# Regression Discontinuity 1: Framework and Application

Instructor: Yuta Toyama

Last updated: 2021-05-18

# Introduction

# Introduction

- **Regression Discontinuity Design (回帰不連続デザイン)**
    - Exploit the discontinuous change in treatment status to estimate the causal effect.
- Example:
    - Threshold of test score for college admission
    - Eligibility of policy due to age.
    - Geographic boundary of two regions.

テレ朝▶news

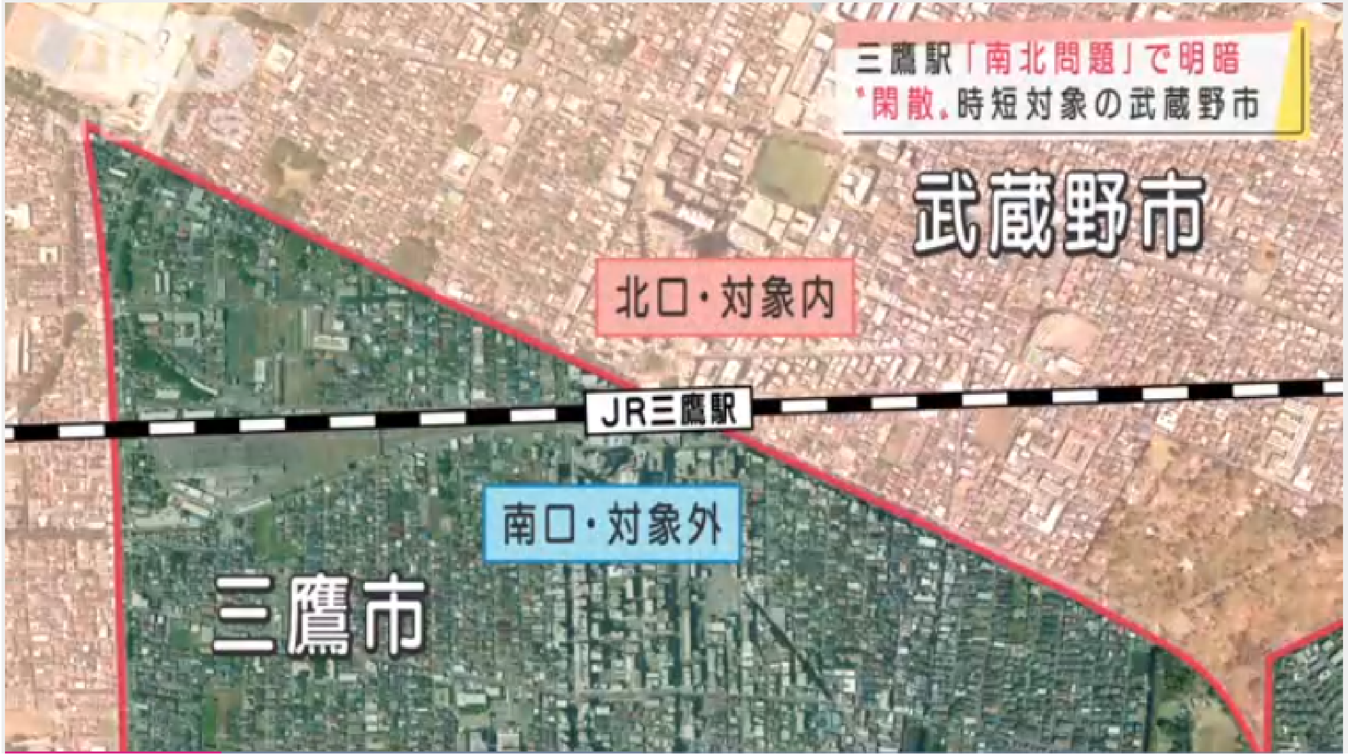
FOLLOW US

テレ朝newsを検索

トップ | 新型コロナ | 速報 | 社会 | 政治 | 経済 | 国際 | スポーツ

# 三鷹駅の「南北問題」　“明暗”分かれ飲食店は…

[2021/04/13 19:45]



# RD Idea

# Course Plan and Reference

- Plan
  - Framework
  - Estimation
  - Application: Shigeoka (2014)
  - Implementation in R

- Reference
  - Angrist and Pischke "Mostly harmless econometrics" Chapter 6
  - R packages: https://sites.google.com/site/rdpackages/rdrobust

