IN THE SUPREME COURT
OF THE STATE OF WASHINGTON

| | |
|---|---|
| STATE OF WASHINGTON,<br><br>Respondent,<br><br>v.<br><br>ALLEN GREGORY,<br><br>Appellant. | NO. 88086-7<br><br>DR. SCURICH'S ANSWERS TO THE COMMISSIONER'S INTERROGATORIES |

Office of Prosecuting Attorney
930 Tacoma Avenue South, Room 946
Tacoma, Washington 98402-2171
Main Office: (253) 798-7400

# RESPONSE TO COMMISSIONER'S INTERROGATORIES

## STATE OF WASHINGTON v. ALLEN EUGENE GREGORY NO. 88086-7

Two provisos should be noted from the outset. First, I responded only to the interrogatories directed towards me. Not responding to other aspects of the Commissioner's Interrogatories should not be interpreted as agreement. Second, I will use the terms *Updated Report*, the *Evaluation of the Updated Report*, and the *Response to Evaluation* in a manner consistent with the Commissioner's Interrogatories described on page 3, footnote 1.

**Interrogatory No. 1 (Directed to Dr. Scurich): Are you aware of cases that are in fact missing from the trial reports, or should your statement be taken as a caveat that you have not independently verified the inclusiveness of the trial reports?**

The latter is correct.

**Interrogatory No. 3 (Directed to Dr. Scurich): Please indicate if you maintain that there is evidence of an error in the data analysis that treats a black defendant as a white defendant and, if yes, specify the nature and location of that evidence.**

I am not aware of such evidence at this time. That statement was intended simply to illustrate how easily and unwittingly a coding error could occur. When I stated "[t]here is evidence that this actually occurred…" it would have been more appropriate to state "[t]here is evidence that this *type of coding error* actually occurred…" *Evaluation of the Updated Report* at 6, footnote 3.

**Interrogatory No. 9 (Directed to Dr. Scurich): To the extent that you were unable to verify the numbers in the last two columns in Table I because of uncertainty about the descriptions, are you able to verify those numbers with the provided clarification? If no, please explain.**

With regard to the third column in Table I., I now understand that "average number of victims" actually means "average number of victims *per case*."

Table I. displays the average number of victims per case decomposed by county.

Not all of the values reported in Table I. are correct. Specifically, Table I. asserts that Pierce County had an average of 3 victims per case. This is incorrect. I have pasted below all of the trial report numbers (according to Professor Becket's variable Trial_Report_Num) from Pierce County and the number of victims per case (according to Professor Becket's variable Vics_Num).

| | Trial_Report_Num | Vics_Num |
|---|---|---|
| 1. | 3 | 1 |
| 2. | 29 | 1 |
| 3. | 34 | 1 |
| 4. | 42 | 2 |
| 5. | 48 | 1 |
| 6. | 62 | 1 |
| 7. | 63 | 1 |
| 8. | 64 | 1 |
| 9. | 65 | 1 |
| 10. | 74 | 1 |
| 11. | 75 | 2 |
| 12. | 84 | 2 |
| 13. | 85A | 1 |
| 14. | 87 | 1 |
| 15. | 95 | 2 |
| 16. | 105 | 2 |
| 17. | 123 | 1 |
| 18. | 130 | 3 |
| 19. | 135 | 1 |
| 20. | 157 | 3 |
| 21. | 166 | 1 |
| 22. | 180 | 1 |
| 23. | 181 | 1 |
| 24. | 182 | 2 |
| 25. | 184 | 1 |
| 26. | 186 | 2 |
| 27. | 190 | 2 |
| 28. | 193 | 1 |
| 29. | 194 | 1 |
| 30. | 204 | 1 |
| | | 2 |

| | | |
|---|---|---|
| 31. | 207 | 1 |
| 32. | 211 | 2 |
| 33. | 212 | 1 |
| 34. | 216 | 1 |
| 35. | 233 | 1 |
| 36. | 240 | 5 |
| 37. | 241 | 1 |
| 38. | 242 | 1 |
| 39. | 243 | 1 |
| 40. | 244 | 1 |
| 41. | 247 | 5 |
| 42. | 248 | 1 |
| 43. | 251 | 2 |
| 44. | 263 | 2 |
| 45. | 269 | 1 |
| 46. | 276 | 2 |
| 47. | 280 | 3 |
| 48. | 281 | 1 |
| 49. | 296 | 2 |
| 50. | 297 | 2 |
| 51. | 302 | 2 |
| 52. | 306 | 1 |
| 53. | 312 | 1 |
| | | Total = 82 |

As is apparent, there are 53 trials and the total number of victims is 82. 82 divided by 53 is 1.547, which rounds to 2, not 3.

With regard to the fourth column in Table I., the “Average Number of Affirmed Aggravators,” I remain uncertain as to what variable Professor Beckett used to calculate the averages.

