BINGHAM McCUTCHEN LLP
DONN P. PICKETT (SBN 72257)
GEOFFREY M. HOWARD (SBN 157468)
HOLLY A. HOUSE (SBN 136045)
ZACHARY J. ALINDER (SBN 209009)
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
Facsimile: (914) 749-8300
dboies@bsfllp.com
STEVEN C. HOLTZMAN (SBN 144177)
FRED NORTON (SBN 224725)
1999 Harrison St., Suite 900
Oakland, CA 94612
Telephone: (510) 874-1000
Facsimile: (510) 874-1460
sholtzman@bsfllp.com
fnorton@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.	CASE NO. 07-CV-01658 PJH (EDL) **DECLARATION OF DANIEL S. LEVY, PH.D. IN SUPPORT OF MOTION NO. 1: TO EXCLUDE TESTIMONY OF DEFENDANTS' EXPERT STEPHEN CLARKE** Date: September 30, 2010 Time: 9 a.m. Place: Courtroom 3 Judge: Hon. Phyllis J. Hamilton **FILED PURSUANT TO DKT. NO. 915**

I, Daniel S. Levy, Ph.D., declare as follows:

I. INTRODUCTION AND ASSIGNMENT

1. My name is Daniel S. Levy. I am the National Managing Director and a founder of Advanced Analytical Consulting Group, Inc. ("AACG"). 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 on the topics of regression analysis, statistical methods, and damages analysis. I perform and review regression analyses for use in reports to government agencies, academic research, business consulting and legal disputes. I, and my company, are currently engaged in consulting projects for Fortune 500 companies in the United States and internationally in which the main purpose of our work is the construction of advanced econometric models, regression analyses, statistical analyses, large-scale sample design and data collection 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. For the past 30 years I have used regression analysis, for most of that time, on a daily basis, discussing results, designing models, programming regressions and delivering results based on regression models to corporate clients and government agencies. I have worked on hundreds of projects where regression analyses of various types have been a central feature of the research.

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 a declaration in support of Oracle's motion to exclude certain of Mr. Clarke's opinions related to his regression analyses. 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.

II. EXECUTIVE SUMMARY

3. I have reviewed the regression analyses Mr. Clarke presented in his report dated

May 7, 2010. Additionally, I have reviewed the portions of his deposition testimony on June 10, 2010 in which he discussed his regression analyses.[1] My findings are that Mr. Clarke's regression methods used to determine variable costs in the OEMEA and OUSA data are based on a series of mistakes and misconceptions that are so fundamental that they render his estimates of variable costs not only unreliable, but entirely useless; the calculations do not conceptually, or in actual fact, measure variable costs. The analyses do not conform to generally accepted scientific methods used to measure how costs change as revenue change. My disagreements with Mr. Clarke's regression analyses are based on the fact that corrections to almost each and every part of his regression analyses have a significant impact on the results and interpretations. I do not suggest minor changes and I do not propose hypothetical, academic exercises to examine what the impact might be. Instead, my disagreements addresses a fundamental methodological issue that produces an important empirical impact to each of his regression analyses.[2] My opinions can be summarized as follows:

a. Although Mr. Clarke's stated goal in performing his regression analysis is to "apportion the fixed and variable costs,"[3] he performs a type of regression, which he calls a **zero intercept technique,** that is incapable of measuring variable costs as distinct from other costs.

b. This single, fundamental error has a significant empirical impact, erroneously inflating the variable costs he is attempting to measure, at times by almost double.

c. Mr. Clarke's zero intercept technique regression is more than simply biased; it is so profoundly incorrect that it will assign the same variable cost to a set of data that has no variable costs as it would to one that has enormous variable costs. And the variable costs

---

[1] All referenced pages from Mr. Clarke's Report are found in Exhibit A, and all references to Mr. Clarke's deposition are found in Exhibit B to the accompanying Declaration of Holly A. House in Support of Oracle's Mo. No. 1: To Exclude Testimony of Stephen Clarke ("House Decl.").
[2] As cited below, these findings are supported by a number of his statements in his deposition, as well as in his report. I reserve the right to update, supplement, and amend this declaration as additional information becomes available.
[3] House Decl. (Clarke Report), p. 244.

measured using Mr. Clarke's method would not be correct for either. Furthermore, for two sets of data that have identical variable costs, Mr. Clarke's zero intercept technique regression will measure them as having drastically differing variable costs.

d. Mr. Clarke defends the quality of these zero intercept technique regressions with his conclusion that the high R^2 implies that the regression fit the data very well.[4] Mr. Clarke is wrong for at least two reasons.

(1) First, Mr. Clarke misunderstands the definition of R^2 upon which he relies; he provides the definition, and citation for one type of R^2, but calculates another. Mr. Clarke calculates an R^2 that has highly inflated values, while in fact, Mr. Clarke's regression method explains virtually none of the change-in-cost to change-in-revenue relationship it purports to measure.

(2) Second, even if Mr. Clarke had interpreted his R^2 correctly, numerous scholarly works warn against using R^2 as an indication that a regression model has been implemented correctly.

e. Mr. Clarke mistakenly reports the predicted total costs associated with his average revenues from his regression as variable costs when they are actually total costs. He compounds this error by incorrectly describing the difference between his predicted total costs and the average across years of the actual total costs in the source data as his measure of fixed costs for OUSA and OEMEA.

f. In addition to these fundamental mistakes, Mr. Clarke does not know about significant and relevant regression techniques and tests that he should have considered or

---

[4] Mr. Clarke also defends his regression based on his calculated t-statistics. However, forcing a regression line to have a zero intercept also results in a t-statistic that is higher than what the formula in the statistics reference cited by Mr. Clarke produces. Macfie and Nufrio, Applied Statistics for Public Policy, p. 446-447, attached to this Declaration as **Exhibit 1**. For his OEMEA regression, the standard t-statistic formula, cited in Macfie and Nufrio, produces a value less than 25 percent of what Mr. Clarke reported. For OUSA, the standard formula, cited in Macfie and Nufrio, produces a t-statistic that is less than 20 percent of what Mr. Clarke reported. Thus, the use of a "zero intercept technique" in combination with his reliance of high R^2 and t-statistics to validate his models appears to have badly misled Mr. Clarke on the adequacy of his regressions.

investigated.

(1) For the SAP regression, Mr. Clarke does not know of a standard regression model known as "fixed effects."[5] Use of this technique would have substantially changed his results.

(2) While Mr. Clarke does not even test for autocorrelation in his total Oracle regression, correcting for autocorrelation that exists in that regression would have substantially changed his results.

(3) There are a significant number of other statistical conditions Mr. Clarke does not check for, which are standard tests that econometricians perform when testing the validity of regression models and which should have been considered here. Mr. Clarke says he did not test for any of them. (e.g., House Decl., Ex. B (Clarke Deposition) at 933:8-17; 934:24-935:2; 939:7-9; 946:9-11; 957:9-12; 958:15-22) A few of these necessary tests are provided below.[6]

g. The unreliability of Mr. Clarke's SAP and Oracle regressions, including the extent to which the results change when corrected for common econometric problems, demonstrates that his results are not reliable.

III. MR. CLARKE'S ERRANT ATTEMPT TO MEASURE THE VARIABLE COSTS OF OUSA AND OEMEA

4. Mr. Clarke estimates the relationship between total costs and total revenue for OUSA and OEMEA in an attempt to measure variable costs in the relevant range of sales at issue

---

[5] Maddala, G.S., Econometrics, p. 138-139, attached to this Declaration as **Exhibit 2**. Kennedy, P., A Guide to Econometrics, Sixth Edition, pp. 281-285, attached to this Declaration as **Exhibit 3**.

[6] Some texts refer to five broad categories of data issues that are tested by experts who use regression analyses and Clarke tests none of them. Most econometrics books spend several chapters, and in some cases entire books are devoted to determining how to test for violations of these broad groupings and assessing whether there is any solution to eliminating or minimizing their damaging effects. His assertions that his high R^2 allows him to ignore these problems are contrary to the science and practice of econometric modeling and regression analysis he purports to be expert in. For more discussion of these issues, see Kennedy, Peter, A Guide to Econometrics, Sixth Edition, p. 42, attached to this Declaration as **Exhibit 3**.

in this case. The change in costs for a change in a unit of revenue is important to Mr. Clarke's analysis because it directly impacts the profit margins Oracle would have earned if Oracle's support sales revenue had been higher (i.e., it directly impacts Oracle's lost profits damages). Mr. Clarke performs this regression using what he calls the "**zero intercept technique**." As discussed in detail below, performing a regression of total costs on total revenue without an intercept, as Mr. Clarke has done, prevents Mr. Clarke from identifying those costs that vary with the relevant change in revenue from those that do not. Although regression techniques performed *with* an intercept can be used for that purpose, ironically Mr. Clarke has disabled the very feature of a regression analysis that allows identification of the variable costs from the other costs. Therefore his methodology and results are meaningless for the purpose for which they are intended.

5. Therefore, instead of estimating variable costs, as Mr. Clarke states he believes he is doing, he is simply estimating an average cost, which includes both fixed and variable costs over the relevant range in revenue.