# Framework

# Framework

- $Y_i$: observed outcome for person $i$

- Define *potential outcomes*
  - $Y_{1i}$: outcome for $i$ when she is treated (treatment group)
  - $Y_{0i}$: outcome for $i$ when she is not treated (control group)

- $D_i$: treatment status is deterministically determined (sharp RD design)

$$D_i = \mathbf{1}\{W_i \geq \bar{W}\}$$

  - $W_i$: running variable (forcing variable).
  - Probabilistic assignment is allowed (**fuzzy RD design**)

# Example: Incumbent Advantage

- Consider the two-candidate elections
  - $D_i$: dummy for incumbent in the election
  - $Y_i$: whether the candidate win in the election
  - $W_i$ : the vote share in the previous election.

- The incumbent status is defined as

$$D_i = \mathbf{1}\{W_i \geq 0.5\}$$

- Idea of RD:
  - Suppose that you won with 51%.
  - You are similar to the guy who lose at 49% (main assumption of RD).
  - If you focus on these people, $D_i$ is as if it were randomly assigned.

# Framework cont.d

- Note that $D_i = \mathbf{1}\{W_i \geq \bar{W}\}$ implies the unconfoundedness

$$(Y_{1i}, Y_{0i}) \perp D_i | W_i$$

- But the overlap assumption does not hold

$$P(D_i = 1 | W_i = w) = \begin{cases} 1 & if\, w \geq \bar{W} \\ 0 & if\, w < \bar{W} \end{cases}$$

- To compare people with and without treatment, we need to rely on some sort of extrapolation around the threshold.

# Linear approach

- Suppose for a moment that

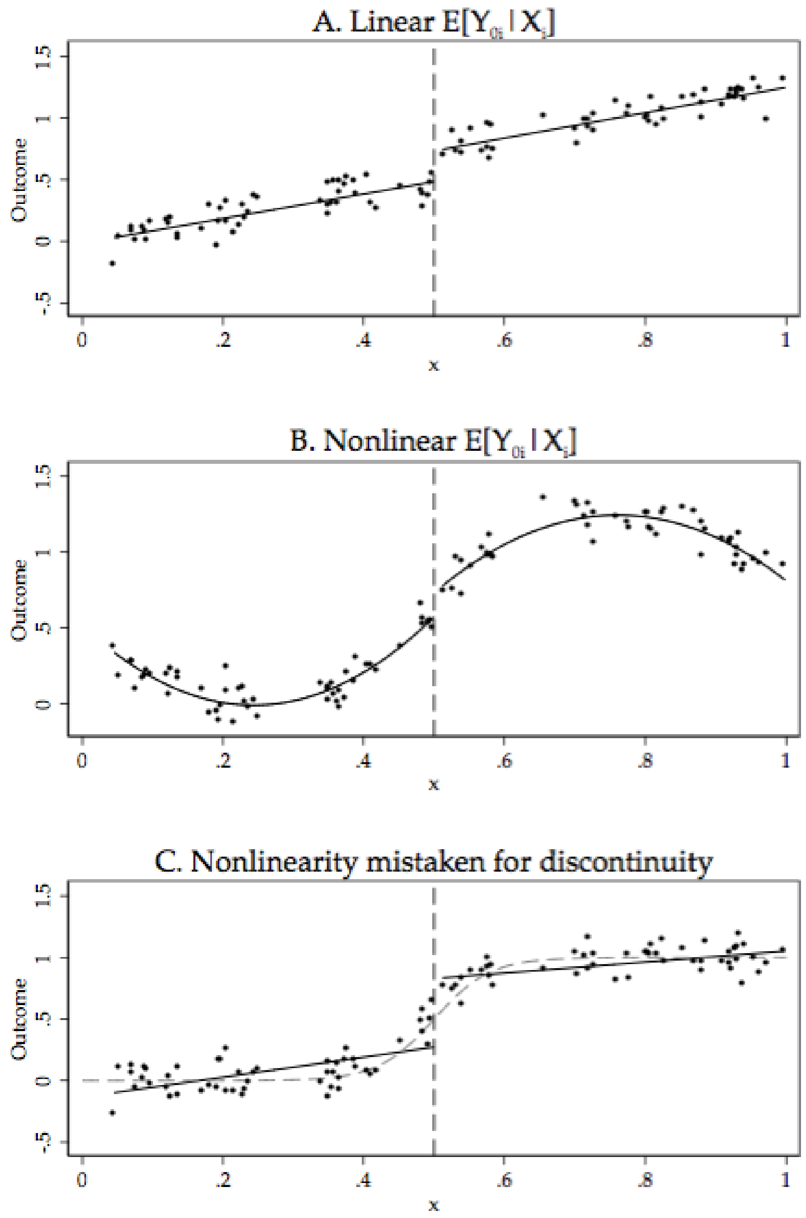
$$Y_{1i} = \rho + Y_{0i}$$
$$E[Y_{0i}|W_i = w] = \alpha_0 + \beta_0 w$$

- This leads to a regression

$$Y_i = \alpha + \beta W_i + \rho D_i + \eta_i$$

  - $\rho$ is the causal effect.