In the *Response to Evaluation*, Professor Beckett states “we consider the number of aggravated circumstances found by the *jury* to be applicable to affirmed aggravators (at 15, footnote 34; emphasis added).” However, the codebook defines the variable AppliedAggCir_Num as the “Number of aggravating circumstances found by the *judge* to have been applicable in this Case (codebook at 45; emphasis added).” I could not find a variable in the codebook that provides “the number of aggravated circumstances found by the jury,” and therefore cannot verify the values in the fourth column of Table I.

**Interrogatory No. 12 (Directed to Dr. Scurich): If it may be assumed that Professor Beckett mistakenly added the "unknown" case to the numerator but not to the denominator in the "death notice filed" column of Table 2, does this result in a difference to the percentage calculation?**

Yes, it would affect the calculated percentages, though with rounding the end result would be the same. If the appropriate numbers are 85/296, the resulting percentage is 28.7%, which rounds to 29%.

**Interrogatory No. 16 (Directed to Dr. Scurich): Using the values identified in the Table 3 note and accompanying narrative, are you now able to replicate the results in the "Death Notice Filed" and "Death Penalty Imposed" columns of Table 3? If not, please explain.**

Yes, following the 65-word description in the Appendix of *Response to Evaluation* at 61 of how the values were originally computed using several different variables (since the germane variables *per se* were not apparently in the codebook or datafile), I was then able to replicate the values in the "Death Notice Filed" and "Death Penalty Imposed" columns of Table 3.

**Interrogatory No. 18 (Directed to Dr. Scurich): In this rerun of the model with a resulting p-value of .062, did you in fact rerun the model without correcting the coding errors? If no, please explain. If yes, please indicate whether and what meaningful information is provided by this rerun of the model, given that Professor Beckett has acknowledged the coding errors.**

The model reported in section 2.3 that yielded a p-value of 0.62 removed the following redundant cases: Gregory ID# 216; Rupe ID# 7; Davic/Davis ID#180 (*Evaluation of the Updated Report* at 25).

I later learned of three coding errors (Spillman ID# 167; Benn ID# 75; Pirtle ID# 132 (*Evaluation of the Updated Report* at 26)). I corrected these coding errors, then re-ran the model reported in section 2.3, and reported the results in section 2.4 (*Evaluation of the Updated Report* at 26). Thus, the redundant cases identified in section 2.3 were removed for the analysis conducted in section 2.4.

Professor Beckett did acknowledge the coding errors, but only after *Evaluation of the Updated Report* identified the coding errors in the first place. At the time the analysis in section 2.4 was conducted it was unknown what effect the coding errors would have on the results.

**Interrogatory No. 19 (Directed to Professor Beckett and to Dr. Scurich): What p-value results if the model reported in Table 7 of the Updated Report is run with the three coding errors corrected, the first sentencing proceedings reported in Trial Reports 7, 180 and 216 removed, and with the logarithmic transformations of variables as set forth in the Response to Evaluation?**

There is an outstanding issue that must be addressed before this question can be answered. Professor Beckett never specifies what constant she used to replace values of 0 when conducting the logarithmic transformations, and a logarithmic transformation of 0 is undefined. The *Response to Evaluation* at 29 suggests that values of 0 were converted to 0.001 and then logarithmically transformed:

> To avoid this, it is common practice to transform all values of the variable by adding a very small number (such as .001) before applying the logarithmic transformation.

However, the value that was actually used in her analyses is never defined at any point in any of the documents produced by Professor Beckett.

I note that the use of 0.001 is arbitrary. The source (a tutorial website) cited in Footnote 58 of *Response to Evaluation* at 29 to support the use of 0.001 in place of 0 for the purpose of conducting a logarithmic transformation does not at any point state that 0.001 should be used in place of 0, much less that the use of this value is "common practice."

Assuming *arguendo* that values of 0 are to be replaced by 0.001 prior to the logarithmic transformation, the model on which Table 7 is based was re-ran, consistent with the specifications delineated in Interrogatory 19. The p-value associated with the race of the defendant is 0.072. See the unaltered output below.

