Econometrics Problem Set 4

Conte, Eustacchi, Kakunuri, Khoso, Wu

21 March 2022

A fundamental question in labor economics is whether a longer duration of unemployment benefits improve the quality of the job after finding employment. However, providing a causal answer to this question is difficult because unemployment duration is correlated with individual characteristics that may also affect job quality. In this exercise, we use a regression discontinuity design for causal identification. Under regular circumstances, an unemployed person receives unemployment benefits for 30 weeks. However, the benefit duration increases to 39 weeks if the person is at least 40 years old when he/she becomes unemployed. The dataset $data\_ps3.dta$ **is a sample of workers who have been laid off and subsequently found employment. The sample contains the following variables:** age **age at layoff** $nonemp$ **non-employment duration in weeks** $jobfind$ **a dummy that equals one if a person had a new job after 39 weeks** $lwage0$ **log monthly wage in previous job** $lwage1$ **log monthly wage in new job**

1) Explain intuitively how this discontinuity in benefit duration can be exploited to estimate the causal effect of benefit duration on wages after re-employment.

In this regression discontinuity design, the running variable is the age of unemployed individuals, where the discontinuity of the sharp RDD is at age of 40, from which the unemployment benefit duration will increase from 30 weeks to 39 weeks. Given that age is independent of subjective decision, we may assume that, by construct, the choice of discontinuity is as good as random.

To estimate the causal effect of the change in benefit duration, we notice that the observable characteristics of age, also our discontinuity running variable, is confounding both the treatment status as well as the potential outcome, while age presumably is not a factor of how treatment of benefit duration affects the wage in re-employment. In other words, backdoor path from this confounding variable is closed by including the running variable in our model. Further, the running variable can be considered as independent of how the treatment would affect the potential outcome, and the choice of the discontinuity cannot be endogenous to treatment assignment. Therefore, we limit the selection bias in the sense that the treatment is exogenous to the assignment rule, individuals will not self-select into the treatment with the on-and-off assignment rule.

To confirm that there is no self-selection, data should show no bunching around the cut-off threshold, which would indicate no evidence of manipulation around the discontinuity for running variable. The running variable then satisfies the continuity assumption, which in fact further implies that the age is not a source of heterogeneity for the state of employment. However, the heterogeneity still endogenously affects the wage for its differences due to individual characteristics other than age. To capture this initial heterogeneous in wage, we use the wage in previous job to proxy such variations. Therefore, the potential outcome is expected to characterise the causal effect of the longer benefit duration on wage in re-employment as closely as possible.

Ideally, around the age threshold, it's only the benefit duration that changes. When individuals re-employed, by comparing those who just above and below the age cut-off, the difference in wage outcome will be considered as the treatment effect of receiving the longer benefit duration.

Additionally, this sharp regression discontinuity design will be informative for the individual characteristics, of which may also affect job quality, since the age cut-off itself with smoothness across the threshold will suggest whether the age is or not one source of heterogeneity for individuals to have different wage when re-employed. The cut-off provides with indication of the exogenous change which will not affect the individual characteristics, say the education, profession, and previous position, etc. Therefore, it offers a comparison of individuals who aged below 40 and above, as well as it compares the difference between whose who receive the treatment and those who not, by which enables us to explore the heterogeneity in individual characteristics other than the age.

2) Write down an estimating equation for a sharp regression discontinuity design whereby you control for the running variable with a second-order polynomial that is allowed to differ above and below the discontinuity. State and explain the identifying assumption that is necessary to interpret your coefficient of interest as causal.

Let age_i index each individual i for our running variable age, and denote the cut-off age for extended employment benefit duration as age_c with $age_c = 40$. The treatment status of having up to 39 weeks of unemployment benefit for each individual, D_i, is given by the assignment rule for this sharp RDD,

$$D_i = \begin{cases} 1 & \text{if } age_i \geq age_i = 40 \\ 0 & \text{if } age_i < age_i = 40 \end{cases} \tag{1}$$

Now suppose a constant treatment effect, where all unemployed individuals aged above 40 would claim extended benefit. Let X_i denote the individual characteristics matrix for age at layoff, age_i, non-employment duration in weeks, $nonemp_i$, new job finding indication dummy, $jobfind_i$, and log monthly wage in previous job, $lwage0_i$. And we have the potential outcomes of log monthly wage in new job, $lwage1_i$, superscript denotes the treatment status,

$$lwage1_i^0 = \alpha_i + \beta_i X_i \tag{2}$$

$$lwage1_i^1 = lwage1_i^0 + \delta_i \tag{3}$$

$$Y_i^0 = \alpha_i + \beta_i X_i \tag{4}$$

$$Y_i^1 = Y_i^0 + \delta_i \tag{5}$$

where $lwage1_i^0 = Y_i^0$ is the potential outcome without treatment, $lwage1_i^1 = Y_i^1$ is the case that treatment occurs at the discontinuity, and δ is the treatment effect parameter at the discontinuity. Thus we have,