6. Mr. Clarke also subtracts his forecasted total cost from a measure of *actual* total costs to get what he believes are the fixed costs. But he actually has nothing of the kind. Mr. Clarke has simply subtracted an estimate of total costs based on his regression from the actual, observed, total costs. Since the regression line in his data does not fit the actual data perfectly, or in fact even very well, there is a difference between Mr. Clarke's estimate of the total costs and the actual total costs. Mr. Clarke attributes this difference between his forecasted total costs and the actual total costs to fixed costs. But it is not; it is simply the difference between his forecasted total costs and the actual, observed total costs.

A. Mr. Clarke's Improper "Zero Intercept Technique"

7. A graphical presentation of the OUSA and OEMEA quarterly total revenue and total cost[7] data analyzed by Mr. Clarke indicates that total costs generally increase with higher

---

[7] Mr. Clarke uses total costs from OEMEA's accounting records and OUSA's accounting records. He says "In my analysis, I analyze the total costs (the dependent variable) against total

(Footnote Continued on Next Page.)

revenue. The purpose of this regression analysis is to help quantify what increase in costs is associated with an increase in sales revenues.

8. As an example, Figure 1 below shows a plot of the OEMEA total costs and total revenue data points used by Mr. Clarke with a regression line that reflects the relationship between total costs and total revenue added through these data points. This is a proper regression *with* an intercept; not the zero intercept technique, Mr. Clarke performed. (Mr. Clarke transforms the data he uses in his regression by dividing both the total costs and total revenues by a data series he calls "U.S. CPI."[8] In this document I work with the same data points that Mr. Clarke used in his regression analysis.)

(Footnote Continued from Previous Page.)

revenues (the independent variable)." House Decl., Ex. A (Clarke Report), p. 278.

[8] For a detailed explanation of Mr. Clark's data series, see *Appendix U-3 - May 7, 2010.xls* to Expert Report of Stephen Clarke, May 7, 2010.

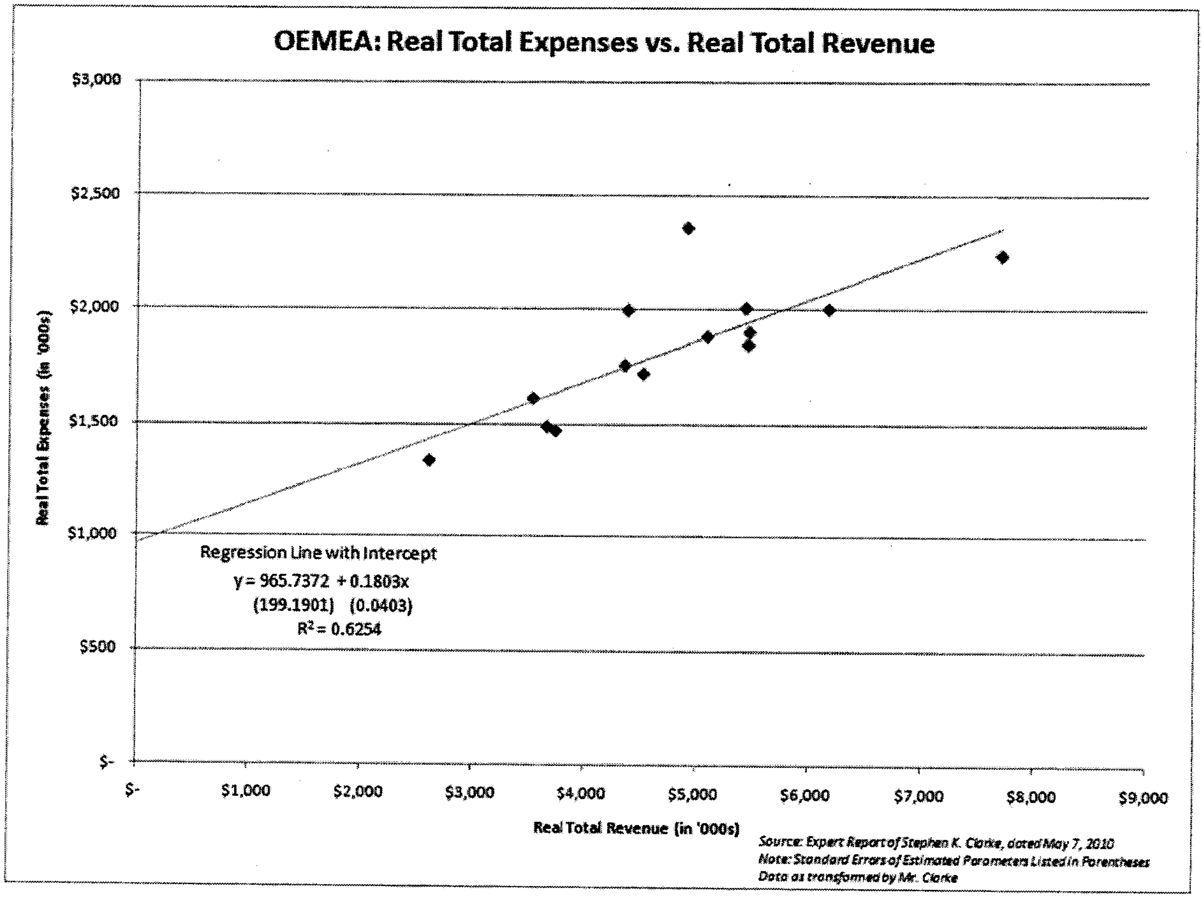


Figure 1

9. The regression line in Figure 1 reflects the general pattern in the data. There is, of course, some distance between each data point and the regression line; this distance is called the "error" by econometricians. Mr. Clarke purports to measure variable costs with an ordinary least squared (OLS) regression. An OLS regression is normally designed to minimize this "error" as much as possible by fitting the regression line that minimizes the total of the distances (squared) from each data point to the regression line. However, Mr. Clarke employs one critical, damaging change to the normal and proper application of the OLS regression technique, which I address below.

10. The equation just below the regression line lists both the value where the line

crosses the vertical axis, 965.74, called the "intercept" and the slope of the line, 0.1803.[9] The slope of the line reflects the amount by which costs increase as revenues increase by one dollar. Equivalently, it reflects the decrease in costs as revenues fall by a dollar. So in this OLS regression, as revenues go up by one dollar, costs increase by about 0.18, or 18 cents. The slope of the line is a piece of information frequently used in the analysis of variable costs. But Mr. Clarke's use of his zero intercept technique measures a different value, which I will discuss in a moment.

11. For the purposes at hand, the intercept in the regressions allows the regression line to fit through the data better. Its specific value may not be of great interest in this case, but it is critical to include the intercept in most settings to obtain a more accurate, unbiased result. Figure 2 shows why. The solid, green line is the same OLS estimate presented in Figure 1. The dashed, red line is the regression line Mr. Clarke estimated using his zero intercept technique. Clearly the regression estimated by Mr. Clarke forces the line to cross the vertical axis at 0, making the line steeper. This is the critical, damaging difference to which I referred above. Mr. Clarke's regression line does not fit the OEMEA data nearly as well as the regression line with the intercept. This means that his line does not reflect the reality presented in the actual data points as well as the regression line with the intercept. A symptom of this problem can be seen in the fact that at the ends of the data particularly there is much more vertical distance ("error") between Mr. Clarke's regression line (dashed, red) and the data points,[10] than between the data

---

[9] The estimated intercept and slope are statistical measures, measured with some standard error. The standard error of the intercept is 199.19. The standard error of the slope is 0.04.

[10] The pattern in Mr. Clarke's residuals should have been a clear warning that his model was badly amiss. As Professor John Rice observes (in a subsection entitled "Assessing the Fit"):

> As an aid in assessing the quality of fit, we make extensive use of the residuals…It is most useful to examine the residuals graphically. Plots of the residuals versus the *x* values may reveal systematic misfit or other ways in which the data do not conform to the standard statistical model. Ideally, the residuals should show no relation to the *x* values, and the plot should look like a horizontal blur.

The pattern in Mr. Clarke's residuals is anything but. The residuals to the right in Figure 2 are all below the zero intercept regression line. The residuals to the left are all above it. Rice, John A.,

(Footnote Continued on Next Page.)

points and the solid green line, which is the standard OLS estimate with an intercept.