**Variables in the Equation**

| | | B | S.E. | Wald | df | Sig. | Exp(B) | 95% C.I.for EXP(B) | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | Lower | Upper |
| Step 1[a] | LNprior | -.053 | .109 | .237 | 1 | .627 | .948 | .766 | 1.174 |
| | Vics_1Total(1) | -.668 | .590 | 1.284 | 1 | .257 | .513 | .161 | 1.629 |
| | AppliedAggCir_Num | .600 | .262 | 5.243 | 1 | .022 | 1.823 | 1.090 | 3.047 |
| | LNmitigating | -.206 | .117 | 3.105 | 1 | .078 | .814 | .647 | 1.023 |
| | Defenses_Num | -.749 | .377 | 3.939 | 1 | .047 | .473 | .226 | .991 |
| | Vics_AnyHostage(1) | .787 | .593 | 1.763 | 1 | .184 | 2.197 | .687 | 7.023 |
| | D_RaceB(1) | 1.415 | .788 | 3.227 | 1 | .072 | 4.115 | .879 | 19.267 |
| | Constant | -1.266 | .789 | 2.572 | 1 | .109 | .282 | | |

a. Variable(s) entered on step 1: LNprior, Vics_1Total, AppliedAggCir_Num, LNmitigating, Defenses_Num, Vics_AnyHostage, D_RaceB.

**Interrogatory No. 21 (Directed to Dr. Scurich): Were only 55 cases included in the analysis when you re-ran the model that appears in Table 7 of the Updated Report, using a logarithmic transformation of prior convictions and number of mitigating circumstances? Is Professor Beckett correct that the inadvertent omission of other cases accounts for your inability to replicate the effect for black defendants when you re-ran the model using logarithmic transformations?**

55 cases were included in the model (*Evaluation of the Updated Report* at 57). As discussed in response to Interrogatory 19, even if it were clear that the variables for prior convictions and number of mitigating circumstances were to be logarithmically transformed, the *Updated Report* never at any point specified what value was used in place of 0 before conducting the transformations. Indeed, the *Updated Report* never discloses that values of 0 were replaced by a constant at all. Nonetheless, the difference in the number of cases included in the model could account for the non-replication.

**Interrogatory No. 22 (Directed to Dr. Scurich): Is Professor Beckett correct that you were unable to replicate the results because you did not transform the variables? If no, please explain.**

The models relevant to Tables 4-6 did not contain logarithmically transformed variables (*Evaluation of the Updated Report* at 36-54). Again, even if it were clear in *Updated Report* that some variables were logarithmically transformed, and I explained in detail why it was not clear (*Evaluation of the Updated Report* at 20), I still could not replicate the analyses conducted by Professor Beckett because she has never articulated what constant she used in place of 0 when conducting the logarithmic transformations.

**Interrogatory No. 23 (Directed to Dr. Scurich and to Professor Beckett). Do you agree with the above general description of MLEs? If not, please indicate what corrections you would make in the description.**

I agree with the description in the quoted material. However, I do not understand what the term “chance variation” refers to with respect to maximum likelihood estimation. Neither the *Reference Manual on Scientific Evidence* nor the Agresti and Finlay text explain the concept in this context.

**Interrogatory No. 24 (Directed to Professor Beckett and Dr. Scurich): Does identifying and removing a data point that is an outlier address the extent to which chance variation accounts for the distribution of the remaining data for purposes of MLEs? If yes, please explain.**

I do not understand the question because I do not know what the term “chance variation” refers to in this context.

**Interrogatory No. 25 (directed to Professor Beckett and Dr. Scurich): Does testing for robustness address the degree to which chance variation accounts for the distribution of the data for purposes of MLEs? If yes, please explain.**

Again, I do not understand the question because I do not know what the term “chance variation” refers to in this context.

**Interrogatory No. 29 (Directed to Dr. Scurich): Is Professor Beckett correct in her assessment that this model variant (categorizing the race of the defendant into white, black, or other) did not use logarithmic transformations of the variables for prior convictions and mitigating circumstances? If yes, do you agree that a model**

**using appropriate transformations results in the statistical output shown in Table C6 in the Response to Evaluation at 80?**

The model reported at 61-64 of *Evaluation of the Updated Report* did not contain logarithmic transformations of the variables for prior convictions and mitigating circumstances.

Assuming the logarithmic transformations are to use .001 in place of 0, I was able to approximately replicate Table6C in *Response to Evaluation* at 80. The unaltered output appears below; the relevant values to compare are in the "B" column of this output with the "Coefficient" column of Table6C. One important discrepancy concerns the p-value for D_RaceB, which is .054 in the output below but .045 in Table6C. This discrepancy is particularly perplexing, given that most of the coefficient values ("B" in my output, "Coefficient" in Table6C) in both analyses are so similar.

