Quantitative Reasoning for Management

Group Problem Set

Instructions: This problem set is worth 15% of your final course grade, and it is **due on November 23. You must submit your solutions to Quercus through the course page no later than 11:59 p.m. (EDT) on the 23rd.** The problem set has three questions (each with multiple parts) and there are 100 marks in total for all of the questions.

Question 1 (30 marks)

You are working for a major grocery store chain, and your company is trying to improve the productivity of its checkout cashiers. The company believes that cashiers would be able to process more customers in a given period of time if they were exposed to extremely efficient cashiers: they would like these highly-efficient cashiers to be within the direct eyesight of the other cashiers, and they hypothesize that this will inspire a greater level of efficiency for all cashiers. To test out this idea, you convince your company to allow you to randomly assign a highly-efficient cashier in a highly-visible place during certain cashier shifts, but not in others. You also collect data on cashier efficiency before and after you implement this experiment. Your data is contained in the spreadsheet EMBA2025GP1.xls; the variables in this data set are:

(i)   Efficiency: is the number of groceries scanned per minute by a particular cashier (which will be your measure of efficiency).

(ii)   Treatment: a dummy variable equal to one if the worker would have a highly-efficient cashier assigned to their shift, and zero if they wouldn't.

(iii)   Post: a dummy variable equal to one in the period after you implement the experiment, and zero in the period before you implement the experiment.

(iv)   ID: This is a variable that assigns a unique numerical value to each cashier in your data set. You will note that you have two data points on each customer – one data point in the "pre" period, and one in the "post".

This question continues on the following page.

(a)   Use <u>only</u> the data on cashiers that have the highly-efficient cashier assigned to their shift, use a regression to determine how exposure to the highly-efficient cashier changed the efficiency of these cashiers between the "pre" and "post" periods. Discuss why this estimate would or would not show a causal effect of your experiment on efficiency. (5 marks)

 (b)   Run a D-i-D regression that compares the efficiency effects of your experiment for cashiers that did and did not have a highly-efficient cashier assigned to their shift. In this case, use the whole sample to run a simple regression whose right-hand-side variables only include: treatment, post and their product. (Please note that you will need to create a variable in your spreadsheet that represents the product of post and treatment). Assess your results by commenting on what they suggest about the causal impact of your experiment on efficiency. (10 marks)

(c)   Suppose that you find out that workers who didn't have the high-efficiency cashier assigned to their shift heard about this experiment, and the absence of the high-efficiency cashier made them worried about the safety of their jobs. If these worries induced them to work harder than normal, then explain how (if at all) this fact would change the interpretation of your results. (15 marks)

<u>Question 2 (30 marks)</u>

You are working for a major supplier of construction materials, and your company is trying to determine the best pricing strategy for its supplies. Given your training in Economics, you know that output demand is negatively related to price, but your company is large enough that it believes it can potentially alter its prices to raise its overall revenues. To test out this idea, you convince your company to allow you to randomly lower prices for some of your repeat customers, but not for others. And suppose you collect data on the sales revenue you collect from customers before and after you implement this pricing experiment. Your data is contained in the spreadsheet EMBA2025GP2.xls; the variables in this data set are:

(i)   Revenue: is the dollar amount of revenue you collect from a sale to a given customer.

(ii)   Treatment: a dummy variable equal to one if the customer will be offered a different price under your pricing experiment, and zero if the customer will not be offered a different price under the pricing experiment.

(iii)   Post: a dummy variable equal to one in the period after you implement the pricing experiment, and zero in the period before you implement the pricing experiment.

(iv)   ID: This is a variable that assigns a unique numerical value to each customer in your data set. You will note that you have two data points on each customer – one data point in the "pre" period, and one in the "post".

(a)   Use <u>only</u> the data on customers that have been offered the special price under the experiment, use a regression to determine how revenues changed for these customers between the "pre" and "post" periods. Discuss why this estimate would or would not show a causal effect of your pricing experiment on revenues. (5 marks)

(b)   Run a D-i-D regression that compares the revenue effects of your pricing experiment for customers that did and did not get offer the arrangement. In this case, use the whole sample to run a simple regression whose right-hand-side variables only include: treatment, post and their product. (Please note that you will need to create a variable in your spreadsheet that represents the product of post and treatment). Assess your results by commenting on what they suggest about the causal impact of your pricing experiment on revenues. (10 marks)

(c)   Suppose that there is a substantial downturn in the construction industry in the post period that impacts all of your customers in an equivalent way. In this case, how (if at all) would this knowledge change the interpretation of your results? (15 marks)

<u>Question 3 (40 marks)</u>

You are running a cruise ship company, and you are trying to improve your customer service ratings amongst repeat customers. In order to do examine the factors that improve customer service ratings, you decide to explore whether or not access to certain amenities on the impacts these ratings. In particular, you believe that spa treatments and personal training sessions will have a direct effect on customer's experience, so you decide to conduct an experiment to test this idea. In your experiment, you randomly provide these services for free for some repeat customers, while you use standard pricing for these services for other repeat customers. You then collected data survey data on all of these repeat customers that reports their experiences on earlier cruises, and you also collect survey information about their experience on this current cruise (when some were given access to free spa and personal training services). Your data is contained in the spreadsheet EMBA2025gp3.xls; the variables in this data set are:

(i)   Customer Service Score: is the customer's numerical evaluation of their experience on the ship, which is measured on a 10-point scale. For this scale, 1 represents the worst possible experience, and 10 represents the best possible experience.

(ii)   Treatment: a dummy variable equal to one if the customer was selected for your program that provided free services, and zero if the customer was not selected for your program.

(iii)   Post: a dummy variable equal to one if the data was collected in the period when your experiment took place, and zero in the period prior to the experiment.

(iv)   ID: This is a variable that assigns a unique numerical value to each customer in your data set. You will note that you have two data points on each customer – one data point in the "pre" period, and one in the "post".

(v)   Wave Height: This is the maximum height (in feet) of the waves encountered by the ship. Note that higher waves are more likely to rock the ship.

This question will be based on regressions drawn from this spreadsheet, and are listed below.

(a)   Run a D-i-D regression that estimates the effect of your experiment on your customers' views by running a simple regression whose right-hand-side variables only include: treatment, post and their product. Assess your results by commenting on what they suggest about the causal impact of your experiment. (10 marks)

(b)   Now run a D-i-D regression whose right-hand-side variables include: treatment, post, the product of post and treatment, and the average wave height. Assess your results by commenting on what they suggest about the causal impact of your experiment on customer service scores. (20 marks)

(c)   Suppose that after you had collected your data, you also learned that the customers in the treatment group, who were randomly assigned free services, also happened to get larger rooms

that were located on higher floors of the ship (with better views). How, if it all, would this impact the interpretation of your findings? (10 marks)