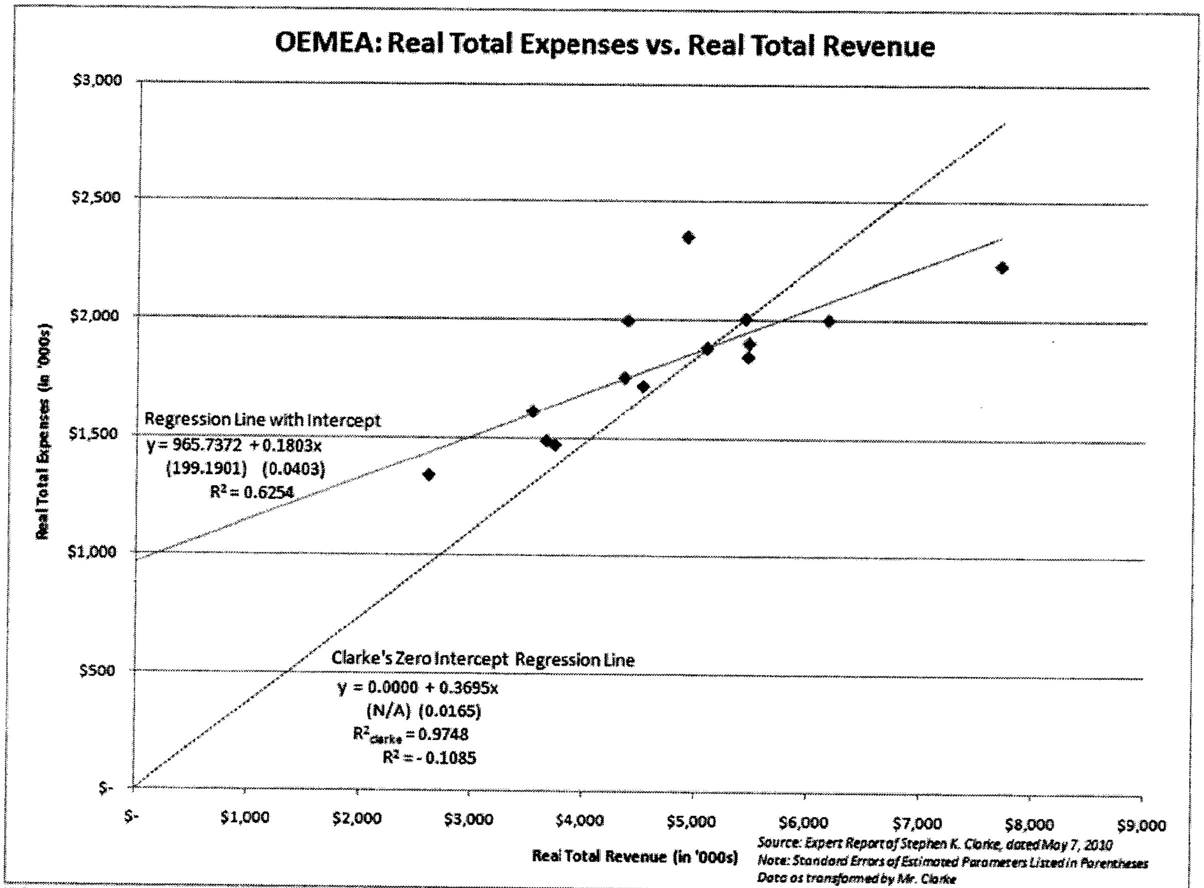


Figure 2

12. By forcing the intercept of the regression line to run through 0, Mr. Clarke has increased the slope of the line, and therefore biased his estimate of OEMEA's variable costs, from 18.03 cents per dollar to 36.95 cents per dollar. Correcting Mr. Clarke's OUSA regression analysis to add an intercept results in a similar reduction in the estimate of variable costs. In other words, Mr. Clarke's regression analysis results in the overstatement of OUSA's and OEMEA's variable costs, and understatement of OUSA's and OEMEA's profit margins applied in his calculation of lost profits damages.[11]

(Footnote Continued from Previous Page.)

Mathematical Statistics and Data Analysis, Second Edition, Belmont, California: Duxbury Press, 1995, p. 515, attached to this Declaration as **Exhibit** 7. (Dr. Spencer's Expert Report cited this source.)

[11] Also presented in this figure are the standard errors for each of the estimated parameters in

(Footnote Continued on Next Page.)

13. There are at least two reasons why the variable costs estimated with the intercept fit this data better than the Mr. Clarke's zero intercept estimate of the variable costs. First, a regression line with an intercept that is allowed to vary so as to fit the data best will always fit the data as well or better than the same regression with the intercept forced through zero. The regression line is determined through a mathematical formula. This formula calculates the intercept and slope such that it allows the regression line to fit the data to minimize the square distances of the data points to the line in aggregate. Forcing the intercept to be 0, as would forcing the intercept to any other value, inherently worsens the fit of the regression line to the data because it prevents the regression from minimizing this distance from the data points to the regression lines (squared).[12] In this case, choosing 0 as the forced intercept inherently forces the slope to be higher, which results in high variable costs, and lower Oracle profit margins.

14. The second reason the standard regression with an intercept fits the data better than Mr. Clarke's zero intercept technique is that the R^2 statistic for Mr. Clarke's zero intercept regression is much lower than the R^2 of a standard OLS regression including an intercept. Although we cannot rely on R^2 alone to tell us whether a regression is specified correctly, we can use it a one measure of how close the data points fall to the regression line. The R^2 is a statistic sometimes used by econometricians. Mr. Clarke, citing Applied Statistics for Public Policy by Macfie and Nufrio, points out in his report that the R^2 "measures the proportion of the total

---

(Footnote Continued from Previous Page.)

parentheses. I have also presented the most common R^2 as discussed in the statistics text Mr. Clarke has cited. This is *not* the R^2 that Mr. Clarke calculates. The R^2 Mr. Clarke calculates is not comparable to the standard R^2 for determining which model fits best. For comparing across regressions, a consistent type of R^2 must be used.

[12] There are some instances in which a regression is estimated without an intercept. These are relatively rare circumstances where it is clear that the pattern in the data *in the relevant range of the estimation* is driven by a relationship that has a zero intercept. At his deposition Mr. Clarke asserted that his **zero intercept technique** regression was justified because if Oracle did not have any revenues in the long run it would eventually stop having costs. (Clarke Deposition, 960:5-962:7). Mr. Clarke's assertion is incorrect as a justification for his regression method. The fact that a company has no costs once it has shut down and its long term contracts have run their course does not mean that in the relevant range of the revenues and costs, the pattern in the data would run through a zero intercept. The fact that this is wrong can be observed in the data. It simply does not fit a pattern that goes through the origin, and there is no theoretical justification that suggests that it should.

variation in the dependent variable (Y) that is explained or accounted for by the total variation in the independent variable (X)."[13] The variation in Y (total costs) that Mr. Clarke is attempting to measure is a variation from the low value 1,337 to the high about 2,354 that is explained by the variation in X (total revenues).[14]

15. If there is a perfect fit of the regression line to the data points, that is all of the data points happen to fall directly on the regression line, the standard R^2 would be 1. However, if there were little or no association between the data points and the regression line, the standard R^2 would be close to 0.

16. Using this standard definition of the R^2 for the regression line *with* the intercept, as I have corrected Mr. Clarke's regression, 63% of the variation in total costs is explained by the variation in total revenue. In contrast, the standard R^2 for the regression line *without* the intercept, used by Mr. Clarke, is ***negative*** 0.11, which means his regression explains virtually nothing of the variation in costs. The regression line *with* the intercept fits the data better than the regression line without the intercept, as performed by Mr. Clarke.[15]

17. Mr. Clarke reports an R^2 for his OEMEA regression of 0.97, not -0.11. The reason he reports a different value is because the R^2 he presents does not "measure [...] the proportion of the total variation in the dependent variable (Y) that is explained or accounted for by the total variation in the independent variable (X),"[16] as stated by Mr. Clarke. The R^2 Mr. Clarke reports reflects a different calculation that has a very different meaning and a non-comparable scale related to whether Mr. Clarke's regression analysis fits the pattern of variable cost presented in the OUSA and OEMEA data. The R^2 calculation used by Mr. Clarke is used in

---

[13] House Decl., Ex. A (Clarke Report), p. 244, n. 1099. Mr. Clarke cites to Applied Statistics for Public Policy, Macfie and Nufrio, p. 398, attached to this Declaration as **Exhibit 1**.

[14] These values are scaled to match the data Mr. Clarke provided as used in his regression analysis.

[15] Indeed, a model with only an intercept, i.e., a flat line, fits the data better than Mr. Clarke's regression.

[16] House Decl., Ex. A (Clarke Report), p. 244, n. 1099. Mr. Clarke cites to Applied Statistics for Public Policy, Macfie and Nufrio, p. 398, attached to this Declaration as **Exhibit 1**.

some situations and can be found in econometrics text books, but it simply means something different than what Mr. Clarke says it means. Using the standard definition of R^2 that Mr. Clarke cites from the Macfie and Nufrio text, his R^2 for the OEMEA and OUSA regressions are much lower than he reports.

B. Examples That Illustrate Mr. Clarke's Zero Intercept Technique Do Not Measure Variable Costs

18. To illustrate the effect of forcing the intercept through zero and the use of the alternative R^2, which I call "R^2_{Clarke}," I turn to a few examples. These examples clearly demonstrate that the zero intercept regression model does not accurately measure variable costs, and that the R^2_{Clarke} will be very large even when the regression line has almost no relationship to the data. These two features of Mr. Clarke's regression analysis render it useless for the purposes of measuring how costs vary with revenues in the relevant range. The unreliability of Mr. Clarke's methods, and resulting irrelevance of his results, are exhibited by the fact that his methods yield measures of slope that are exactly the same for patterns of data that obviously have very different slopes. While the following examples, including those discussed in Section III.C below, are based on Mr. Clarke's OEMEA regression, the issues demonstrated equally apply to Mr. Clarke's OUSA regression.

19. For example, Figure 3 provides an example of a set of data points where there is little to no variation in costs as revenues change; there are virtually no variable costs over the range of data. But Mr. Clarke's **zero intercept technique** provides a different answer.