**Variables in the Equation**

| | | B | S.E. | Wald | df | Sig. | Exp(B) | 95% C.I.for EXP(B) | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | Lower | Upper |
| Step 1[a] | LNprior | -.091 | .103 | .792 | 1 | .373 | .913 | .746 | 1.116 |
| | Victim1_vs_mult(1) | -.726 | .591 | 1.510 | 1 | .219 | .484 | .152 | 1.540 |
| | AppliedAggCir_Num | .642 | .266 | 5.808 | 1 | .016 | 1.900 | 1.127 | 3.202 |
| | LNTotMitigating | -.222 | .115 | 3.723 | 1 | .054 | .801 | .639 | 1.004 |
| | Defenses_Num | -.801 | .383 | 4.382 | 1 | .036 | .449 | .212 | .950 |
| | Vics_AnyHostage(1) | .739 | .599 | 1.521 | 1 | .217 | 2.094 | .647 | 6.775 |
| | D_RaceB(1) | 1.516 | .786 | 3.718 | 1 | .054 | 4.555 | .975 | 21.276 |
| | D_RaceNotBW(1) | -.178 | .903 | .039 | 1 | .844 | .837 | .143 | 4.915 |
| | Constant | -1.200 | .831 | 2.083 | 1 | .149 | .301 | | |

a. Variable(s) entered on step 1: LNprior, Victim1_vs_mult, AppliedAggCir_Num, LNTotMitigating, Defenses_Num, Vics_AnyHostage, D_RaceB, D_RaceNotBW.

Bear in mind that TableC6 contains three redundant cases (i.e., Trial Reports 7, 180, 216) as described in Interrogatory No. 19. When these three cases are removed, the p-value for D_RaceB is 0.067. The unaltered output appears below.

**Variables in the Equation**

| | | B | S.E. | Wald | df | Sig. | Exp(B) | 95% C.I.for EXP(B) | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | Lower | Upper |
| Step 1[a] | LNprior | -.048 | .109 | .192 | 1 | .661 | .953 | .769 | 1.181 |
| | Vics_1Total(1) | -.645 | .594 | 1.181 | 1 | .277 | .525 | .164 | 1.680 |
| | AppliedAggCir_Num | .616 | .267 | 5.307 | 1 | .021 | 1.852 | 1.096 | 3.127 |
| | LNmitigating | -.207 | .118 | 3.087 | 1 | .079 | .813 | .645 | 1.024 |
| | Defenses_Num | -.729 | .381 | 3.659 | 1 | .056 | .482 | .229 | 1.018 |
| | Vics_AnyHostage(1) | .808 | .597 | 1.835 | 1 | .176 | 2.243 | .697 | 7.223 |
| | D_RaceB(1) | 1.470 | .804 | 3.346 | 1 | .067 | 4.349 | .900 | 21.015 |
| | D_RaceNotBW(1) | -.318 | .833 | .146 | 1 | .702 | .727 | .142 | 3.725 |
| | Constant | -1.073 | .934 | 1.322 | 1 | .250 | .342 | | |

a. Variable(s) entered on step 1: LNprior, Vics_1Total, AppliedAggCir_Num, LNmitigating, Defenses_Num, Vics_AnyHostage, D_RaceB, D_RaceNotBW.

However, there is a major technical problem with using the two "dummy variables" D_RaceB and D_RaceNotBW (*Response to Evaluation* at 79) to test whether black defendants are more likely than white or other race defendants individually to receive a death sentence.

The correlation of D_RaceB and D_RaceNotBW is -.221. Unaltered output appears below.

**Correlations**

| | | D_RaceNotBW | D_RaceB |
|---|---|---|---|
| D_RaceNotBW | Pearson Correlation | 1 | -.221** |
| | Sig. (2-tailed) | | .000 |
| | N | 293 | 293 |
| D_RaceB | Pearson Correlation | -.221** | 1 |
| | Sig. (2-tailed) | .000 | |
| | N | 293 | 293 |

**. Correlation is significant at the 0.01 level (2-tailed).

Consider the following passage from the peer-reviewed article *Primer on Multiple Regression Coding: Common Forms and Additional Case of Repeated Contrasts* in the journal *Understanding Statistics*:

> However, many textbooks also seem to imply that the parameter estimates associated with each of the vectors are direct tests of the a priori contrast of interest. That is, the implication is that if the vectors are used as simultaneous predictors of the dependent variable, then the regression weights and their associated significance tests are direct tests of contrasts that the researcher specified. In truth, though, the parameter estimates associated with each contrast coded vectors are direct tests of the a priori contrasts in only a limited set of circumstances; *only when the vectors are orthogonal—that is, when the codes are uncorrelated—will the desired contrasts be directly tested*. Under all other cases, the tests of the individual parameters are partially confounded by the mere inclusion of nonorthogonal vectors and hypotheses. This occurs primarily because these contrasts are partially redundant (which is why their codes are correlated), and the MR [multiple regression] procedure by design produces parameter estimates and hence, significance tests that reflect each vector's unique contribution. (at 52.; emphasis added)

As noted above, D_RaceB and D_RaceNotBW are correlated (i.e., nonorthogonal). Therefore, they do not actually test the contrasts they purport to test, as the passage quoted directly above makes clear.

