Scientific Evidence* ("FJC Manual"), 116, 126-127 (2d ed, 2000). The lower boundary of a 90% confidence interval is related to the significance level of the .05 one-tailed hypothesis test.

[2] Dr. Spencer had no similar finding about my retrofit random sample.

other sample can be much greater than when a single sample is drawn. In completely random samples, if there is a 10% chance of some condition arising in a single, randomly drawn sample, there is nearly a 20% chance (19% to be exact) that one or the other of two samples will exhibit the condition that would be expected to occur 10% of the time in a single sample.[3]

Generally Accepted Documentation of Sampling

9. Below is a list of some sampling methods that provide documentation of the sample selection process. These examples show that many documentation processes for sampling do not maintain recoverable seeds. Some maintain only a description of the sampling process. Many of the sampling processes described below were used for extremely important samples and are relied upon for significant research. Some sampling processes below had significant impacts on the courses of hundreds of thousands of lives.

10. The National Household Education Surveys use a survey sample firm, MSG Genesys, to generate random numbers for some studies. These random numbers are generated and the process is documented in a way that is very similar to the process I followed – the set of random numbers is generated, the list of items picked is documented.

11. Members of my team called MSG Genesys under my direction. MSG Genesys has proprietary software to generate random numbers. We tested the software, which did not produce a recordable seed.

12. Knowledge Networks, a well-known source of survey information for academic and government researchers, performs extensive sampling.[4] Their sample selection process for selection into their panel does not provide a seed. Instead they document the selection process by maintaining a record of the individuals that were sampled, which can then

---

[3] Dr. Spencer agreed with the mathematics behind these probabilities at his deposition (Spencer Deposition, P. 241:4-14.

[4] For one example of a scholarly publication by Harvard professors and researchers based on data obtained from Knowledge Networks see J.K. Hammitt, K. Haninger, and N. Treich, "Effects of Health and Longevity on Risk Tolerance," Geneva Risk and Insurance Review 34(2): 117-139, 2009. This paper is about the critically important issue of how the nation's retirement investments are influenced by evolving health and longevity.

be used to identify the individuals who were chosen.

13. I reviewed a recording of the nationally televised random selection process used to set the order in which men would be drafted into the military during the Vietnam War.[5] The result of the Vietnam Era random selection process has been used in economic research specifically because it is known to be random.[6] The Vietnam Era sample selection process did not use a random seed. Dates of the year were put into individual capsules, poured into a bowl and randomly selected by hand. The outcome of the selection process was used as the record of the selection.

14. The *Handbook of Research Methods in Public Administration* observes the following: "Many popular computer packages such as Excel, SPSS, and Stata also include a random number generator function." The authors provide no admonition to avoid random number generator functions that do not start with a seed.[7]

15. Cornell Professor William Trochim, author of the web-based *Research Methods Knowledge Base,*[8] provides the following random sample selection process: "Many computer programs can generate a series of random numbers. Let's assume you can copy and paste the list of client names into a column in an EXCEL spreadsheet. Then, in the column right next to it paste the function =RAND() which is EXCEL's way of putting a random number between 0 and 1 in the cells. Then, sort both columns -- the list of names and the random number -- by the random numbers. This rearranges the list in random order from the lowest to the highest random number. Then, all you have to do is take the first hundred names in this sorted list."

---

[5] Available at www.youtube.com/watch?v=zVwUEABV9mg (last viewed August 18, 2010).
[6] For scholarly research based on this sample, published in one of the top academic journals in the field of economics See Joshua D. Angrist, "Lifetime Earnings and the Vietnam Era Draft Lottery: Evidence from Social Security Administrative Records," The American Economic Review, vol. 80, No. 3, June 1990, P. 313-336.
[7] Yang, Kaifeng, Gerald J. Miller, Handbook of Research Methods in Public Administration,2nd ed, Boca Raton, Florida: CRC Press, 2008, 217.
[8] Available at http://www.socialresearchmethods.net/kb/sampprob.php (last visited August 18, 2010). Professor Trochim has a Ph.D. in Psychology from Northwestern University.

Documents in Support

16. Attached as Exhibit A is a true and correct copy of excerpts from Federal Judicial Center, *Reference Manual on Scientific Evidence*. Second Edition. 2000.

17. Attached as Exhibit B is a true and correct copy of excerpts from Agresti A. and B. Finlay, Statistical Methods for the Social Sciences, Upper Saddle River, New Jersey: Prentice Hall, 2009.

18. Attached as Exhibit C is a true and correct copy of Joshua D. Angrist, "Lifetime Earnings and the Vietnam Era Draft Lottery: Evidence from Social Security Administrative Records," The American Economic Review, vol. 80, No. 3, June 1990.

19. Attached as Exhibit D is a true and correct copy of J.K. Hammitt, K. Haninger, and N. Treich, "Effects of Health and Longevity on Risk Tolerance," Geneva Risk and Insurance Review , 2009.

20. Attached as Exhibit E is a true and correct copy of excerpts from Yang, Kaifeng, Gerald J. Miller, Handbook of Research Methods in Public Administration, Second Edition. Boca Raton, Florida: CRC Press, 2008.

21. Attached as Exhibit F is a true and correct copy of excerpts from Pindyck R. S. and D.L. Rubinfeld, Econometric Models and Econometric Forecasts, Fourth Edition, Boston, MA: Irwin McGraw-Hill, 1998.

22. Attached as Exhibit G is a true and correct copy of a print out "Probability Sampling" an article published on the website *Research Methods Knowledge Base*, available at http://www.socialresearchmethods.net/kb/sampprob.php (last visited August 19, 2010).

I declare under penalty of perjury under the laws of the United States that the foregoing facts are true and correct, and that this Declaration was executed on August 19, 2010, in Boston, MA.

Daniel S. Levy, PhD