- This approach relies on linear extrapolation. May not be good.
  - What if $E[Y_{0i}|W_i = w]$ is nonlinear?


A. Linear $E[Y_{0i}|X_i]$
B. Nonlinear $E[Y_{0i}|X_i]$
C. Nonlinearity mistaken for discontinuity
Outcome
x
-.5
0
.5
1
1.5
0
.2
.4
.6
.8
1

# A more general approach

- Allowing for nonlinear effect of the running variable $W_i$

$$Y_i = f(W_i) + \rho\mathbf{1}\{W_i \geq \bar{W}\} + \eta_i$$

- A function $f(\cdot)$ might be a $p$th order polynomial.

$$f(W_i) = \beta_1 W_i + \beta_2 W_i^2 + \cdots + \beta_p W_i^p$$

  - nonparametric approach later.

# Implementation in Regression

- Consider

$$E[Y_{0i}|W_i = w] = f_0(W_i - \bar{W})$$
$$E[Y_{1i}|W_i = w] = \rho + f_1(W_i - \bar{W})$$

  where $\tilde{W}_i = W_i - \bar{W}$ is a normalization.
- Then the regression equation is

$$\begin{aligned} Y_i = \alpha + \beta_{01}\tilde{W}_i + \cdots + \beta_{0p}\tilde{W}_i^p \\ + \rho D_i + \beta_1^* D_i \tilde{W}_i + \cdots + \beta_p^* D_i \tilde{W}_i^p + \eta_i \end{aligned}$$

- When running regression, need to focus on the sample around threshold.
- How close the sample should be to the threshold can be taken care by statistical procedure.