An alternative approach is to use Professor Beckett's variable D_RaceOrdinal, which the *Codebook* defines at 17 as "Defendant's race – 3 categories (1=White or Caucasian; 2=Black or African American; 3=Other race)." Because D_RaceOrdinal is a single variable with multiple, mutually exclusive categories, there can be no correlation among the "dummy variables." Also, given the concerns described in the Commissioner's Interrogatories at 29-30 about "parsimonious models" (i.e., models with the fewest variables), the use of a single variable with multiple categories, as opposed to multiple variables, is appropriate.

I used D_RaceOrdinal in my original model (*Evaluation of the Updated Report* at 23), albeit without the logarithmic transformations of the variables for prior convictions and mitigating circumstances. I re-ran the model reported at 23 of *Evaluation of the Updated Report* with both variables

logarithmically transformed and the three coding errors discussed in Interrogatory 19 corrected as well as the three duplicate cases discussed in Interrogatory 19 removed. The unaltered output appears below.

**Variables in the Equation**

| | | B | S.E. | Wald | df | Sig. | Exp(B) | 95% C.I.for EXP(B) | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | Lower | Upper |
| Step 1[a] | LNprior | -.054 | .110 | .240 | 1 | .624 | .947 | .763 | 1.176 |
| | Vics_1Total(1) | -.670 | .590 | 1.287 | 1 | .257 | .512 | .161 | 1.628 |
| | AppliedAggCir_Num | .599 | .263 | 5.181 | 1 | .023 | 1.820 | 1.087 | 3.047 |
| | LNmitigating | -.206 | .117 | 3.113 | 1 | .078 | .814 | .648 | 1.023 |
| | Defenses_Num | -.751 | .380 | 3.915 | 1 | .048 | .472 | .224 | .993 |
| | Vics_AnyHostage(1) | .781 | .601 | 1.690 | 1 | .194 | 2.184 | .673 | 7.087 |
| | D_RaceOrdinal | | | 3.232 | 2 | .199 | | | |
| | D_RaceOrdinal(1) | 1.407 | .796 | 3.121 | 1 | .077 | 4.083 | .857 | 19.452 |
| | D_RaceOrdinal(2) | -.058 | .901 | .004 | 1 | .949 | .944 | .161 | 5.522 |
| | Constant | -1.249 | .832 | 2.256 | 1 | .133 | .287 | | |

a. Variable(s) entered on step 1: LNprior, Vics_1Total, AppliedAggCir_Num, LNmitigating, Defenses_Num, Vics_AnyHostage, D_RaceOrdinal.

Consistent with my previous finding, neither the main effect for D_RaceOrdinal nor any of the contrasts are statistically significant (p-values .199, .077, .949; all 95%CIs for Exp(B) include the value 1. See response to Interrogatory 33 below).

Bear in mind that D_RaceOrdinal is a variable that was created by Professor Beckett. The *Response to Evaluation (e.g.,* at 33 or 79) never states that this variable is inappropriate for testing whether black defendants are more likely than white or other race defendants individually to receive a death sentence.

**Interrogatory No. 30 (Directed to Dr. Scurich): Do you maintain that your model variant using the DefRaceXVicRace variable demonstrates that there are no racial effects for the defendant with respect to the imposition of the death penalty? If yes, please explain the theoretical basis for a model that includes a DefRaceXVicRace variable if we may assume that consideration of race in assessing**

**whether the defendant should receive the death penalty is an illegitimate factor-whether that consideration is of the race of the defendant alone, the race of the victim alone, or the race of the defendant in combination with the race of the victim.**

The non-significance of DefRaceXVicRace implies that defendant race as well as victim race is not related to death sentences in the model presented on page 24 of *Evaluation of the Updated Report*. As noted below in response to Interrogatory 31, this finding holds even after the technical issues noted in Interrogatory 19 are corrected.

Regarding the theoretical basis of the variable DefRaceXVicRace, note that Table 3 of *Updated Report* at 22 disaggregates the data by both race of defendant and race of victim concurrently, and the associated text refers to different "racial configurations" of defendant and victim race (*Updated Report* at 22). These "configurations" refer to an interaction in statistical parlance. This led me to infer that an interaction of the race of the defendant and the race of victim should be accounted for in the model.

There are two options if one wants to model such an interaction. One is to include a main effect variable for race of defendant, a main effect variable for race of victim, and an interaction variable in the model. The second approach is to create a single variable with all possible combinations of defendant and victim race.

I utilized the second approach, which is defensible for at least two reasons. First, it requires only a single variable as opposed to three variables. The use of fewer variables comports with the notion of "parsimonious models" that the Commissioner's Interrogatories discusses at 29. Second, the approach facilitates interpretation, since it readily allows a direct comparison of certain combinations of defendant and victim race.