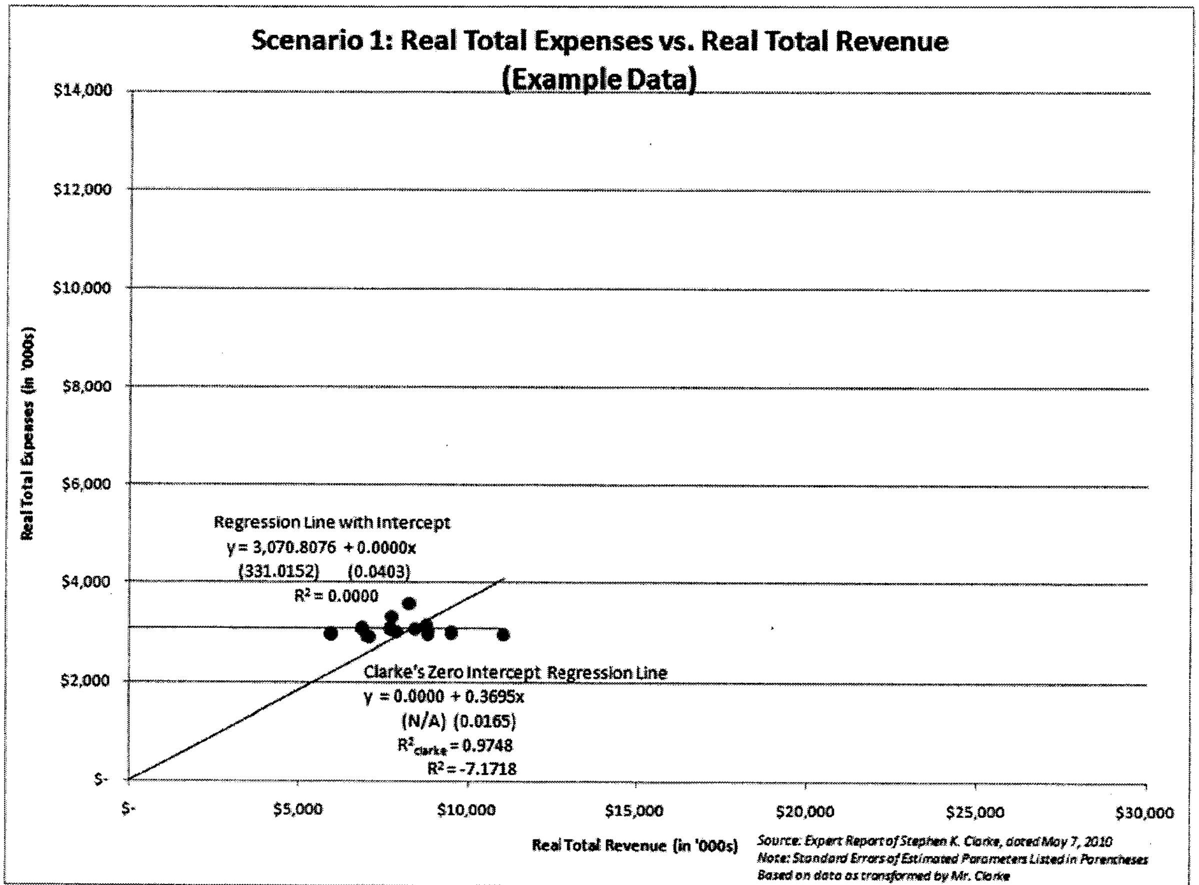


Figure 3

20. Despite the fact that there is very little change in costs with a change in revenue in these data, Mr. Clarke's zero intercept regression model estimates that costs vary by 36.95 cents for each dollar change in revenue, exactly the same variation in costs associated with the change in revenues Mr. Clarke measured in the OEMEA data. The regression with the intercept provides a very different slope than the one produced with Mr. Clarke's zero intercept technique. In addition, Mr. Clarke's R^2_{Clarke} indicates (to him) that his regression line fits this data very well, with an R^2_{Clarke} =0.97. Again, this is exactly the same R^2_{Clarke} that Mr. Clarke measured in the OEMEA data. The standard R^2 for Mr. Clarke's zero intercept regression running through the data depicted in Figure 3 is ***negative*** 7.17. This means the regression line forced through the origin explains little of the variation in costs. Clearly, Mr. Clarke's zero intercept regression technique is not connected to the relationship between costs and revenues observed in the data. Not only does his method produce a completely incorrect estimate of the slope of this data, his

calculated R^2_{Clarke} also erroneously indicates that his regression line fits the data very well.

21. This example is not an isolated case. Figure 4 shows two additional cases alongside the original OEMEA data. Each set of colored data points is estimated as a separate set of example data. In each case, Mr. Clarke's zero intercept regression model and R^2_{Clarke} methods produce the exact same measured relationship between costs and revenues as his OEMEA analysis. The sets of example data include both 0 relationship between the change in costs and the change in revenues and a very large relationship between the change in costs and the change in revenue – that is, very different from the OEMEA data – but Mr. Clarke's zero intercept technique produces results that are unaltered by these different patterns of variable costs. Furthermore, the R^2_{Clarke}= 0.97 is also impervious to the differences presented in the three additional data sets plotted in Figure 4. One of the example data sets in Figure 4 even has a *negative* slope (i.e., costs decrease as revenues increase). But again, Mr. Clarke's zero intercept technique regression measures the relationship between changing costs and changing revenues as 0.3695, same as all the other sets of data in Figure 4.

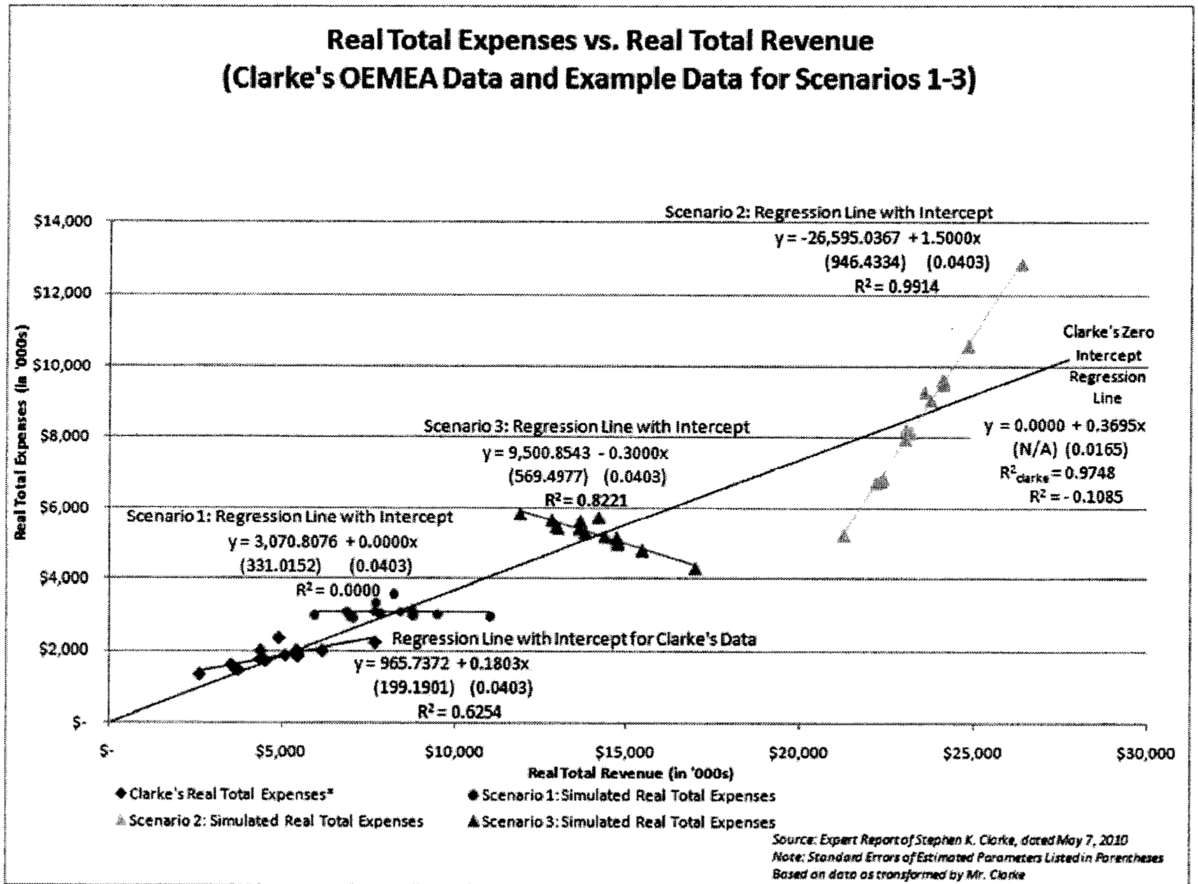


Figure 4

22. It is obvious that these four sets of data do not depict the same variation in costs with variation in revenue, but Mr. Clarke's zero intercept technique regression, depicted by the heavy black like emanating out of the origin, measures them as having exactly the same variable cost.

C. Mr. Clarke's Zero Intercept Technique Produces Differing Variable Costs When Variable Costs Are The Same

23. I do not want to leave the impression that Mr. Clarke's zero intercept technique will always produce the same relationship between costs and revenues. Figure 5 presents two sets of data, which when estimated using the Clarke zero intercept technique and R^2_{Clarke}, would be estimated to have differing variable costs and R^2_{Clarke}. The first set of data is the original OEMEA data that Mr. Clarke used in his zero intercept regression.