# Example

# Effects of the minimum age drinking law

FIGURE 4.1
Birthdays and funerals

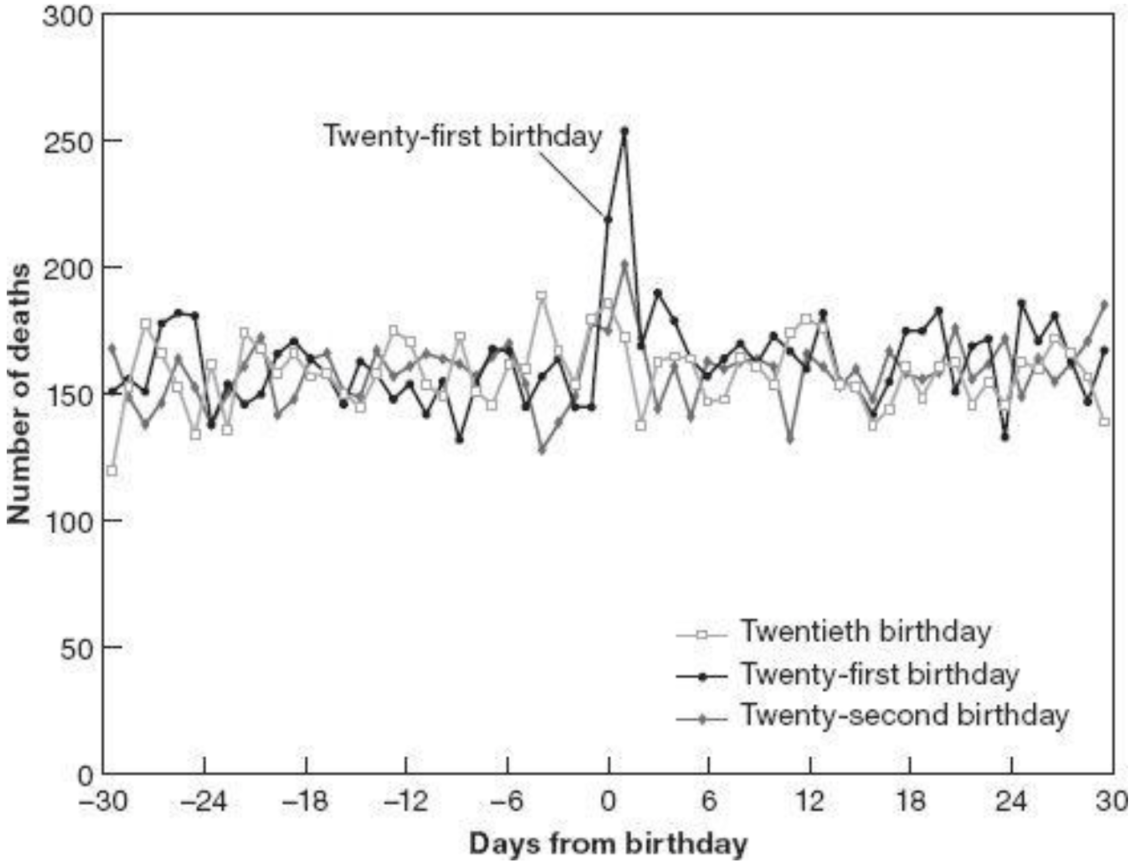
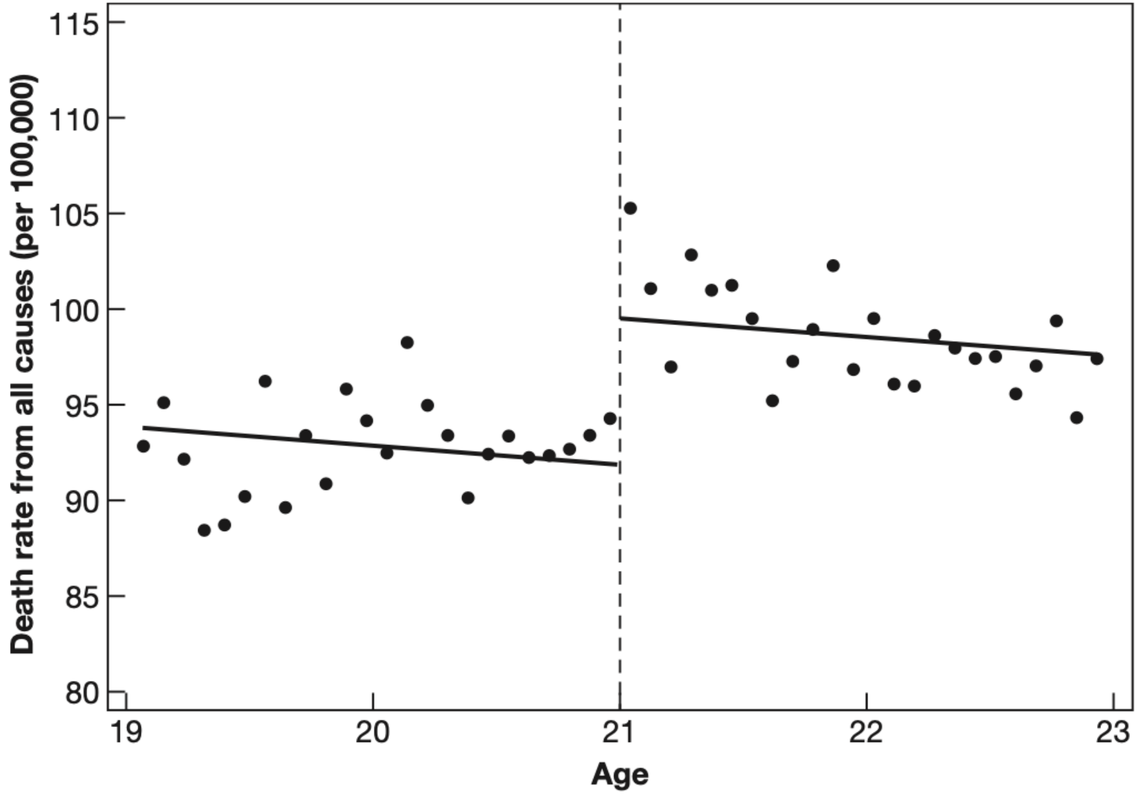