**Interrogatory No. 31 (Directed to Dr. Scurich): Is Professor Beckett correct in her assessment that this model variant did not use logarithmic transformations of the variables for prior convictions and mitigating circumstances and included only 60 sentencing proceedings?**

The model reported on page 24 of *Evaluation of the Updated Report* is based on 60 cases and did not contain logarithmic transformations. However, if the issues noted in Interrogatory 19 are corrected, DefRaceXVicRace is still not

statistically significant (p-values = .164, .660, .536; all 95%CIs for Exp(B) include the value 1.). See the unaltered output below.

**Variables in the Equation**

| | | B | S.E. | Wald | df | Sig. | Exp(B) | 95% C.I.for EXP(B) | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | Lower | Upper |
| Step 1[a] | LNprior | -.070 | .129 | .292 | 1 | .589 | .933 | .724 | 1.202 |
| | Vics_1Total(1) | -.580 | .704 | .680 | 1 | .410 | .560 | .141 | 2.223 |
| | AppliedAggCir_Num | .623 | .286 | 4.738 | 1 | .030 | 1.864 | 1.064 | 3.267 |
| | LNmitigating | -.206 | .133 | 2.415 | 1 | .120 | .813 | .627 | 1.055 |
| | Defenses_Num | -.638 | .402 | 2.513 | 1 | .113 | .528 | .240 | 1.163 |
| | Vics_AnyHostage(1) | .251 | .646 | .151 | 1 | .697 | 1.286 | .363 | 4.558 |
| | DefRaceXVicRace | | | 3.613 | 2 | .164 | | | |
| | DefRaceXVicRace(1) | -.677 | 1.539 | .193 | 1 | .660 | .508 | .025 | 10.387 |
| | DefRaceXVicRace(2) | 1.062 | 1.717 | .383 | 1 | .536 | 2.893 | .100 | 83.715 |
| | Constant | -.693 | 1.658 | .175 | 1 | .676 | .500 | | |

a. Variable(s) entered on step 1: LNprior, Vics_1Total, AppliedAggCir_Num, LNmitigating, Defenses_Num, Vics_AnyHostage, DefRaceXVicRace.

If one follows the first approach described above of modeling the defendant race, victim race, and the interaction (i.e., two main effect variables and an interaction term), the same result obtains: neither the race of the defendant, the race of the victim, nor the interaction are significant (p-values = .977, .295, .327, respectively; all 95%CIs for Exp(B) include the value 1.). See the unaltered output below.

**Variables in the Equation**

| | | B | S.E. | Wald | df | Sig. | Exp(B) | 95% C.I.for EXP(B) | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | Lower | Upper |
| Step 1[a] | LNprior | -.036 | .111 | .103 | 1 | .748 | .965 | .776 | 1.200 |
| | Vics_1Total(1) | -.654 | .609 | 1.153 | 1 | .283 | .520 | .158 | 1.715 |
| | AppliedAggCir_Num | .652 | .274 | 5.662 | 1 | .017 | 1.919 | 1.122 | 3.283 |
| | LNmitigating | -.249 | .130 | 3.639 | 1 | .056 | .780 | .604 | 1.007 |
| | Defenses_Num | -.708 | .381 | 3.457 | 1 | .063 | .493 | .234 | 1.039 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Vics_AnyHostage(1) | .853 | .608 | 1.972 | 1 | .160 | 2.348 | .713 | 7.724 |
| | D_RaceB(1) | -.047 | 1.605 | .001 | 1 | .977 | .954 | .041 | 22.157 |
| | Vics_RaceW(1) | -.902 | .862 | 1.095 | 1 | .295 | .406 | .075 | 2.197 |
| | D_RaceB(1) by Vics_RaceW(1) | 1.857 | 1.895 | .961 | 1 | .327 | 6.407 | .156 | 262.813 |
| | Constant | -.697 | .968 | .519 | 1 | .471 | .498 | | |

a. Variable(s) entered on step 1: LNprior, Vics_1Total, AppliedAggCir_Num, LNmitigating, Defenses_Num, Vics_AnyHostage, D_RaceB, Vics_RaceW, D_RaceB * Vics_RaceW .

Note that the analysis above used the variables Vics_RaceW and D_RaceB simply to be consistent with the variables used by Professor Beckett (*Response to Evaluation* at 82).

