Foundations of Econometrics - Part II
Final Exam Solutions

Exam$_2$024.*pdf*

December 10, 2024

Question 1

The following table is included in the paper "Do Conditional Cash Transfers Improve Child Health? Evidence from PROGRESA's Control Randomized Experiment" by Paul Gertler, published in the *American Economic Review* in May 2004. The table represents averages of different covariates for treated and control samples. The treatment is providing the conditional cash transfers implemented in the PROGRESA policy, and it was randomized across villages. The outcome is child's height.

Given this information, do you think that comparing treated and control average heights provides a good estimate for the average treatment effect? Do you need to control by any regressors? Why?

TABLE 1—PRE-INTERVENTION DESCRIPTIVE STATISTICS FOR THE MORBIDITY SAMPLE OF CHILDREN AGE 0–35 MONTHS AT BASELINE

Variable	Treatment	Control	*p* value for difference
Child was ill in last 4 weeks (=1)	0.330	0.323	0.771
Age	1.625	1.612	0.914
Male (=1)	0.511	0.491	0.091
Father's years of education	3.803	3.840	0.980
Mother's years of education	3.495	3.829	0.062
Father speaks Spanish (=1)	0.942	0.929	0.276
Mother speaks Spanish (=1)	0.935	0.917	0.443
Own house (=1)	0.923	0.917	0.465
House has electricity (=1)	0.644	0.711	0.091
Hectares of land owned	0.809	0.791	0.553
Male daily wage rate (pesos)	30.483	31.219	0.370
Female daily wage rate (pesos)	27.258	27.844	0.493
Sample size:	4,519	3,306	

Notes: This table reports descriptive statistics for the sample of children age 0–35 months at baseline before the intervention. The *p* values in the third column are for the test of the hypothesis that the means of the treatment and control groups are equal and are adjusted for inter-cluster correlation at the village level.

Answer: We would want to check to ensure that there is no statistically significant difference in the characteristics of the control and treatment. That is, we check the magnitude of the differences and also the p-value. For 95 percent confidence interval, we want a p-value below 0.05. The only variable close to this is Mother's years of education, with a p-value of 0.062, so it is only a significant difference at the 90 percent confidence level. As a result, we don't need to control for any regressor because there isn't a real threat to randomization - we see no systemic differences between treatment and control.

Question 2

Rita Almeida and Marta Faria, in their paper "The wage returns to on-the-job training: evidence from matched employer-employee data" published in the *IZA Journal of Labor and Development* in 2014, are interested in analyzing the effect of on-the-job training on different outcomes such as worker productivity and wages. Whether a worker receives training is observable, but possibly not random.

For this reason, they use propensity score matching. In particular, they estimate their propensity score based on several worker (including education, gender, age, tenure with the firm, potential experience, marital status, occupation) and firm characteristics (including firm size, age, export intensity, foreign ownership, education of the workforce, degree of technological adoption).

Are you convinced by this approach? Why?

Answer: We are not convinced that this approach is the best possible. While their propensity score matching may do a good job handling observable co variates, we think there is a high level of unobserved differences in the groups that may create bias. For example, ability or motivation may relate to both the uptake (whether the worker joins the program) and also the outcome (productivity and wages). As a result, we think these unobservable co-variates would lead to a positive bias. Workers who are highly able or motivated would have potentially experienced better outcomes even without the program, so this model may overstate the impact of the program on wages and productivity.

Question 3

Amazon is interested in estimating the effect of advertisement on purchases and they hire us for that purpose. They give us weekly data on sales and advertisement for that purpose. Our first approach is to classify weeks into high advertisement or not based on the number of ads displayed on that week. We call high advertisement weeks treated and low advertisement weeks controls.

Do you expect the comparison of treated and control weeks to provide a valid causal effect? Why? Explain briefly, connecting the empirical context described and the assumptions discussed in the course.

Answer: We do not expect a casual effect. Firstly, the treatment is not randomly assigned, nor is there a conditional adjustment to handle this lack of randomness. The treatment is simply a classification of high and low volume advertising weeks, but we know factors like holidays and seasonality may influence both ad volume (push more during holidays) as well as sales (people spend more during the holidays). In theory, a half-decent agent like Amazon would not alter between high and low advertising weeks at random - it is highly targeted and specific.

Question 4

Continuing with the Amazon example, we have now access to Amazon's app. Through the app, we can randomly vary the amount of ads that individuals receive displayed on their app, and we observe whether or not individuals watch the ads that are displayed in their apps. The more ads are displayed, the more likely that the individual sees the advertisement at least once (our treatment variable).