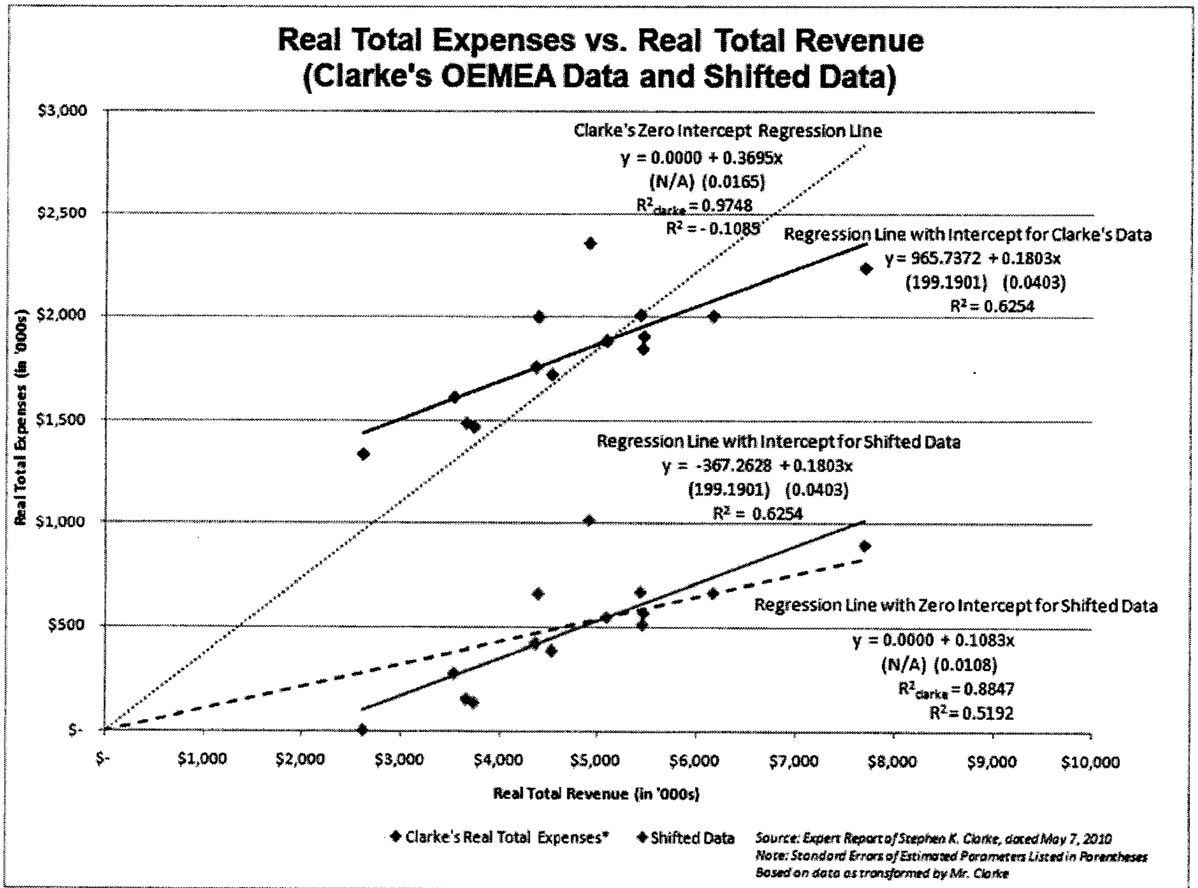


Figure 5

24. The set of data further down the vertical axis is simply the OEMEA data shifted vertically down the page. They are otherwise exactly the same. The slope between any pair of points in the upper set of data is exactly the same as the slope between analogous pairs of points in the lower set of data. Clearly, the change in costs associated with the change in revenue is the same for both sets of points. The standard OLS regression with an intercept reflects the fact that the slope of the relationship between costs and revenues is exactly the same in both the upper and lower set of points. Both sets of data reflect an actual change in costs of 18.03 cents for a one dollar change in revenues. In addition, the standard R^2 ("$R^2_{standard}$") depicts the exact same fit of the regression line to the data. $R^2_{standard} = 0.63$. The patterns in these two sets of data are identical so any valid measure of the relationship between the change in costs and the change in revenues of these patterns should be the same as well. A regression analysis of this same data using Mr. Clarke's inappropriate zero intercept technique and R^2_{Clarke}, paints an entirely different

picture. Using Mr. Clarke's zero intercept regression methods on the upper set of data (the actual OEMEA data) produces his original findings; a slope of 0.3695. Alternatively for the lower set of data, Mr. Clarkes zero intercept regression model produces a slope of 0.1083 measuring a much lower 11 cent change in costs per dollar change in revenues. Mr. Clarke's zero intercept technique is clearly producing nonsensical results; incorrect for both the upper and the lower set of data. By forcing the intercept to zero, Mr. Clarke imposes a relationship on the data that does not actually exist in the data. It biases the estimates of the relationship between the change in revenues and the change in costs. This is a critical number in Mr. Clarke's calculation of variable costs. Once this number has been biased in such a fundamental fashion, the rest of his calculations used to determine variable costs are hopelessly fouled. There are more steps to Mr. Clarke's calculation of variable costs in which he adds further errors and misinterpretations, which further damage his calculation, but already at this point his calculation of the relationship between the change in costs and the change in revenue is irrevocably damaged. His results are not a reflection of the relationships in the data, as he claims, but rather simply the aftermath of errant assumptions and methods.

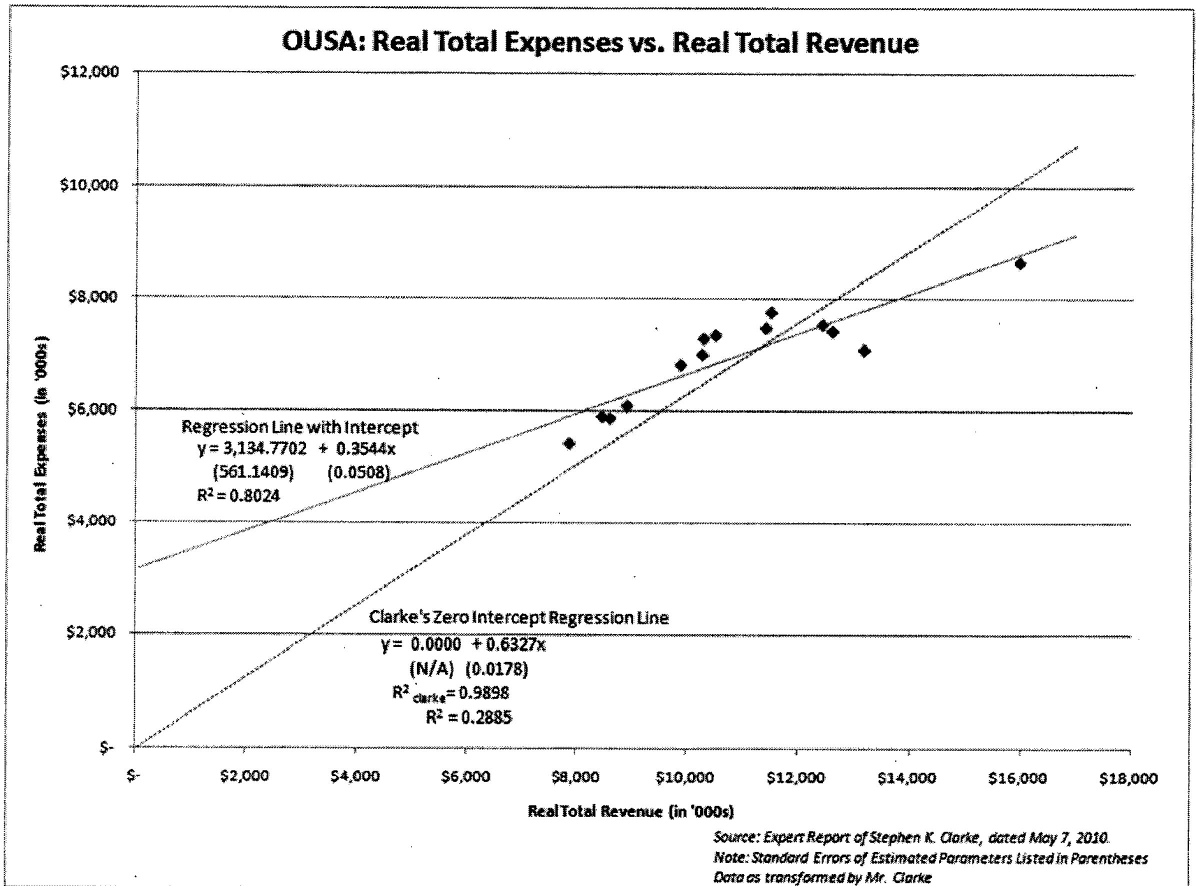


Figure 6

25. Similar to what Mr. Clarke does in his OEMEA regression analysis, Figure 6 shows that he also forces the intercept in the OUSA regression to be zero (dashed red line). Again, this causes the slope of the line to erroneously increase, and therefore the variable costs to increase from 35.4 cents per dollar to 63.3cents per dollar. Once again, Mr. Clarke's regression line does **not** fit the data as well as the regression line with the intercept. This implies that the regression line with the intercept, which estimates that costs vary by 35.4 cents per dollar of revenue, is a better reflection of reality presented in the actual data points. Just as with the OEMEA regression, in this case the standard R^2 of the line with an intercept ($R^2_{standard}$=0.802) is higher than the standard R^2 of the line with the intercept forced through zero, $R^2_{standard}$=0.289. Further, as explained above, the R^2 reported by Mr. Clarke is not what he claims it to be as described above.