$$Y_i = Y_i^0 + (Y_i^1 - Y_i^0)D_i \tag{6}$$

$$Y_i = \alpha_i + \beta_i X_i + \delta_i D_i + \upsilon_i \tag{7}$$

and we have the local treatment at discontinuity as follows:

$$\delta_i = \lim_{X_i \to age_c} E[Y_i^1 | X_i = age_c] - \lim_{X_i \leftarrow age_c} E[Y_i^0 | X_i = age_c] \tag{8}$$

$$= \lim_{X_i \to age_c} E[Y_i | X_i = age_c] - \lim_{X_i \leftarrow age_c} E[Y_i | X_i = age_c] \tag{9}$$

$$= E[Y_i^1 - Y_i^0 | X_i = age_c] \tag{10}$$

Now suppose a nonlinear relationship for individual characteristics matrix,

$$E[Y_i^0 | X_i] = f(X_i) \tag{11}$$

and we have,

$$Y_i = f(X_i) + \delta D_i + \eta_i \tag{12}$$

which is valid when the smoothness of $f(X_i)$ on X_i distribution space is satisfied. Note that this is the nonlinear counterpart of a linear relationship in equation (7).

However, given the treatment will affect the age profile of outcome on each treated individual, we do not observe the potential outcome for the same treated individuals if there were no treatment (i.e., the counterfactual is in range of $X_i > age_c$). By employing higher-order polynomials, we approximate the age profile with $f(X_i)$ for different X_i functions for values above and below the cut-off with their corresponding interaction terms with treatment status D_i. Note that $E[Y|X] = E[Y^0|X] + (E[Y^1|X] - E[Y^0|X])D$, and we rearrange the two second-order polynomials on each side of threshold to obtain,

$$E[Y_i^0|X_i] = f_0(X_i) = \lambda_i + \beta_{01i}\ \widetilde{X}_i + \beta_{02i}\ \widetilde{X}_i^2 \tag{13}$$

$$E[Y_i^1|X_i] = f_1(X_i) = \lambda_i + \beta_{11i}\ \widetilde{X}_i + \beta_{12i}\ \widetilde{X}_i^2 + \delta_i \tag{14}$$

$$Y_i = \lambda_i + \beta_{1i}\ \widetilde{X}_i + \beta_{2i}\ \widetilde{X}_i^2 + \gamma_{1i}\ \widetilde{X}_i\ D_i + \gamma_{2i}\ \widetilde{X}_i^2\ D_i + \delta_i D_i + \eta_i \tag{15}$$

where $\widetilde{X}_i$ is the running variable that recentred at $X_i - age_c$ with change in coefficients of intercept term. Coefficient estimates of interest at discontinuity give $\gamma_1 = \beta_{11} - \beta_{01}$, and $\gamma_2 = \beta_{12} - \beta_{02}$, where subscript p and q ($p = 0, 1$; $q = 1, 2$) in β_{pq} and γ_{pq} denote the treatment status and the index for higher-order coefficient parameter, respectively. Note that for simplicity, we replace $\widetilde{X}_i$ with X_i for the following discussion. Equation (15) is reduced to (7), when assuming the age profile $f(X_i)$ is linear or is the same for values above and below the cut-off. Namely, $\beta_{11} = \beta_{01}$ and $\beta_{12} = \beta_{02}$, so $\gamma_1 = 0$ and $\gamma_2 = 0$.

Now consider the benefit duration as in an accumulation amount for individuals to claim when unemployed, for those who stay in unemployed less than 30 weeks, they are indifferent on the benefit duration, and so they are not the intended treatment group. However, for those who will stay unemployed from 31st week onward, the longer benefit duration means a higher amount of benefit based on some criteria of the entitlement to the additional amount, in the current setting it is the age of 40. Therefore, we are interesting in the effect of higher unemployment benefit on wage outcome in re-employment, particularly for those who stay unemployed for more than 30 weeks.

Supposedly, the individual characteristics that may not just affect job quality, but also affect their state of employment, heterogeneity then gives the treatment in unemployment benefit with three subcompliant population: always-takers, whose age is greater than 40; never-takers, whose age is less than 40 when layoff, or stay unemployed less than or equal to 30 weeks; compliers, who stay unemployed more than 30 weeks, claiming the standard amount if younger than 40 when layoff, or claiming the additional amount if older than 40 when layoff. Hence, the true relationship for wage in re-employment for each individuals depends on both the treatment status as well as the state of employment,

$$Y_i = \alpha_i + \beta_i X_i + \delta_i D_i + \Delta_i age_i D_i + \epsilon_i \tag{16}$$

Assuming perfect compliance, when every unemployed claims their benefit fully and accordingly based on the age and the unemployment duration criteria, our interest of intention-to-treat (ITT) lies upon the additional benefit amount on wage outcome for the compliers who stay unemployed for more than 30 weeks,

$$\begin{aligned} Reduced\ Form &= E[Y_i|D_i = 1, i \in complier] - E[Y_i|D_i = 0, i \in complier] \\ &= \widehat{\delta}_i \end{aligned} \tag{17}$$

which gives the effect of more benefit on wage difference in re-employment after more than 30 weeks of staying unemployment. Thus, the reduced form identifies the treatment effect on wage difference for 30-more-weeks unemployed individuals who are around the threshold of age cut-off that have found new job, by looking at the wages between those who claimed higher total amount of unemployment benefit and those who claimed standard amount.