**Interrogatory No. 3[3] (Directed to Dr. Scurich); Please clarify the object of your statement concerning the interpretation of the confidence intervals that include the value of 1 and indicate whether you agree with Professor Beckett that a correction is needed in your comment about interpreting the confidence intervals.**

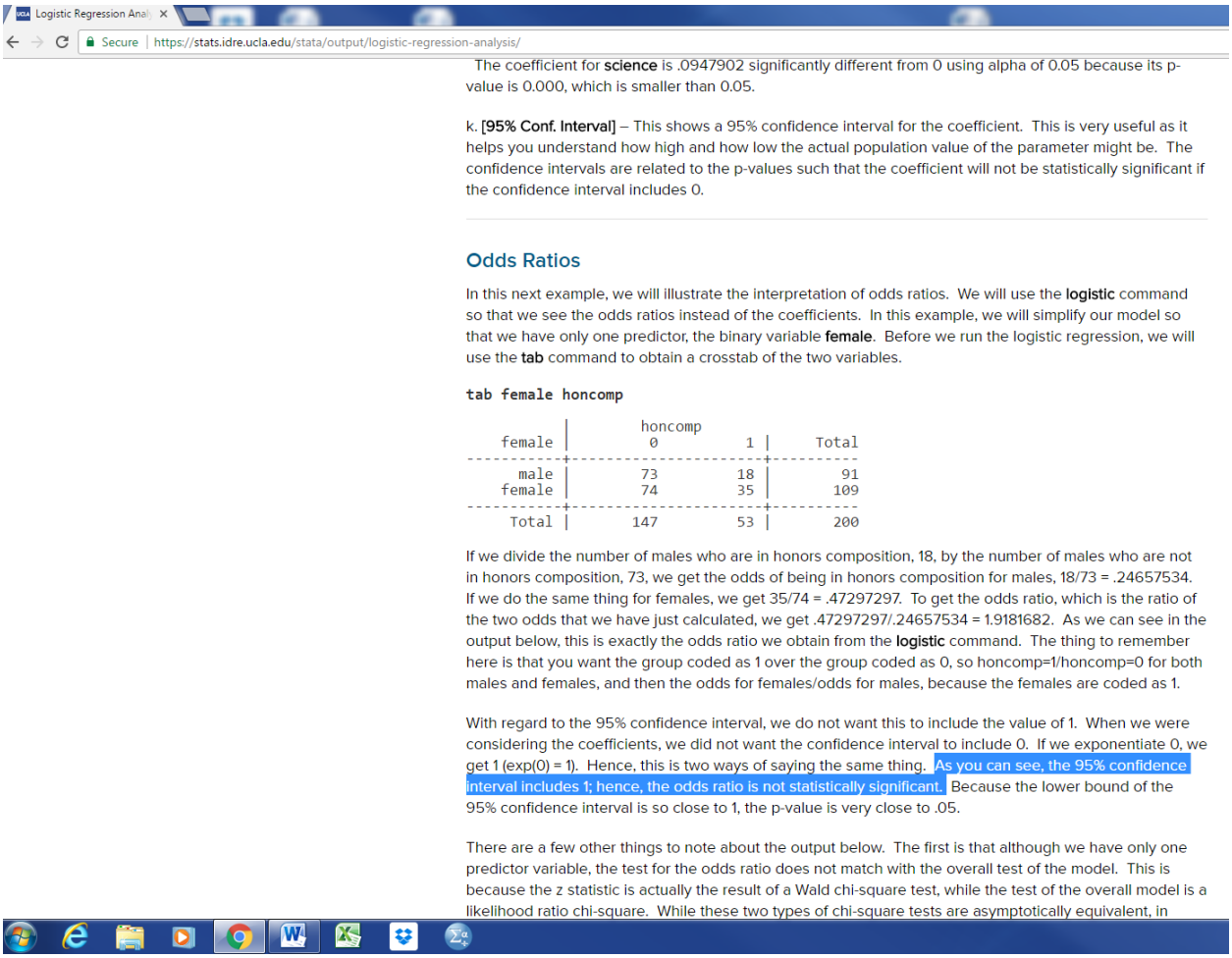
My statement about confidence intervals for Exp(B) containing the value of 1 being interpreted as "not 'significantly' predictive" is accurate, and therefore I disagree with Professor Beckett that a correction is necessary.

Professor Beckett refers to a tutorial website to buttress her contention that confidence intervals containing the value of 0, not 1, should interpreted as not statistically significant (*Response to Evaluation* at 57). The hyperlink in *Response to Evaluation* at 54, footnote 96 is no longer valid; I believe it is now at https://stats.idre.ucla.edu/stata/output/logistic-regression-analysis/. Specifically, a subsection entitled "Parameter Estimates", subsection k from the website reads:

> k. [95% Conf. Interval] – This shows a 95% confidence interval for the coefficient. This is very useful as it helps you understand how high and how low the actual population value of the parameter might be. The confidence intervals are related to the p-values such that the coefficient will not be statistically significant if the confidence interval includes 0. (quoted – with some alterations – in *Response to Evaluation* at 57).

However, I was explicit that the confidence intervals in my analyses referred to Exp(B), which I explained is the "exponentiation of the logarithmic (natural log) beta parameter" or "[i]n short, it is an odds ratio," as the Commissioner's Interrogatory notes at 37.