We implement two different treatments. In the first one we randomly classify customers in two groups, and we double the number of ads displayed to the group assigned to treatment. In the second one, we display a different (random) amount of ads to each customer.

Briefly explain the different treatment effects that can be identified using these two experiments, how, and why.

Answer: Both of these treatments are valid and return different measurements. They both are random so we can compare the group means.
The first will tell us the treatment effect of strictly doubling advertising, since we are comparing the group with double to the normal group.
The second will be more detailed - it can tell us the marginal effect of advertising. Since advertising level is randomly assigned across some spectrum, we can analyze the effects of different ad levels and compare, instead of just comparing doubled advertising exposure to regular.

Question 5

Building on "Do corporate income tax cuts decrease labor share? Regression discontinuity evidence from China" by Bing Li and co-authors, published in *Journal of Development Economics* in May 2021, we are interested in analyzing how corporate income tax cuts decrease the firms' employment to output ratio.

We want to implement a regression discontinuity design based on group 4 from the table (offshore service outsourcing firms), that is, focusing on a sample of the selected 21 cities, the running variable is the offshore service outsourcing (OSO) revenue divided by total revenue.

Would you implement a sharp or a fuzzy design? Why? What treatment effect will you identify?

Table 1. China's corporate tax rate system, 2010–2013.

Group	Tax rate (%)	Firms	Main qualification requirements
1	25	Normal rate	Firms with no tax credits
2	20	Small- and micro-enterprises tax credit	For service firms: annual taxable income≤300,000, employees≤80, and asset≤100,000,000
3	15	High-technology firms tax credit	R&D intensity≥6% if sales<50,000,000; R&D intensity≥4% if sales≥50,000,000 & sales <200,000,000; R&D intensity≥3% if sales≥ 200,000,000
4	15	Offshore service outsourcing firms	Offshore service outsourcing revenue/total revenue≥ 50% in 21 cities
5	15	Western development tax credit	Located in China's western regions, and main business revenue from government-supported industry/revenue≥70%
6	0 or 12.5	Software and integrated circuit firm tax credit	For start-ups, exemption from tax for two years from the first profit-making year and a preferential tax rate of 12.5% for the subsequent three years
7	Miscellaneous	–	–

Notes: The monetary unit is RMB Yuan. In group 2, the threshold for manufacturing firms is different. In group 3, R&D intensity is defined as the ratio of R&D input to total revenue. There are several special cases not included in groups 1–6, which we label as the miscellaneous group 7.

Answer: We would need a fuzzy design here. A sharp design requires a strict jump in treatment from 0 to 1 around a cutoff, notated as:

$$P(D_i = 1|Z_i = Z) = \begin{cases} 0 & \text{if } Z < z_0 \\ 1 & \text{if } Z \geq z_0 \end{cases}$$

The running variable in this case is the offshore service outsourcing, so the treatment (corporate tax rate) should jump sharply when the OSO hits the cutoff.

However, from the table, we can see there are other categories in which a firm could be applied the same tax rate (15 percent), even if they have yet to hit the OSO cutoff. That is, a firm may not strictly jump from 0 to 1 around the OSO cutoff, it could earlier for example. This means it would be best to use a fuzzy design.

Question 6

Building on the 2012 *American Economic Review* paper by Petra Moser and Alessandra Voena entitled "Compulsory Licensing: Evidence from the Trading with the Enemy Act", we analyze the effect of compulsory licensing on domestic invention. During World War I, the US enacted TWEA, an act that allowed US firms to violate enemy-owned patents if they contributed to war effort.

To measure the effects of licensing, they compare, before and after, the number of patents by domestic inventors across subclasses of (organic chemistry) technologies that were differentially affected by TWEA. Chemical inventions in all these subclasses, however, were affected by increase in tariff barriers.

Does a difference-in-differences approach (exposed to TWEA vs not) provide a consistent estimate in this case? Why? If the UK was exposed to the same tariff increase but did not implement TWEA, can you suggest a method to obtain a valid causal effect?

Answer: The difference-in-difference estimate is not consistent here because tariffs effect firms that are both and not exposed to TWEA.
The core assumption of difference in difference is the parallel trends assumption: without treatment, the trend of the treatment group would have followed the same of the control group.
The tariffs acted as a shock that impacted both groups, but we do not know if treated and non treated firms responded in the same way. Thus, it cofounds and may violate the parallel trends assumption.
Now if the UK had been exposed to the same tariffs but did not implement TWEA, it may be a good option for DiD. This gives us a group that experienced the same tariff shock but without TWEA. We could calcuate the DiD for both the US and UK and then compare this difference.