D. Econometric Literature Warns Against Using a Zero Intercept

26. Many econometric texts warn about problems that occur when the intercept of a regression is restricted to zero. Most point out that econometric tests can be performed to determine whether the intercept really is zero. These texts also state that restricting the intercept to zero, without analyzing whether such a restriction is actually reflective of the process driving the pattern in the data, can severely bias the estimated relationship from reality. Furthermore, it is well-known that even if the pattern in the data does point through zero, allowing the regression to estimate the intercept will not bias the measure of variable costs. The estimated variance may be large.[17]

27. Professors Snedecor and Cochran write the following in their book, Statistical Methods:

> "This model [zero intercept model] should not be adopted without careful inspection of the data, since complications can arise. If the sample values of *X* are all some distance from zero, plotting may show that a straight line through the origin is a poor fit, although a straight line that is not forced to go through the origin seems adequate. The explanation may be that the population relation between *Y* and *X* is curved, the curvature being marked near zero but slight in the range within which *X* has been measured…It is sometimes useful to test the null hypothesis that the line, assumed straight, goes through the origin."[18]

28. In fact, plots of Mr. Clarke's data for OUSA and OEMEA in figures above demonstrate that assuming the constant is zero is very likely to be erroneous for these data. Further, Mr. Clarke does not test whether or not the constant should be included. He simply, and

---

[17] "One serious drawback with regression through the origin is that, if the intercept β_0 in the population model is different from zero, then the OLS estimators of the slope parameter will be biased. The bias can be severe in some cases. The cost of estimating an intercept when β_0 is truly zero is that the variances of the OLS slope estimators are larger." Wooldridge, J.M., Introductory Econometrics, Fourth Edition, p. 83. Also, "Obtaining an estimate of β_1 using regression through the origin is not done very often in applied work, and for good reason: if the intercept $\beta_0 \neq 0$, then $\hat{\beta}_1$ is a biased estimator of β_1." Wooldridge, J.M., Introductory Econometrics, Fourth Edition, p. 59, attached to this Declaration as **Exhibit 4**.

[18] Snedecor, G.W. and W.G. Cochran, Statistical Methods, Sixth Edition, p. 166, attached to this Declaration as **Exhibit 5**.

improperly, *assumes* it is zero.

29. Similarly, Professor Kennedy observes:

> "Sometimes economic theory suggests that the intercept in a regression is zero…Practitioners usually include an intercept, however. Why? It is possible that a relevant explanatory variable was omitted, creating bias. This bias can be alleviated (but not eliminated) by including an intercept term; no bias is created by including an unnecessary intercept."[19]

30. Mr. Clarke's assumption that the intercept is zero is clearly one that econometricians warn against. By excluding an intercept, without testing whether it really should be excluded, Mr. Clarke has biased his analyses to yield results that would not be arrived at by an experienced econometrician. Perhaps most importantly as discussed in detail above, by removing the intercept, Mr. Clarke has dismantled the mechanism in regression analysis that allows the variable costs to be measured separately from other costs.

E. Mr. Clarke Does NOT Estimate Variable and Fixed Costs

31. Mr. Clarke next uses the estimated slopes of the OUSA and OEMEA regressions as if they reflect an attempt to measure what he calls the relevant margin. But after all of Mr. Clarke's regression machinations, he still has done nothing that separates the variable costs from the rest of the total costs. He has simply estimated a regression that shows the relationship between the total costs and total revenues. The regression line that Mr. Clarke estimates was as follows: Total Costs = 0.3695 x Total Revenue + error. Having zeroed out the intercept, the slope of Mr. Clarke's regression line (0.3695) is relegated to measuring the average cost, which includes both fixed and variable costs. It is not a measure of variable costs.

32. Mr. Clarke reports his estimate of fixed costs to be the difference between the average total costs in his source data and his estimate of total costs (which he incorrectly claims is an estimate of variable costs). There is no sense in which the difference between these two measures of *total costs* could be construed to be the difference between the total cost and the

---

[19] Kennedy, Peter, A Guide to Econometrics, Sixth Edition, p. 109-110, attached to this Declaration as **Exhibit 3**.

variable costs because there is no calculation performed by Mr. Clarke that separates out the variable costs. At this point Mr. Clarke's analysis is estimating no value of fixed or variable costs that can be used for any purpose. His calculation simply reflects the difference between a predicted value for total costs on a regression line and the actual average value of those same total costs. The depth of the errors in Mr. Clarke's calculations goes far beyond creating numbers that are biased, or flawed or ones that could have been estimated much better. The numbers Mr. Clarke calculates are completely useless for his purposes because they simply do not measure how costs change with revenues.

IV. MR. CLARKE'S SAP AND ORACLE REGRESSIONS ARE UNRELIABLE

33. Mr. Clarke's presentation of his regression analyses for SAP and for Oracle in his report and in his deposition testimony clearly indicates that he lacks the expertise necessary to apply and interpret the results of regression analyses appropriated. He is not an expert in the field of econometrics. This finding is supported by a number of facts.

A. Lack of Knowledge of Fixed Effects

34. Mr. Clarke indicated in his deposition that he is not familiar with a fundamental concept in econometrics known as "fixed effects." House Decl., Ex. B, (Clarke Depo.) at 935:3-4. Mr. Clarke acknowledges that the SAP data is what is referred to as "panel data", that is, data for sixteen different subsidiaries across the globe. His regression analysis fails to take into account the panel feature of the data, that is, he *assumes* that relationship between expenses and revenues are identical across all sixteen SAP subsidiaries. Econometric theory dictates that he should have tested whether there are differences across these subsidiaries. When a fixed effects regression (with fixed effects for each subsidiary) is performed using Mr. Clarke's SAP data, it lowers his coefficient of log of real total revenues from 0.95 to 0.80.[20]

35. Statisticians and economists have developed a variety of methods to allow experts

---

[20] The standard error of the new estimates I performed can be seen in Appendix 1. The standard error of the estimated slope is .01. The regression is included in Appendix 1. The F-test between Mr. Clarke's model and the fixed effects model indicates that the fixed effects model statistically significantly fits the data better than Mr. Clarke's model at more than the .01 significance level.

to test for differences in intercepts across subsets of the data. One such model is known as the fixed effects model. Basic econometrics textbooks discuss the use of fixed effects models.[21] The fact that Mr. Clarke initially said he had not heard of fixed effects calls into question his ability to use regression analysis appropriately for this data. Later in his deposition testimony, Mr. Clarke said he had heard of fixed effects but had never used it.[22] His only justification for not using a fixed effects model is that the R^2s are high in his model.[23] However, no such justification exists. High R^2s -- particularly the type used by Mr. Clarke– are not indicative of a good model or a good estimate of the variable of interest.[24]

36. Econometrics texts that discuss fixed effects show that ignoring fixed effects will cause the expert or researcher to incorrectly estimate the slope of the regression line that fits the data.

37. The following graph, a copy of an example from a well-known statistical text written by G. S. Maddala, shows how failure to include fixed effects can bias the estimate of variable costs upward.

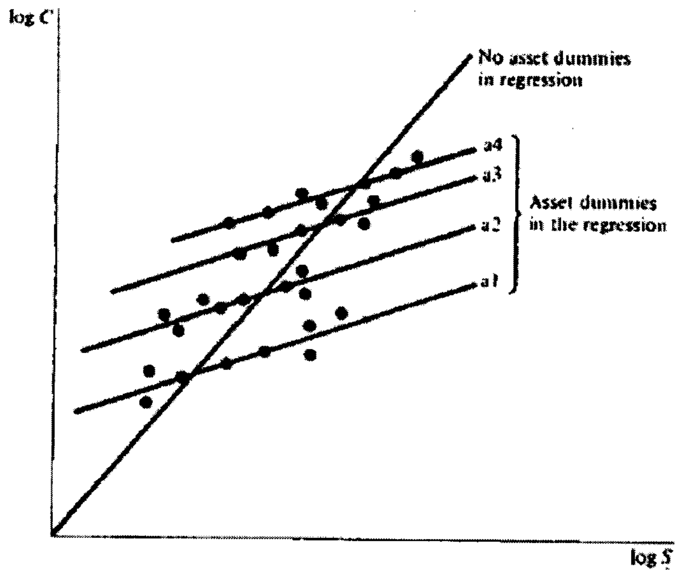
---

[21] Maddala, G.S., Econometrics, p. 138-139, attached to this Declaration as **Exhibit 2**; Kennedy, P., A Guide to Econometrics, Sixth Edition, pp. 281-285, attached to this Declaration as **Exhibit 3**.

[22] House Decl., Ex. B (Clarke Depo.) at 943:23-944:7.

[23] House Decl., Ex. B (Clarke Depo.) at 944:16-18.

[24] Kennedy, P., A Guide to Econometrics, Sixth Edition, p. 27, attached to this Declaration as **Exhibit 3**.



38. When all of the data is estimated together in a single regression, the slope of the regression is errantly estimated, in this case, to be much greater than actual. When separate fixed effects for each set of data are estimated, represented by the multiple regression lines, the slope is more accurately estimated to be flatter. The same pattern occurs in the SAP data that Mr. Clarke used. When fixed effects are included in Mr. Clarke's regression, the slopes of the regression lines, which are now more accurately estimated, become flatter.[25] Mr. Clarke's failure to even investigate the bias he imposed on his regression estimates by failing to even test for the benefits of using fixed effects once again demonstrates his lack of expertise, and renders his analysis faulty and his regression estimates unreliable.

