Examples of RDD in Applied Economic Research

Dr. Kamiljon T. Akramov
IFPRI, Washington, DC, USA

Regional Training Course on Applied Econometric Analysis

June 4-15, 2018, WIUT, Tashkent, Uzbekistan

Do benefits of additional medical expenditures exceed their costs? (Almond et al. QJE, 2010)

- RDD allows to compare health outcomes and medical treatment provision for newborns on either side of the very low birth weight threshold at 1,500 grams
- Study finds that newborns with birth weights just below 1,500 grams have *lower* one-year mortality rates than do newborns with birth weights just above this cutoff, even though mortality risk tends to decrease with birth weight
- One-year mortality falls by approximately one percentage point as birth weight crosses 1,500 grams from above
- Infants with birth weight < 1,500 grams receive more medical treatment and their hospital costs higher by $4,000 relative to mean hospital costs of $40,000 for infants with birth weight just above 1,500 grams
- Assuming observed medical spending fully captures the impact of the “very low birth weight” designation on mortality, the study estimates suggest that the cost of saving a statistical life of a newborn with birth weight near 1,500 grams is on the order of $550,000 in 2006 dollars

Economic impact of unionization (DiNardo & Lee, QJE 2004)

- Estimation of economic impacts of unionization is difficult due to selection bias
- Unions could organize at highly profitable enterprises that are more likely to grow and pay higher wages
- Union elections
  - If employers want to unionize, board holds election
  - 50% or less means the employer doesn't have to recognize the union, and
  - 50% + 1 means the employer is required to "bargain in good faith" with the union
- Multiple establishment-level datasets that represent establishments that faced organizing drives in the United States during 1984-1999

Economic impact of unionization DiNardo & Lee, QJE 2004 (cont.)

- The paper applies RD design to estimate the impact of unionization on business survival, employment, output, productivity, and wages
- Paper essentially compares outcomes for employers where unions barely won the election with those where the unions barely lost
- The analysis finds small impacts on all outcomes
- The results suggest that-at least in the study period-the legal mandate that requires the employer to bargain with a certified union has had little economic impact

Effects of class size on test scores (Angrist & Levy, QJE 1999)

- Fuzzy RD design is used to estimate the effects of class size on student's test scores
- School class size- Maimonides' rule
  - No more than 40 kids in a class in Israel
  - 40 kids in school means 40 kids per class
  - 41 kids means two classes with 20 and 21 kids
- Multiple discontinuities: causal variable of interest, class size, takes on many values
- Nonlinear relationship between the local number of students and the class size predicted by Maimonides' rule to estimate the impact of class size on student performance, and evaluate the effect of being just below the number of students for whom an additional teacher would be brought up, and of being just above this number
- First stage exploits jumps in average class size
- Finding: smaller class size increases test scores
- The results have shown highly irregular patterns in class size that are precisely mirrored in student achievement: a reduction in predicted class size of ten students is associated with a 0.25 standard deviation increase in fifth-graders' test scores

RD examples from literature (cont.)

- Anderson and Magruder (2012) and Lucas (2012)
  - Yelp.com ratings have an underlying continuous score
  - Distribution determines cutoff points for 1 to 5 stars
  - Effect of an extra star on future reservations and revenue
- Anderson et al. (2012)
  - Young adults lose their health insurance as they age (older than 18 and in college but different after ACA)
  - Age changes the probability of having health insurance (fuzzy design)

Paper by Raffaello Bronzini and Eleonora Iachini (AEJ: Economic Policy, 2014)

- The paper uses sharp RDD to evaluate a unique R&D subsidy program implemented in northern Italy
- Firms were invited to submit proposals for new projects and only those which scored above a certain threshold received the subsidy.
- It compares the investment spending of subsidized firms with that of unsubsidized firms

Main questions in empirical research

- What is the policy question?
- What is the causal relationship of interest?
- What is the dependent variable and how is it measured?
- What is (are) the key independent variable(s)?
- What is the data source?
- What is the identification strategy?
- What is the mode of statistical inference?
- What are the main findings?

Policy question

- Governments spend substantial financial resources to support private R&D activities
  - Direct government funding of private R&D in OCED countries amounts about 0.1% of GDP ($16.5 trillion), excluding tax incentives – $16.5 billion
- Economic rationale
  - Market failure
  - Liquidity constraints
- Do R&D investment subsidies actually work, i.e., increase private R&D expenditures?
- In theory, public subsidies are expected to increase private R&D investment by reducing the cost of capital and increasing expected investment profitability
  - Inframarginal versus marginal projects
- Empirical research yield mixed results
- Do benefits of additional government expenditures on investment subsidies exceed their costs?

