Quiz 2 PSY 5013 Name: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

You have until Friday, March 27 at midnight to complete this quiz.

You are to work on the quiz ON YOUR OWN. No help is allowed – other than the usual non-person resources you always have.

Finally, you may put your completed quiz in the Assignment BOX that states QUIZ 2.

How fast can evolution occur in nature? Are evolutionary trajectories unique or predictable? In 1980, a European Union (EU) fly (Drosophila subobscura) was accidentally introduced into North America. In Europe, the fly’s wing size systematically varies with latitude, suggesting an evolutionary adaptation. After allowing two decades for the introduced North American flies to spread over the continent, flies were captured and the hypothesis of speedy evolution was examined by comparing the wing sizes at different latitudes between NA and EU flies.

The data are given below:

continent latitude Wing Size

Females Males

na 35.5 901 797

na 37 896 806

na 38.6 906 812

na 40.7 907 807

na 40.9 898 818

na 42.4 893 809

na 45 913 810

na 46.8 915 819

na 48.8 927 800

na 49.8 924 823

na 50.8 930 814

eu 36.4 905 789

eu 39.3 889 803

eu 41.3 915 812

eu 43.4 930 820

eu 45.5 895 808

eu 47.3 926 815

eu 48.5 944 855

eu 50.4 925 842

eu 52.1 920 819

eu 56.1 934 839

1. (+22 points Total)

Define the variables as follows:

Y = Wing Size

X1 = Latitude

X2 = dummy code for Continent 1 = NA 0 = EU

X3 = dummy code for Sex 1 = M 0 = F

a. (+4) Write the full and restricted models which, in a models comparison framework, would evaluate the null hypothesis that latitude - controlling for continent and sex - has a significant relationship with wing size.

b. (+4) How many degrees of freedom would exist for full and restricted models in part a above?

c. Suppose you were to see the following SAS code in your program editor window:

PROC REG;

MODEL Y = X1 X2 X3;

DEMO: Test X2=0, X3=0;

c1) (+4) In WORDS, what is the hypothesis being tested in the test statement labeled DEMO?

c2) (+4) Write the full and restricted models used to evaluate the DEMO hypothesis.

d. (+4) Write out the expected wing size for a female North American fly captured at a latitude of 45 degrees in terms of the model parameters (we don’t have numerical estimates yet) from the full model in part A above.

2. To evaluate the speedy adaptional hypothesis, we need to evaluate whether or not the rates of wing change as a function of latitude vary between EU and NA flies. We may do this by including an interaction term X4 - where X4 is the interaction between latitude and continent. In your favorite program, run a multiple regression model - with wing size as the DV - that includes the linear effects of sex, continent, latitude, and the latitude by continent interaction.

Please answer the following questions (+3 points each).

2a. Which of the effects modeled has the most influence on wing size and how do you know this?

2b. What is the value of the multiple correlation for this model and what is it's sign or direction of influence?

2c. What is the *F-value* for testing whether or not there are different latitude slopes by continent? What is the companying p-value and squared partial correlation?

2d. Write out the full prediction equation for the model with the estimated parameters in place of the coefficients.

2e. What is the estimated numerical value of Root MSE for this analysis? In words, what is the meaning of this number?

2f. Which observation number has the largest residual? What is the predicted value and observed value associated with this observation?

2g. In words, interpret the coefficient for the interaction term in this model.

2h. What is the numerical value of the t-statistic that would result for the interaction term if *all of the coefficients were standardized coefficients?*

2i. What is the numerical value of E(R) - E(F) = Δfit for the hypothesis that the interaction term does not influence wing size?

2j. What is the numerical value of the expected wing size for a female North American Fly captured at 45 degrees latitude in this sample?

Bonus questions (+2 pts each)

2k. Using parameter estimates from the interaction model fit in for question 2, what is the estimated intercept and latitude slope for female NA flies?

2l. Using parameter estimates from the interaction model fit in for question 2, what is the estimated intercept and latitude slope for female EU flies?

2m. Using parameter estimates from the interaction model fit in for question 2, what is the estimated intercept and latitude slope for male NA flies?

2n. Using parameter estimates from the interaction model fit in for question 2, what is the estimated intercept and latitude slope for male EU flies?

3. (15 points total) Consider a multiple regression in which we have 4 variables: A response variable Y and 3 explanatory variables: x1, x2, and x3.

Use proper notation for all parts of the question. That is, use

r2y1.2 for (squared) partial correlations,

r2y(1.2), for (squared) semi-partial correlations, and

r2y1 for (squared) simple correlations.

R2y.12 for (squared) multiple correlations

a. (+5) Write out the squared multiple correlation between Y and x2 and x3 in terms of a sum of squared simple correlations and squared semi-partial correlations.

b. (+5) Write out the squared partial correlation between Y and x2 controlling for x1 and x3 in terms of the squared semi-partial correlation between Y and x2 controlling for x1 and x3.

c. (+5) Write the squared partial correlation between Y and x3 controlling for x1 and x2 as a function of squared multiple correlations only.

4. (20 points) As usual, here is output in which most everything has been erased. For each blank you can fill in, you get 1 point credit. The SAS Code is given as MODEL Y = x1 x2;

Correlation

Variable x1 x2 y

x1 1.0000 0.3890 \_\_\_\_\_\_

x2 0.3890 1.0000 0.2317

y 0.2565 0.2317 1.0000

Analysis of Variance

Sum of Mean

Source DF Squares Square F Value Pr > F

Model \_\_\_ \_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_ 4.58 0.0126

Error 97 \_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_

Corrected Total \_\_\_ \_\_\_\_\_\_\_\_\_

Root MSE 1.03178 R-Square \_\_\_\_\_\_

Dependent Mean 11.99894 Adj R-Sq \_\_\_\_\_\_

Coeff Var 8.59895

Parameter Estimates

Squared

Parameter Standard Standardized Semi-partial

Variable DF Estimate Error t Value Pr > |t| Estimate Corr Type I

Intercept \_ 12.00578 0.10538 \_\_\_\_\_\_ <.0001 \_\_\_\_\_\_\_ .

x1 \_ 0.22421 \_\_\_\_\_\_\_ 1.86 0.0659 0.19599 \_\_\_\_\_\_\_

x2 \_ \_\_\_\_\_\_ 0.12275 1.48 0.1432 0.15550 \_\_\_\_\_\_\_

Parameter Estimates

Squared

Partial

Variable Corr Type II

Intercept .

x1 0.03445

x2 \_\_\_\_\_\_\_