B. Incorrect Interpretation of High R^2s

39. Mr. Clarke purports to apply principles and methods in accordance with professional standards, and yet reaches a conclusion that true experts in the field would not reach. For example, he repeatedly claims that R^2 of the regression is a good way of determining

---

[25] For a detailed discussion of fixed effects that would apply exactly as in Mr. Clarke's SAP regression analysis, see Maddala, G.S., Econometrics, p. 139, attached to this Declaration as **Exhibit 2**.

that he has a meaningful relationship between total revenue and total costs.[26] The R^2 of a regression is a measure of "goodness of fit" – that is, a measure of how well the variation in the dependent variable is explained by the explanatory variables. However, basic econometrics textbooks caution researchers against using the R^2 as a means of determining whether the coefficient of the independent variables is meaningful.

40. Professor Kennedy states that:

> "In general, econometricians are interested in obtaining 'good' parameter estimates where 'good' is not defined in terms of R^2. Consequently the measure of R^2 is not of much importance in econometrics. Unfortunately, however, many practitioners act as though it is important, for reasons that are not entirely clear, as noted by Cramer (1987, p. 253): 'These measures of goodness of fit have a fatal attraction. Although it is generally conceded among insiders that they do not mean a thing, high values are still a source of pride and satisfaction to their authors, however hard they may try to conceal these feelings.'"[27]

C. Incorrect Interpretation of High R^2 in the Presence of Autocorrelation

41. Mr. Clarke states the summary statistics of his models but does not explain how the models work or interpret the results correctly. For example, Mr. Clarke says that he does not need to check for autocorrelation because his t-scores and R^2s are high but also acknowledges that autocorrelation causes high R^2 and t-scores.[28]

42. Autocorrelation is something that experts and researchers check for and, when identified, they use different techniques to correct for autocorrelation. The R^2 cannot be used as a way to determine *whether* autocorrelation should be corrected for. In fact, econometricians are particularly suspicious of high R^2s in the presence of autocorrelation and warn against interpreting them positively.[29]

---

[26] House Decl., Ex. B (Clarke Depo.) at 934:10-19 ; 934:24-935:2; 935:12-18; 944:8-18; 948:13-24.

[27] Kennedy, P., A Guide to Econometrics, Sixth Edition, p. 27, attached to this Declaration as **Exhibit 3**.

[28] House Decl., Ex. B (Clarke Depo.) at 934:10-19; 948:13-949:4.

[29] "What is a high R^2? There is no generally accepted answer to this question. In dealing with time series data, very high R^2s are not unusual, because of common trends. Ames and Reiter

(Footnote Continued on Next Page.)

43. Mr. Clarke dismisses the importance of the fact that there is autocorrelation in his data because he observes high R^2's. However, he fails to realize that a trained econometrician would, in fact, do the opposite – that is, a trained econometrician would search for techniques to correct for autocorrelation particularly when the R^2 is high.[30]

44. The Durbin-Watson statistic is a technical calculation that informs experts on whether or not they need to worry about autocorrelation. In Mr. Clarke's total Oracle regression, the Durbin-Watson statistic is relatively close to 0, 0.86, which indicates autocorrelation. In this instance, methods such as first differences are used to correct for this issue. There are at least two possible ways to correct for the autocorrelation in Mr. Clarke's regression. These corrections reduce his estimate of the coefficient of log of real total revenues from 0.79 by at least 22%.[31] As Professor Maddala states in his book: "However, if the Durbin-Watson statistic is very low, it often implies a misspecified equation no matter what the value of R^2 is. In such

---

(Footnote Continued from Previous Page.)

(1961) found, for example, that on average the R^2 of a relationship between a randomly chosen variable and its own value lagged one period is about 0.7, and that an R^2 in excess of 0.5 could be obtained by selecting an economic time series and regressing it against two to six other randomly selected economic time series." Kennedy, P., A Guide to Econometrics, Sixth Edition, p. 26, attached to this Declaration as **Exhibit 3**. Also, Professor Maddala's book says "Another important thing to note is that usually with time-series data one gets good R^2's when the regressions are estimated with the levels y_t and x_t but one gets poor R^2's if the regressions are estimated in first differences $(y_t - y_{t-1})$ and $(x_t - x_{t-1})$. Since usually a high R^2 is considered as proof of a strong relationship between the variables under investigation, there is a strong tendency to estimate the regression in levels rather than the first differences. This is sometimes called the 'R^2 syndrome.' However, if the Durbin-Watson statistic is very low, it often implies a misspecified equation, no matter what the value of the R^2." Maddala, G.S., Econometrics, p. 92, attached to this Declaration as **Exhibit 2**.

[30] "One compelling reason for taking first differences of trending variables is the phenomenon of **spurious regression**. It should be obvious that if two variables, say y_t and x_t, both trend upward, a regression of y_t on x_y is very likely to find a significant relationship between them, even if the only thing they have in common is the upward trend." R. Davidson and J. G. McKinnon, Estimation and Inference in Econometrics, p. 670-671, attached to this Declaration as **Exhibit 6**. These econometricians are simply stating that a regression of two variables that are observed over time, such as Revenues and Expenses, will likely produce a high R^2 but that is not indicative of a good regression model. This relationship could be spurious. One way to avoid misinterpretation is to use something like a first differences approach.

[31] The 22% reduction brings the coefficient down to .61 (Standard Error of estimate is .0445). The regression results for these two methods are presented in Appendix 2. The Durbin-Watson statistics for both alternative specifications are better than Mr. Clarke's.

cases one should estimate the regression equation in first differences."[32]

45. However, Mr. Clarke claims that he does not need to worry about this because his R^2s are so high. This is despite the fact that he acknowledged more than once that autocorrelation can cause high R^2s.

46. When asked why he does not use first differences, Mr. Clarke's response is – "Because my – my reading of this data with the t stat where it was, was that I didn't need to do that." (House Decl., Ex. B (Clarke Depo.) at 958:25-959:2) In fact, Mr. Clarke does need to use some technique for correcting for autocorrelation in his total Oracle data because a mis-specified regression can result in misleading t-statistics. Results correcting for this problem would demonstrate that he has over estimated the variable costs and show he has significant gaps in his understanding of regressions.

V. ADDITIONAL ERRORS IN MR. CLARKE'S REGRESSION ANALYSIS

A. Assumes the Results of his Regressions Apply for his Purpose

47. On more than one occasion during his deposition, Mr. Clarke stated that he assumed the results of his regression were appropriate for his purpose. Yet Mr. Clarke performed no analysis to determine the reasonableness of his regression results. ["Q: Once you ran the regression analyses and developed the relationship between the revenues and costs, did you do any further investigation of that relationship, or did you just – not just, but did you accept the results of the regressions? A: I assumed that the results of my analysis were appropriate for my purposes." (House Decl., Ex. B (Clarke Depo.) at 925:1-9)] However, experts who use regression analysis investigate the relationships among the variables they use to ensure that the regression analysis is meaningful. Then they interpret the results to inform them of what can be gleaned from the data. Professor Kennedy discusses the need to test a model. He says "the model is continually respecified until a battery of diagnostic tests allows a researcher to conclude that the model is satisfactory on several specific criteria (discussed in general notes), in which

---

[32] Maddala, G.S., Econometrics, p. 92, attached to this Declaration as **Exhibit 2**.

case it is said to be "congruent" with the evidence."[33]

B. Does not report confidence intervals

48. Defendant's own sampling expert, Dr. Bruce Spencer, says that estimated numbers reported without confidence intervals or standard errors are improper and appear to look like "unconditional truth".[34] Although Mr. Clarke reports sample errors for his estimates in his regression results, he does not discuss the effect these sampling errors have on his measurement of damages. House Decl., Ex. B (Clarke Depo.) at 949:22-950:7. Based on Dr. Spencer's standards, Mr. Clarke fails miserably to report his estimates with the appropriate standard errors.

C. Lack of Understanding of his Log-Log Model

49. Mr. Clarke does not understand the relationship between the variables in his log-log model. He is not able to clearly explain the impact of the change in the coefficient of the log of revenue in his regression models for SAP and total Oracle. (House Decl., Ex. B (Clarke Depo.) at 949:5-21). Mr. Clarke contradicts himself between his deposition, where he says that the intercept value from his log-log regression is meaningless, and his report, where he claims that the interpretation of the intercept is that of fixed costs. (House Decl., Ex. B (Clarke Depo.) at 962:10-963:2 and Ex. A (Clarke Report) at p. 244). The intercept of a regression is not meaningless. It has a role in the regression and cannot be ignored in his calculation of predicted values as Mr. Clarke does.

D. Lack of Understanding of F-Test

50. Mr. Clarke says the F-test does not apply and is not used to check for a model specification because his analysis has only one variable. (House Decl., Ex. B (Clarke Depo.) at 940:3-11) This statement is incorrect. For example, in the context of a fixed effects model, an F-test allows the expert to check whether all the fixed effects intercepts are the same. It is not

---

[33] For a broader discussion of how econometricians test models, see the section entitled "Test, Test, Test." Kennedy, P., A Guide to Econometrics, Sixth Edition, p. 73, attached to this Declaration as **Exhibit 3**.

