ECON 7130 - Microeconomics III

Spring 2016

Notes for Lecture #8

Today:

- Regression Kink Design
- Bunching at Kinks
- Examples of papers exploiting kinks

Regression Kink Design

- General idea: RDD exploits discontinuities in likelihood of treatment as a function of some continuous variable (the assignment variable). RKD is going to exploit discontinuities in treatment effects as a function of of the first derivative of the assignment variable; Δtreatment/Δassignment variable.
  - As with RDD, RKD will exploit policies that create arbitrary cutoffs. Now we'll be looking at discontinuities in the first derivative, or kinks. E.g. The income tax schedule in the US, phaseouts of tax credits, bulk pricing.
  - The regression kink design relies on a regressor (such as a policy variable) that is a deterministic function of a "behaviorially endogenous variable" that cannot be replaced by a plausible instrument (because the outcome of interest is also associated behaviorally) (Card et al., 2009).
  - Unlike RD designs, the running variable does not determine whether or not a treatment is given. Rather, the slope of the treatment function changes at a kink point.
- Can have fuzzy or sharp RKD
- The regression analysis is run first by retrieving β_1 in this equation:

$$E[Y|V=v] = \alpha_0 + \sum_{p=1}^{p}[\alpha_p(v-k)^p + \beta_p(v-k)^p * D], \tag{1}$$

  - Where Y is the outcome variable of interest; k is the kink point; v is the running variable, and D allows the slopes to change at the kink.
  - The β_1 retrieves the change in slope in outcomes at the kink point in the treatment variable.
- The causal impact per unit is found by dividing β_1 by the difference in derivatives around the kink which is calculated mathematically rather than econometrically.
  - This step may not be intuitive. Think about it a little bit.
  - The kink in outcomes needs to be scaled by the kink in treatment.
  - Consider an outcome where there's a modest kink. Perhaps the kink in treatment is also modest, or perhaps it is fairly striking. We account for the degree of the kink by dividing the outcome slope change by the treatment slope change.
  - The estimand can be considered the "treatment on the treated" (TT) effect as from a randomized experiment, as long as certain conditions are met regarding the distribution of all other predetermined factors (Card et al. (2009)).
- Here's a nice graph that may shed some more light on RK designs: This is from Camille Landais' paper on unemployment insurance. Here, benefits increase with highest quarterly earnings until some maximum level. If you earn more than a certain amount, you get the same level of unemployment benefits regardless of how much more you earn. Camille exploits this. He finds this kink in the duration of unemployment when plotted on the highest quarterly earnings of the individual (this is the running variable). He then divides the change in the slope in unemployment duration by the change in the slope of the benefits level at that kink.

- Identification:
  - Rules are often arbitrary so provide a natural experiment where treatment and control status is determined in a way that is uncorrelated with unobservables
  - Exploit this for quasi-random assignment to treatment/control
  - Key assumptions:
    * Kink in treatment results in a discontinuous change in the slope of the outcome variable
    * There is not a discontinuity in a covariate that cause the discontinuity in the slope of the outcome variable at that value of the forcing variable
    * The outcome variable's slope is continuous in the forcing variable - in order to identify the average treatment effect, we need to make a smoothness assumption
    * There is a unavoidable need for extrapolation, because by design there are no units with $X_i = k$ for whom we observe the same treatment amount as for those below k. We therefore will exploit the fact that we observe units with covariate values arbitrarily close to k.
    * No manipulation of forcing variable (e.g., leading to bunching around kink points)
- Graphical analysis
  - RKD lends itself especially well to graphical analysis
  - Graphical inspection is key
  - 3 types of graphs should always be produced, where assignment variable is graphed against:
    1. The outcome
    2. Other covariates
    3. Density of cases
  - 1 should show a discontinuity at the kink points and no other discontinuities that can't explain
  - 2 and 3 should show no discontinuities at kink points - if so, be worried about identification
  - If you can't see the main result with such a simple graph, it's probably not there
  - If you see a discontinuity in 2 or 3, be worried
- RKD in Stata
  - Just run OLS with polynomials
- Examples of discontinuities in treatment effects
  - Unemp compensation policies
  - Tax credit phase-ins and phase outs (EITC, Child Tax Credit, Saver's Credit, etc)
  - Non-linear tax schedules (manipulation??)
  - Non-linear pricing (e.g. bulk pricing) (manipulation??)
  - Fuel economy policies (e.g. CAFE standards)

<u>Bunching analysis</u>

- Often the term "notch" is used for kinks here (and with RKD??) - I'll use kink, but when searching the literature, use both
- Idea: Estimate elasticities by looking at observed behavior around kink points
- E.g. Number of individuals moving just below income level to trigger higher tax rate will tell you something about the elasticity of taxable income

