Econometrics I, FALL SEMESTER 2019-2020

Department of Economics, Rochester Institute of Technology
Instructions for Paper 1: 100 Points
Instructor: Dr. Bríd G. Hanna

- *This paper requires you to carry out linear regression analysis on a data set of your own choosing and to analyze the regression estimates.*
- ***You should not use a dataset that we have used in class nor one that was used in an assignment/an exam.*** *If you do, then you are required to use a regression model that is significantly different from the model analyzed in class / in the assignment. If your paper does not satisfy these requirements, I will reduce your Paper 1 grade by an amount that I deem appropriate.*
- *Write your report on your model and your estimation results in an MS word or PDF document and submit that, together with the R code, and the associated R output, to the appropriate myCourses drop-box by* ***11.59 PM on Friday October 18th, 2019.***
- *It will be helpful to paste your R code and R output into your paper.*
- *Your paper must have the following structure:*

Section 1 (20 points): Describe fully the theory / intuition / logic for the relationship that your paper studies. That is, describe the **causal** relationship(s) that you are examining, describe the form of those relationships (e.g., are they linear or non-linear, do they interact with other relationships?) and provide logic for why it (they) exist(s) in reality and why they exist in that form.

Section 2 (30 points): Describe the linear regression model that you will use to test your logic from section 1 and make sure that the structure / specification of this model is consistent with that logic. For this purpose, do the following:

- State the name of the data file that you are using, if you are using one from the myCourses zip file. Give the source of data if you are using an alternative data source.
- Describe which variables you are using as independent and dependent variables.
- You may have to experiment with different forms of the model. You do not have to describe all of those models but you should briefly explain why you chose your regression model.
- Be specific about the definitions of each of the variables and the units in which they are measured.
- Be sure to point out the **non-linearities** that are built into your econometric model and make sure that these have been motivated in section 1.

- You must experiment with **at least one type of non-linearity** in your model (e.g., a quadratic or an interaction variable).
- Finally, you must explicitly state what section 1 predicts for the signs of each of the parameters in your population regression equation.

R Code (10 points): Next, you must estimate the regression model using R. (These 10 points are for merely using R code to produce regression output and for submitting the code and output.)

Section 3 (30 points): Describe the estimation results for the final version of your regression model. That is,

- Which of the explanatory variable parameter estimates are statistically significant and which are not?
- Then, *precisely* interpret those statistically significant parameter estimates. (*For an example of what is meant by precise interpretation, see the various "R Code & Output" documents on myCourses.*)

Section 4 (10 points): Assess your results by describing how well or how poorly your results support your logic from section 1.

- If your logic is not supported by the regression results, suggest and explain at least one potential problem with the sample data, the regression model, and / or your logic that may explain this inconsistency between your prediction and your result. (For example, you might suspect that you have a non-random sample, or that your regression model exhibits some type of omitted variable bias. Explain these potential problems in detail.)
- If your logic is supported by the regression results, propose and explain one adjustment to your logic or your regression model that may improve its fit with the data but that is currently not feasible to implement due to lack of data, for example.