Now back to the heterogeneity in the sense of structural distribution, suppose a business follows a worker-supervisor-manager-chief pyramid-like organisational structure, where the age is positively associated with position in hierarchy, the number of peers is negatively associated with position. By this construct, with a proportionate layoff, younger individuals with lower rank in position are more likely to lose their job. Therefore, the state of employment and the duration of unemployment is considered conditional on the running variable of age. In other words, the treatment effect is also conditioning on the structural distribution, where the state of employment and the duration of unemployment depend on the age (i.e., unemployment duration distribution is conditioning on age distribution). Therefore, the first stage gives the difference in unemployment duration that based on the treatment assignment, which is the actual treatment for different unemployment duration,

$$\begin{aligned} First\ Stage &= E[nonemp_i|D_i = 1] - E[nonemp_i|D_i = 0] \\ &= \widehat{\Delta_i} \end{aligned} \tag{18}$$

The first stage takes into account the fact that, younger the individuals the more likely for them to lose their job. Since with more job experience, workers will move to senior level within the organisation, and less "replaceable" they will become. Given the same wage, more skilled employees are less likely to become unemployed in a layoff, when considering both their productivity and network. For those who fall into low wage range with older age, heterogeneity implies that their potential outcome after the treatment will be different to start with. Similarly, this argument holds for re-employment, it is easier for skilled workers with more job experience to find a new job after the layoff. That says, in the sense of heterogeneity, the distribution of 30-more-weeks unemployed is different from the distribution of 30-less-weeks unemployed, and is different from the distribution of employed, and finally is different from the distribution of entire labour supply.

Given the discontinuity design and variables of interest, our treatment effect of interest is to look at individuals whose local treatment effect at discontinuity divided by difference in the actual treatment. We estimate with the Wald estimator, from (17) and (18), we have,

$$Wald\ Estimator = \frac{E[Y_i|D_i = 1, i \in complier] - E[Y_i|D_i = 0, i \in complier]}{E[nonemp_i|D_i = 1] - E[nonemp_i|D_i = 0]} \tag{19}$$

$$\frac{\widehat{\delta_i}}{\widehat{\Delta_i}} = \lim_{X_i \to age_c} E[Y_i|X_i = nonemp_i] \quad - \lim_{X_i \leftarrow age_c} E[Y_i|X_i = nonemp_i] \tag{20}$$

3) Carry out two tests for the validity of the RD design: 1) plot the density of age at layoff; 2) plot the log previous wage against the age at layoff. For each test, produce a scatter plot with bin size 4 months (i.e. each point summarizes the average value on the vertical axis for workers whose age at layoff falls within that bin). For easier visual inspection, the plot should contain a vertical line at the discontinuity and lines for the second-order polynomials above and below the cutoff. What do those graphs tell us about the validity of the RD design?

Vaild. discontinuity is important around the threshold, for observation outside the band-with, there is selectivity of observations. The treatment effects is around the cut off and since this is of interest to us, we can say that the RD design is valid. From which the potential outcomes differ, it is outside the bandwidth of our interest, where the heterogeneity bias more significantly (strictly wrong, but intuitively the concept, will rephrase) .

4) Produce the main results graphically. We focus here on three outcomes, namely non-employment duration, the probability of finding a job within 39 weeks and log wages in the new job. Use the same binned scatters as in exercise 3). Interpret your results.

5) Focusing on the same three outcomes, report the coefficient of interest (i.e. the coefficient of a dummy for being above or below the discontinuity) of the regression outlined in exercise 2). Produce a regression table with five panels

- **The reduced-form effect at the discontinuity in the full sample using the regression from exercise**
- **The reduced-form effect at the discontinuity with a bandwidth of ±5 years in age at layoff.**
- **The reduced-form effect at the discontinuity in the full sample using a linear control for the running variable, allowing for different slopes above and below the discontinuity.**
- **The reduced-form effect at the discontinuity in the full sample, controlling for the running vari- able with a fourth-order polynomial and, allowing for different parameters above and below the discontinuity**
- **The reduced-form effect at the discontinuity based on the optimal bandwidth computed based on the producedure by Calonico et al. (2014). You can use their Stata/R pack-**

age *rdrobust*. The package offers multiple procedures to calculate the bandwith. Use the default procedure. Report the estimated bandwidth along with each coefficient. Interpret and discuss the differences between the panels.

check for functional form

bandwidth robustness check