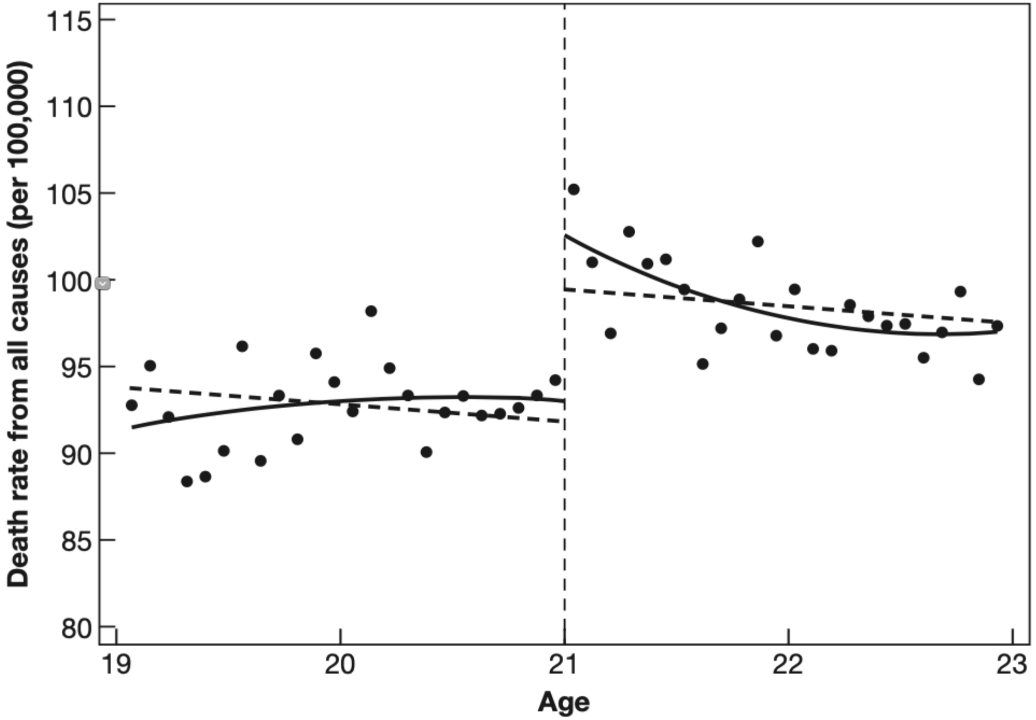
FIGURE 4.2
A sharp RD estimate of MLDA mortality effects




*Notes:* This figure plots death rates from all causes against age in months. The lines in the figure show fitted values from a regression of death rates on an over-21 dummy and age in months (the vertical dashed line indicates the minimum legal drinking age (MLDA) cutoff).