- DRAW kinks budget constraint and show how will bunch
- DRAW how distribution will look with bunching
- Recall that with RDD/RKD, we were in trouble if individuals manipulated the forcing variable. Now we are using this manipulation to estimate parameters that govern behavior
- Identification:
  - The x elasticity of y (compensated) is approximately equal to %change in density of y at the kink/% change in x at kink, where the kink is in the schedule for x
  - $e = \frac{dy/y}{dx/x} \simeq \frac{\text{excess mass at kink point}}{\%\text{ change in } x \text{ at kink}}$
  - Assumes agent fully optimize and manipulate the variable y in response to kinks in x
  - Assumes smooth distribution of y in the absence of the kink
  - Either assumes homogenous elasticity or measures average elasticity for sample population
- Graphical analysis:
  - Lends itself well to graphical analysis
  - Graph want to show: histogram of density over x, with notation showing kink points in x
- How to get estimates of elasticity:
  - Find amount of bunching:
    * Limit sample to those with $x \in [x - 2\delta, x + 2\delta]$, where δ is your bandwidth
    * Choose δ by looking at graph and see what need to capture bunching (also will move around for robustness checks)
    * Calculate the fraction of sample in three regions: $H(x) = \int_{x-\delta}^{x+\delta} h(x)dx$, $H(x)_- = \int_{x-2\delta}^{x-\delta} h(x)dx$, $H(x)_+ = \int_{x+\delta}^{x+2\delta} h(x)dx$, where $h(x)$ is the density of x (so to find $H(x)$ you are just summing up the number of obs in each of these regions and dividing by the total number of obs in the three regions)
    * $\hat{h}(x) = H(x)/\delta$
    * Excess bunching $= B = H(x) - H(x)_- - H(x)_+$ (note that the width of $H(x)$ is twice the width of either of the others)
    * Note that this assumes a uniform distribution around the kink - you can get more fancy (see Chetty et. al 2011)
  - Calculate elasticity as $e = B/\%\Delta$x at kink
    * The $\%\Delta$x at kink should be given policy parameters
    * E.g. with kink in income tax schedule going from τ_0 to τ_1, $\%\Delta$x at kink $= \frac{\tau_1 - \tau_0}{1 - \tau_0}$
- Standard errors can be calculated 2 ways:
  1. Bootstrapping: random samples from sample and find std. errors in estimates by calculating std dev of all estimates
  2. Delta Method: $Var(G(X)) = Var(X) * [G'(\mu)]^2$, where μ is mean of X
     - Find variance in outcome variable (e.g. income)
     - Find how elasticity changes as e.g. income changes by numerical derivative
- Caveat to all this: Need data with little measurement error
- Bunching in Stata:
  - Produce histograms

- Calculate fraction of population in different regions around kink
    - This gives the excess bunching around kink
    - Use this to calculate elasticity

- Think about all the kinks mentioned above - can do bunching analysis on them if you think people are manipulating behavior in response to kink

RKD: Example 1, Card, Lee, Pei, and Weber, "Nonlinear Policy Rules and the Identification and Estimation of Causal Effects in a Generalized Regression Kink Design", (NBER WP, 2012):

- Question: (mostly methodological about RKD, but do have example): Does the level of unemployment benefits in Austria affect the length of unemployment?
- Problem with straight forward answer: length of unemployment is endogenous - likely omitted variable
- Background:
    - Unemployment benefits based on income in base period.
    - The benefit formula for unemployment exhibits 2 kinks
        * There is a minimum benefit level (that is not binding for people with very low earnings)
        * Then benefits are 55% of the earnings in the base period
        * There is a maximum benefit level that varies year to year
    - People with dependents get small supplements (thus several parallel lines in graph)
    - Note that errors in administrative data mean some don't get benefits that correspond exactly to formula
- Data:
    - Administrative data from Austria: Austrian Social Security Database
    - Data records employment and unemployment spells on a daily basis for all individuals employed in the Austrian private sector
    - Merge ASSD with UI benefits data - date of claim, claim amount, duration of unemployment
    - Sample includes 369,566 unemployment spells initiated between 2001 and 2008
- Basic model:
    - Fuzzy RKD since policy rule is employed with error
    - Consider local linear and local quadratic models
    - Robustness checks with varying bandwidth
- Identification:
    - Assume covariates distributed in a smooth fashion around kink points (i.e., over income)
    - Then variation around kink in outcome variable (duration of unemp) caused by kink in benefits schedule
    - SHOW GRAPHS FROM PAPER
- Results:
    - People with higher base earnings have less trouble finding a job (negative slope).
    - There is a kink: the elasticity of joblessness with respect to the benefit rate is on the order of 1.5

Bunching, Example 1, Saez, "Do Taxpayers Bunch at Kink Points?" , (*AEJ: Economic Policy*, 2010):

- Question: What is the the elasticity of taxable income?
- Problem: can't directly estimate with cross-sectional data - a person's marginal tax rate is endogenous to the amount of income they earn. Progressive taxes make it hard to identify effects without panel.
- Solution: Leverage manipulation of taxable income around tax brackets to identify the elasticity.
- Idea/Identification:
  - With kinks in budget constraint, not all indifference curves will be tangent to BC - there will be bunching at/below kink points where marginal tax tax rates increase.
  - The amount of bunching at kinks will imply the elasticity of taxable income
  - SHOW Panel A figure
- Note: relies on assumptions that individuals know a lot about structure of tax code and optimize (need to have assumption of no manipulation that was made for RKD/RDD to be broken)
- Data: IRS public use files for individual income tax filers, 1960-2004 (benefit is less reporting error than survey data)
- Basic Model: structural model of taxpayer optimization which implies that amount of bunching is a function of the elasticity of taxable income and distribution of tax payers
- Results:
  - Finds bunching around:
    - First kink point of the Earned Income Tax Credit (EITC), especially for self-employed
    - At threshold of the first tax bracket where tax liability starts, especially in the 1960s when this point was very stable
  - SHOW GRAPHS
  - However, no bunching observed around all other kink points
- Why not more bunching?
  - True intensive elasticity of response may be small
  - Randomness in income generation process
  - Lack of information and salience
  - Frictions: Adjustment costs and institutional constraints (Chetty et al. QJE'11)