**V506 (Connect) Final Exam Study Guide**

The final exam will have the same format as the midterm exam, with short answer questions. Some of the questions will be conceptual in nature, others will require that you perform calculations, and finally some will require programming in R. The exam will cover bivariate and multiple regression. As was the case with the midterm, you will be given several days to complete the exam, and will submit a word document with your answers on Canvas. Please remember to show your work wherever relevant and to include any R code and output. You are allowed to use any of the course materials but are not allowed to communicate with anybody other than myself or the TAs or to use external websites. Below is a list of topics that you should be familiar with for the exam:

**Correlation:**

* What does correlation measure?
* How is the correlation coefficient calculated? Know how to calculate this in R.
* What values can the correlation take on? How are these values interpreted?
* What is the role of correlation in OLS regression?

**Regression**:

* What does OLS stand for? How are the OLS coefficients estimated?
* What are the assumptions for OLS regression (review the textbook)
* What are “control variables?” Why do we need to include these variables in an OLS model? What kinds of problems might occur if we omit an important independent variable from our model?
* What is R-squared and adjusted R-squared? How are these calculated and how do we interpret them? Why do we use adjusted R-squared with multiple regression?
* What is the intercept term in OLS regression? How do we interpret this term?
* Know how to generate a scatterplot between two variables and add the regression line
* Know how to interpret coefficients, t-statistics, and p-values from an OLS model.
* Know how to conduct hypothesis tests and construct confidence intervals regarding OLS coefficients
* Know how to interpret dummy variables in OLS, and how to include these variables in an OLS model in R.
* What is the “dummy trap,” why is it a problem, and how do we avoid it?
* Know how to generate predicted values from an OLS model given a set of values for the independent variables.
* Know how to conduct and interpret an F-test
* Know how to estimate and interpret an OLS model in R
* Know how to estimate and interpret an interaction term in R
* Know how to generate an effects plot to show to effects from an interaction term in R
* Know how to conduct and interpret nested F-tests in R