Pol Sci 630: Problem Set 8: Dummy Variables and Interactions (Part 2)

Prepared by: Anh Le (anh.le@duke.edu)

Due Date: Tue, Oct 20, 2015, 10 AM (Beginning of Lab)

1 Interaction (8 points)

a)

Download FDI, tariff, and GDP data from WDI for all countries, year 2010
(`indicator = c("BX.KLT.DINV.CD.WD", "TM.TAX.MRCH.SM.AR.ZS", "NY.GDP.MKTP.CD")`).
Clean the data as usual. Run the following regression, showing a stargazer result
table.

$$\log(FDI) = \beta_0 + \beta_1 tariff + \beta_2 \log(gdp) + \beta_3 tariff \times \log(gdp) \tag{1}$$

b)

Mathematically, what is the marginal effect of tariff on logfdi (i.e. taking partial
derivative with regards to tariff)? (Hint: This would be a function of loggdp)

Plugging in the number, what's the marginal effect of tariff on logfdi, holding
loggdp at its median value? Note: Use `\Sexpr()` to extract coefficients from
the model, do not hand write your calculation.

c)

Using `ggplot2`, plot the marginal effect of tariff on logfdi (y-axis) against differ-
ent values of loggdp (x-axis). (Hint: Create a data frame, in which one variable
is the values of loggdp, the other variable is the corresponding marginal effect
given that value of loggdp. This data frame is the data that makes up your
plot. The plot is just a line.)

d)

With log fdi on the y-axis, tariff on the x-axis, plot the effect of tariff on log fdi
when log gdp is at the 25%, 50%, and 75% percentile. Brownie point if you do
this in ggplot2.

(Hint: The plot should have 3 lines, each according to a value of log gdp. This is the plot you saw in last lab, with confidence interval included)
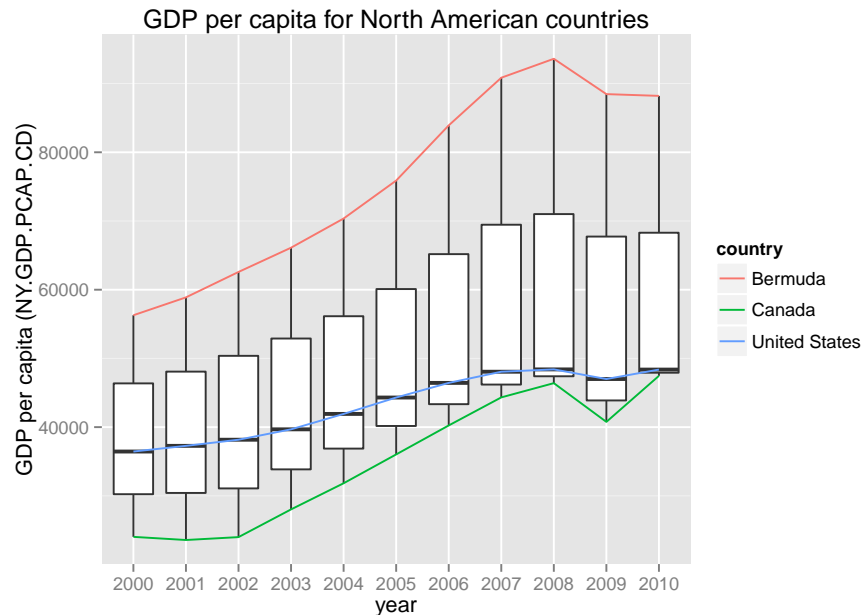
e) Interpretation

Interpret the result (i.e. statistical significance and effect size) using information from both the table and the two plots.

2 ggplot2 (4 points)

Plot this. Note: DO NOT look at the .Rnw code that generates the plot.

If you did try but couldn't figure out, you can look at the code for hints, but then you have to add comments to explain what the code does.

3 ANOVA (4 points)

a)

Load the diamond dataset in R (`data(diamonds)`). With \verbprice' as the dependent variable, run 1) one-way ANOVA on `cut`; 2) two-way ANOVA on `cut` and `clarity` and their interaction.

Interpret the table (i.e. which factor is important in determining the diamond's price?)

b)

What is the expected price for a diamond that has:

- Ideal cut, VS1 clarity

- Very Good cut, VVS1 clarity

Note: You'll have to do dummy regression like last homework, not ANOVA. Also you need to remove the order from the factors `cut` and `clarity`. See this SO answer. Remember to use `\Sexpr{}`, don't write down answers by hand (see how I use `\Sexpr{}` from last solution if confused.)