The website Professor Beckett refers to has another subsection for Odds Ratios directly beneath subsection k. A screenshot appears below, and the relevant portion vindicating my position is highlighted:



Logistic Regression Anal

Secure | https://stats.idre.ucla.edu/stata/output/logistic-regression-analysis/

The coefficient for **science** is .0947902 significantly different from 0 using alpha of 0.05 because its p-value is 0.000, which is smaller than 0.05.

k. **[95% Conf. Interval]** – This shows a 95% confidence interval for the coefficient. This is very useful as it helps you understand how high and how low the actual population value of the parameter might be. The confidence intervals are related to the p-values such that the coefficient will not be statistically significant if the confidence interval includes 0.

## Odds Ratios

In this next example, we will illustrate the interpretation of odds ratios. We will use the **logistic** command so that we see the odds ratios instead of the coefficients. In this example, we will simplify our model so that we have only one predictor, the binary variable **female**. Before we run the logistic regression, we will use the **tab** command to obtain a crosstab of the two variables.

```
tab female honcomp
```

| female | honcomp 0 | 1 | Total |
|---|---|---|---|
| male | 73 | 18 | 91 |
| female | 74 | 35 | 109 |
| Total | 147 | 53 | 200 |

If we divide the number of males who are in honors composition, 18, by the number of males who are not in honors composition, 73, we get the odds of being in honors composition for males, 18/73 = .24657534. If we do the same thing for females, we get 35/74 = .47297297. To get the odds ratio, which is the ratio of the two odds that we have just calculated, we get .47297297/.24657534 = 1.9181682. As we can see in the output below, this is exactly the odds ratio we obtain from the **logistic** command. The thing to remember here is that you want the group coded as 1 over the group coded as 0, so honcomp=1/honcomp=0 for both males and females, and then the odds for females/odds for males, because the females are coded as 1.

With regard to the 95% confidence interval, we do not want this to include the value of 1. When we were considering the coefficients, we did not want the confidence interval to include 0. If we exponentiate 0, we get 1 (exp(0) = 1). Hence, this is two ways of saying the same thing. As you can see, the 95% confidence interval includes 1; hence, the odds ratio is not statistically significant. Because the lower bound of the 95% confidence interval is so close to 1, the p-value is very close to .05.

There are a few other things to note about the output below. The first is that although we have only one predictor variable, the test for the odds ratio does not match with the overall test of the model. This is because the z statistic is actually the result of a Wald chi-square test, while the test of the overall model is a likelihood ratio chi-square. While these two types of chi-square tests are asymptotically equivalent, in

In case the image gets distorted when printed, the highlighted sentence states: "As you can see, the 95% confidence interval includes 1; hence, the odds ratio is not statistically significant."

Professor Beckett is mistaken about the use of confidence intervals with regard to Exp(B). Her own source supports my explanation that the relevant value is 1, not 0.

Do note the output in response to Interrogatory No. 19; and in particular, notice that the 95% Confidence Interval (C.I.) for the Exp(B) associated with defendant race contains the value of 1.

I declare under penalty of perjury that the foregoing is true and correct to the best of my knowledge and belief. Signed on July 10, 2017.

NICHOLAS SCURICH, PH.D.

# PIERCE COUNTY PROSECUTING ATTORNEY

## July 11, 2017 - 1:33 PM

## Transmittal Information

| | |
|---|---|
| **Filed with Court:** | Supreme Court |
| **Appellate Court Case Number:** | 88086-7 |
| **Appellate Court Case Title:** | State of Washington v. Allen Eugene Gregory |
| **Superior Court Case Number:** | 98-1-04967-9 |

**The following documents have been uploaded:**

- 880867_Other_20170711131146SC041439_9747.pdf
  This File Contains:
  Other - Answer to Interrogatories
  *The Original File Name was Scurich Interrogatory and Cover.pdf*

**A copy of the uploaded files will be sent to:**

- aravind@orrick.com
- changro@seattleu.edu
- cstubbs@aclu.org
- jessica.levin@gmail.com
- jneeb@co.pierce.wa.us
- john.wolfe@orrick.com
- kprocto@co.pierce.wa.us
- levinje@seattleu.edu
- lila@washapp.org
- lobsenz@carneylaw.com
- mrshapiro@orrick.com
- nf@neilfoxlaw.com
- robinson@sgb-law.com
- sea_wa_appellatefilings@orrick.com
- talner@aclu-wa.org
- wapofficemail@washapp.org

**Comments:**

---

Sender Name: Heather Johnson - Email: hjohns2@co.pierce.wa.us
**Filing on Behalf of:** Michelle Hyer - Email: PCpatcecf@co.pierce.wa.us (Alternate Email: )

Address:
930 Tacoma Ave S, Rm 946
Tacoma, WA, 98402
Phone: (253) 798-7875

**Note: The Filing Id is 20170711131146SC041439**