Program

- “Regional Program for Industrial Research, Innovation and Technological Transfer” implemented in Emilia-Romagna (Italy)
- The regional government subsidizes the R&D expenditure of eligible firms through grants, the grant may cover up to
  - 50% of the costs of industrial research projects
  - 25% for precompetitive development projects; the 25% limit is extended by an additional 10% if applicants are SMEs
- The maximum grant per project is €250,000
- Duration of the investment is from 12 to 24 months

Causal relationship of interest, dependent and key independent variables

- Relationship between government R&D subsidies and private R&D activity (expenditures)
- Dependent variable
  - Natural candidate would be R&D investment, but not available
  - Net investment calculated from the balance-sheet data as annual differences in tangible or intangible assets net of amortization
- Key independent variables
  - Binary treatment variable for an R&D subsidy
  - Score – total 100 points
    - technological and scientific (max. 45 points)
    - financial and economic (max. 20 points)
    - managerial (max. 20 points);
    - regional impact (max. 15 points)
    - Only projects deemed sufficient in each category and which obtain a total score of at least 75 points receive the grants

Identification strategy

- Goal is to evaluate whether subsidized firms would not have made the same amount of R&D outlays without the grants
- Subsidized and nonsubsidized firms can differ in terms of unobserved characteristics correlated with the outcome
- Therefore, the variable identifying recipient firms in empirical analysis can be endogenous
- To deal with the endogeneity issue, paper exploit the funds' assignment mechanism
- Only those receiving a score equal to or above a given threshold (75 out of 100) were awarded grants

Identification strategy (cont.)

- The paper applies a sharp RDD comparing the performance of subsidized and nonsubsidized firms with scores close to the threshold
- By letting the outcome variable be a function of the score, the average treatment effect of the program is assessed through the estimated value of the discontinuity at the threshold

Empirical specification

$$Y_i = \alpha + \beta T_i + (1 - T_i) \sum_{p=1}^{3} \gamma_p (S_i)^p + T_i \sum_{p=1}^{3} \gamma_p' (S_i)^p + \varepsilon_i$$

where

Y_i is the outcome variable;

T_i = 1 if firm i is subsidized (all firms with score $\geq$ 75) and T_i = 0 otherwise;

S_i = $Score_i - 75$;

γ_p and γ_p' are the parameters of the score function and allowed to be different on the opposite side of the cutoff to allow for heterogeneity of the function across the threshold;

ε_i is the random error.

Data

- Balance sheet data provided by Cerved group, which collects data from all Italian corporations
  - Start-up costs, R&D and advertising costs, costs of patents, software, and other intellectual property rights, licenses and trademarks, costs of ongoing intangible assets, etc.
- Administrative data from Emilia-Romagna Region
  - Name, score, planned investment, grants assigned, subsidies revoked and renunciations
- Pooled data from two invitations (2004 & 2005)
  - 1246 firms: 557 treated and 689 untreated
  - 411 unsubsidized firms that didn’t receive a score in 2005 were excluded
- Final sample included 357 industrial (254 treated and 103 untreated) and 111 service firms (61 treated and 50 untreated)

Estimation

- First, a third order polynomial model was estimated on the full sample
- Second, equation was estimated through local regressions around the cutoff point using two different sample windows
  - Firms with scores between 52 and 80 (50% of the baseline sample)
  - Firms with scores between 66 and 78 (35% of the baseline sample)
- Third, paper estimated the discontinuity using other nonparametric techniques, namely the kernel regressions using two bandwidths, 30 and 15 points of the score

Estimation (cont.)

- The OLS estimates of the parameter β measures the value of the discontinuity of function $Y(S_i)$ at the cutoff point, corresponding to the unbiased estimate of the causal effect of the program
- A coefficient β equal to zero would signal complete crowding-out of private investment by public grants
  - This would mean that firms reduced private expenditure by the amount of the subsidies received and the investment turned out to be unaffected by the program
- A positive coefficient would show that overall treated firms invested more than untreated firms, plausibly thanks to the program, and that total crowding-out did not occur

Main findings

- Overall, no significant increase in investment
- Substantial heterogeneity in the program's impact
- Small enterprises increased their investments—by approximately the amount of the subsidy they received—whereas larger firms did not

Data and Stata codes

- Data and Stata codes are in the folder