[34] House Decl., Ex. I (Expert Report of Bruce Spencer, March 17, 2010) at p. 43.

dependent on whether you have one explanatory variable or many explanatory variables.[35] Mr. Clarke also says that the F-statistic reported in his own regression results is meaningless, which is not true. (House Decl., Ex. B (Clarke Depo.) at 942:5-14) His claim that it is meaningless is further indication of his lack of understanding of regression analysis.[36]

VI. CONCLUSION

51. I have found that Mr. Clarke's regression analyses are unreliable and unusable for the purpose for which they were intended. These criticisms are not based on small, minor changes to his regression models. Rather, the issues I found with Mr. Clarke's regression analysis reflect his lack of knowledge of the fundamentals of econometrics, which have a significant impact on his estimate of the relevant profit margins estimated by Mr. Clarke. Mr. Clarke's lack of econometric knowledge leads him to make numerous errors in his analysis, which prevent him from accomplishing the main goal of his regression analysis, estimation of how costs change as revenues change. His numerous errors, which have a significant empirical impact on his results, his lack of knowledge of the econometric tools that he attempts to use, and his reliance on baseless assumptions render his regression analyses at best unreliable and unusable and at worst, in the case of his OEMEA and OUSA regressions, completely meaningless.

---

[35] Dr. Kennedy says "An F test, structured in the usual way, can be used to test whether or not the vector with elements α_0, α_1, and α_2 is equal to the zero vector." Kennedy, P., A Guide to Econometrics, Sixth Edition, p. 238, attached to this Declaration as **Exhibit 3**.

[36] Dr. Kennedy says "A special case of the F statistic is automatically reported by most regression packages – the F statistic for the 'overall significance' of the regression." This F statistic tests the hypothesis that all the slope coefficients are zero. The constrained regression in this case would have only an intercept." Kennedy, P., A Guide to Econometrics, Sixth Edition, p.63, attached to this Declaration as **Exhibit 3**. Indeed, Macfie and Nufrio, whom as I noted above Mr. Clarke cites when defining R^2, make essentially the same observation in their discussion of simple regression models with only one explanatory variable. Macfie and Nufrio, op. cit. at pp. 451-454, attached to this Declaration as **Exhibit 1**. In particular, the authors observe: "Although the F-test has greater use in the evaluation of multivariate regression equations, it can easily be introduced and applied to simple regression analysis." This explanation reinforces my conclusion that Mr. Clarke is completely wrong in deeming the F-statistic to be meaningless.

VII. APPENDIX 1 - SAP FIXED EFFECTS REGRESSION

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.9976
R Square	0.9953
Adjusted R Square	0.9951
Standard Error	0.1127
Observations	448

Calculation of Significance of Fixed Effects with Chow test

	SS	df	
Restricted SS regression*	1143.36		
Unrestricted SS reg	1146.53		
Numerator	3.17	15	0.2113
Denominator - SS of resid	5.47	431	0.0127
		F statistic	16.65
		P-value	0.0000

ANOVA

	df	SS	MS	F	Significance F
Regression	16	1146.533176	71.6583235	5645.01213	0
Residual	431	5.47115519	0.0126941		
Total	447	1152.004332			

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	2.93	0.20	14.55	0.00	2.53	3.33	2.53	3.33
LN(Revenue)	0.80	0.01	59.68	0.00	0.77	0.82	0.77	0.82
AG	0.12	0.03	3.74	0.00	0.05	0.18	0.05	0.18
AU	-0.49	0.04	-10.82	0.00	-0.57	-0.40	-0.57	-0.40
CA	-0.41	0.04	-9.92	0.00	-0.49	-0.33	-0.49	-0.33
CH	-0.38	0.04	-9.99	0.00	-0.46	-0.31	-0.46	-0.31
DE	-0.14	0.03	-4.59	0.00	-0.20	-0.08	-0.20	-0.08
FR	-0.33	0.04	-8.64	0.00	-0.41	-0.26	-0.41	-0.26
HR	-0.98	0.09	-11.13	0.00	-1.15	-0.80	-1.15	-0.80
IT	-0.46	0.04	-11.13	0.00	-0.55	-0.38	-0.55	-0.38
JP	-0.32	0.04	-8.71	0.00	-0.39	-0.24	-0.39	-0.24
NL	-0.50	0.04	-11.74	0.00	-0.58	-0.41	-0.58	-0.41
NZ	-0.90	0.07	-13.35	0.00	-1.04	-0.77	-1.04	-0.77
PS	-0.44	0.04	-10.27	0.00	-0.52	-0.35	-0.52	-0.35
SG	-0.52	0.05	-10.02	0.00	-0.63	-0.42	-0.63	-0.42
ST	-0.63	0.06	-11.08	0.00	-0.74	-0.52	-0.74	-0.52
SW	-0.54	0.05	-10.38	0.00	-0.64	-0.44	-0.64	-0.44

Ln Real Revenue*	0.95
Percentage drop comapred to Clarke's results	16%

* See Clarke's Appendix M-9 - May 7, 2010.pdf

VIII. APPENDIX 2 - TOTAL ORACLE REGRESSIONS CORRECTING FOR AUTOCORRELATION

Table 1. First Difference Adjustment

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.8869
R Square	0.7866
Adjusted R Square	0.7823
Standard Error	0.0603
Observations	52

Durbin-Watson statistic	2.4094

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	0.6696	0.6696	184.2745	0.0000
Residual	50	0.1817	0.0036		
Total	51	0.8513			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	0.0074	0.0084	0.8836	0.3812	-0.0095	0.0243	-0.0095	0.0243
X Variable 1	0.4944	0.0364	13.5748	0.0000	0.4213	0.5676	0.4213	0.5676
Percentage drop comapred to Clarke's results*	37%							

X Variable (X_t): (ln(Real Revenue_t) - ln(Real Revenue_t-1))

Dependent Variable (Y_t): (ln(Real Expenses_t) - ln(Real Expenses_t-1))

Table 2. Quasi First Difference Adjustment

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.8892
R Square	0.7907
Adjusted R Square	0.7865
Standard Error	0.0707
Observations	52

Durbin-Watson statistic	1.5960

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	0.9426	0.9426	188.8349	0.0000
Residual	50	0.2496	0.0050		
Total	51	1.1922			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	0.3021	0.0544	5.5500	0.0000	0.1927	0.4114	0.1927	0.4114
X Variable 1	0.6114	0.0445	13.7417	0.0000	0.5221	0.7008	0.5221	0.7008
Percentage drop comapred to Clarke's results*	22%							

X Variable (X_t): (ln(Real Revenue_t) - 0.57 ln(Real Revenue_t-1))

Dependent Variable (Y_t): (ln(Real Expenses_t) - 0.57 ln(Real Expenses_t-1))

Note: the lag multiplier (0.57) is based on Maddala, p. 92

Lag multiplier = .5 (2.0 - dw from simple model), where dw from the simple model = 0.86).

* See Clarke's Appendix U-1 - May 7, 2010.xls

IX. APPENDIX 3 – INFORMATION CONSIDERED

52. *ORACLE USA, INC., a Colorado corporation, ORACLE INTERNATIONAL CORPORATION, a California corporation, ORACLE EMEA LIMITED, an Irish private limited company, and SIEBEL SYSTEMS INC., a Delaware corporation ,Plaintiffs ,v. SAP AG, a German corporation, SAP AMERICA, INC., a Delaware corporation, TOMORROWNOW, INC., a Texas corporation, and DOES 1-50, inclusive, Defendants, Fourth Amended Complaint for Damages and Injunctive Relief.* In United States District Court, Northern District of California, San Francisco Division, Filed August 19, 2009.

53. Expert Report of Stephen Clarke, dated May 7, 2010, pages 243-246, 271-273, 276-281.

54. Clarke's *Appendix U-1 – May 7, 2010.x; Appendix U-2 – May 7, 2010.xls; Appendix U-3 - May 7, 2010.xls; Appendix M-1 – M-8 – May 7, 2010.xls; Appendix M-9 – May 9, 2010.pdf*

55. Deposition of Stephen Clarke, June 10, 2010, pages 921-967.

56. Davidson, R. and J.G. MacKinnon, Estimation and Inference in Econometrics. New York, New York: Oxford University Press. 1993.

57. Kennedy, Peter, A Guide to Econometrics. Sixth Edition. Malden, Massachusetts: Blackwell Publishing. 2008.

58. Maddala, G.S., Econometrics. New York, New York: McGraw-Hill Book Company. 1977.

59. Maddala G.S. and K. Lahiri, Introduction to Econometrics. West Sussex, England: John Wiley & Sons Ltd. 2009.

60. Macfie, B.P. and P.M. Nufrio, Applied Statistics for Public Policy. Armonk, New York: M.E. Sharpe, Inc. 2006.

61. Rice, John A., Mathematical Statistics and Data Analysis, Second Edition, Belmont, California: Duxbury Press, 1995

62. Snedecor, G.W. and W.G. Cochran, Statistical Methods. Sixth Edition. Ames, Iowa: The Iowa State University Press. 1967.

63. Wooldridge, J.M., Econometric Analysis of Cross Section and Panel Data. Cambridge, Massachusetts: The MIT Press. 2002.

64. Wooldridge, J.M., Introductory Econometrics. Mason, Ohio: South-Western Cengage Learning. 2009.

I declare under penalty of perjury under the laws of the United States that the foregoing is true and correct and that this declaration is executed on August 19, 2010 at Boston, Massachusetts.

Daniel S. Levy, Ph.D.