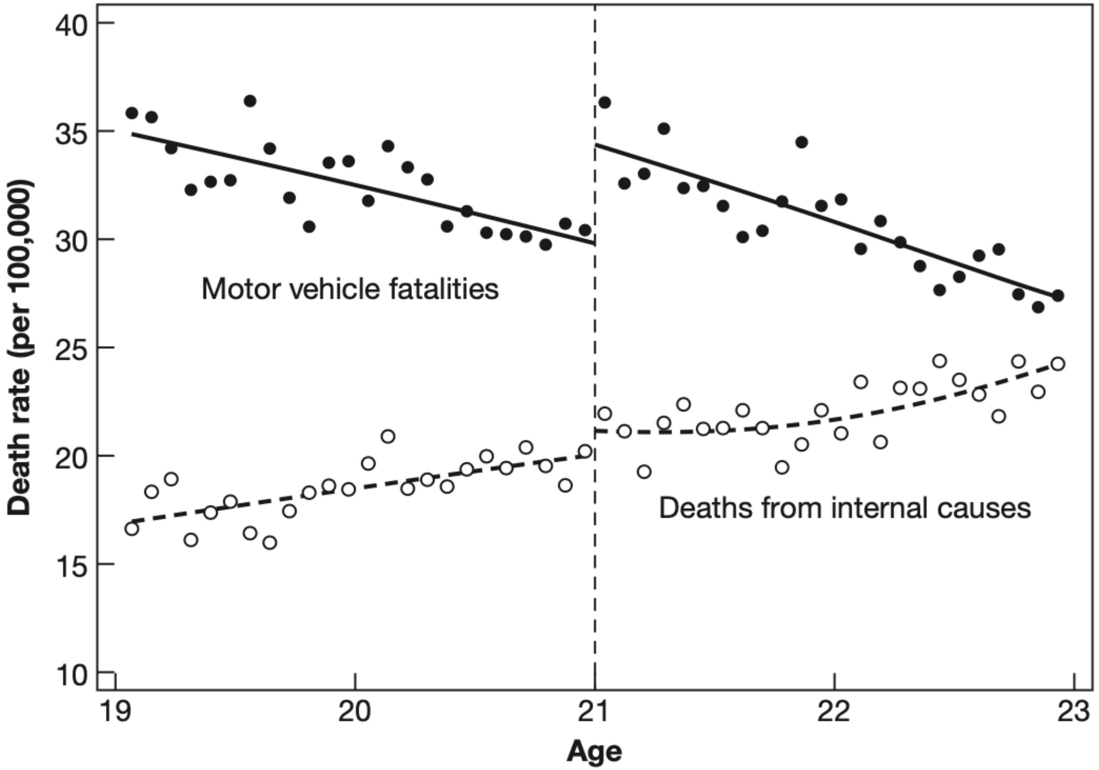
tel:115%20110%20105%20100

FIGURE 4.4
Quadratic control in an RD design




*Notes:* This figure plots death rates from all causes against age in months. Dashed lines in the figure show fitted values from a regression of death rates on an over-21 dummy and age in months. The solid lines plot fitted values from a regression of mortality on an over-21 dummy and a quadratic in age, interacted with the over-21 dummy (the vertical dashed line indicates the minimum legal drinking age [MLDA] cutoff).

Figure 4.5
RD estimates of MLDA effects on mortality by cause of death




*Notes:* This figure plots death rates from motor vehicle accidents and internal causes against age in months. Lines in the figure plot fitted values from regressions of mortality by cause on an over-21 dummy and a quadratic function of age in months, interacted with the dummy (the vertical dashed line indicates the minimum legal drinking age [MLDA] cutoff).

# Validation of Assumptions

# Validation of Assumptions for RD

- Key assumption 1: STUVA. No spill over of treatment across threshold.
  - See next slide.

- Key assumption 2: Continuity of potential outcomes at the threshold.

# Violation of STUVA

# Continuity of

- The key assumptions : Both $E[Y_{1i}|W_i = w]$ and $E[Y_{0i}|W_i = w]$ are continuous at the threshold $w = \bar{W}$.

- This is not directly testable because we cannot observe $Y_{1i}$ below the threshold.

- There are two common approaches that support this assumption:
  1. Covariate test
  2. Density test (no bunching in the running variable).

# Covariate Test

- The underlying idea of RDD: Comparing outcomes right above and right below $\bar{W}$ provides a comparison of treated and control agents who are similar due to the assumed continuity in conditional distributions.

- If this is a valid comparison, then we would expect that covariates $X$ also change smoothly as we pass through the threshold.

- Run the RDD on the covariate $X$.
- If we found the discontinuity, it suggests that the conditional expectation of $Y$ on $W$ may not be continuous either.
- If $X$ has a direct effect on $Y$, the discontinuity in $E[Y_i|W]$ at $\bar{W}$ will confound the treatment effect.
- Example:
  - $Y$ hours worked,
  - $D$: older-than-65 discounts,
  - $W$: age, $X$: social security benefit (non-work income)

# Density Test, or No Bunching

- Manipulation if agents know about the institutional details
  - If schools scoring lower than $w = 50$ on standardized tests get labeled as dysfunctional, we might see many schools to be right above 50

- In this case, we observe **bunching** around the threshold.
  - Agents are "manipulating" treatment assignment around the threshold.
  - Density of $W_i$ is discontinuous at $\bar{W}$

- We would expect that $E[Y_{1i}|W_i = w]$ would be also discontinuous.

- McCrary (2008) suggests a test of the null hypothesis that the density of $W_i$ is continuous at $\bar{W}$.

# Digression: Bunching Analysis (集積点分析)

- Bunching itself is an interesting economic phenomenon. It can be used to analyze a different question.

# Example: Ito and Sallee (2018, REStat)

# Empirical Paper

# Empirical Paper: Health Demand

- "The Effect of Patient Cost Sharing on Utilization, Health, and Risk Protection" by Hitoshi Shigeoka 2014 AER'

# Policy Issue: Medical Expenditure

- Medical expenditures are rising.
  - due to an aging population and coverage expansion
  - acute fiscal challenge to governments!

- Current expenditure on health (to GDP) in 2018 according to OECD Health Statistics 2019
  - U.S.A. (16.9%), Switzerland (12.2%), Germany (11.2%), France (11.2%), Sweden (11.0%), Japan (10.9%)...

- One main strategy is higher patient cost sharing, that is, requiring patients to pay a larger share of the cost of care.

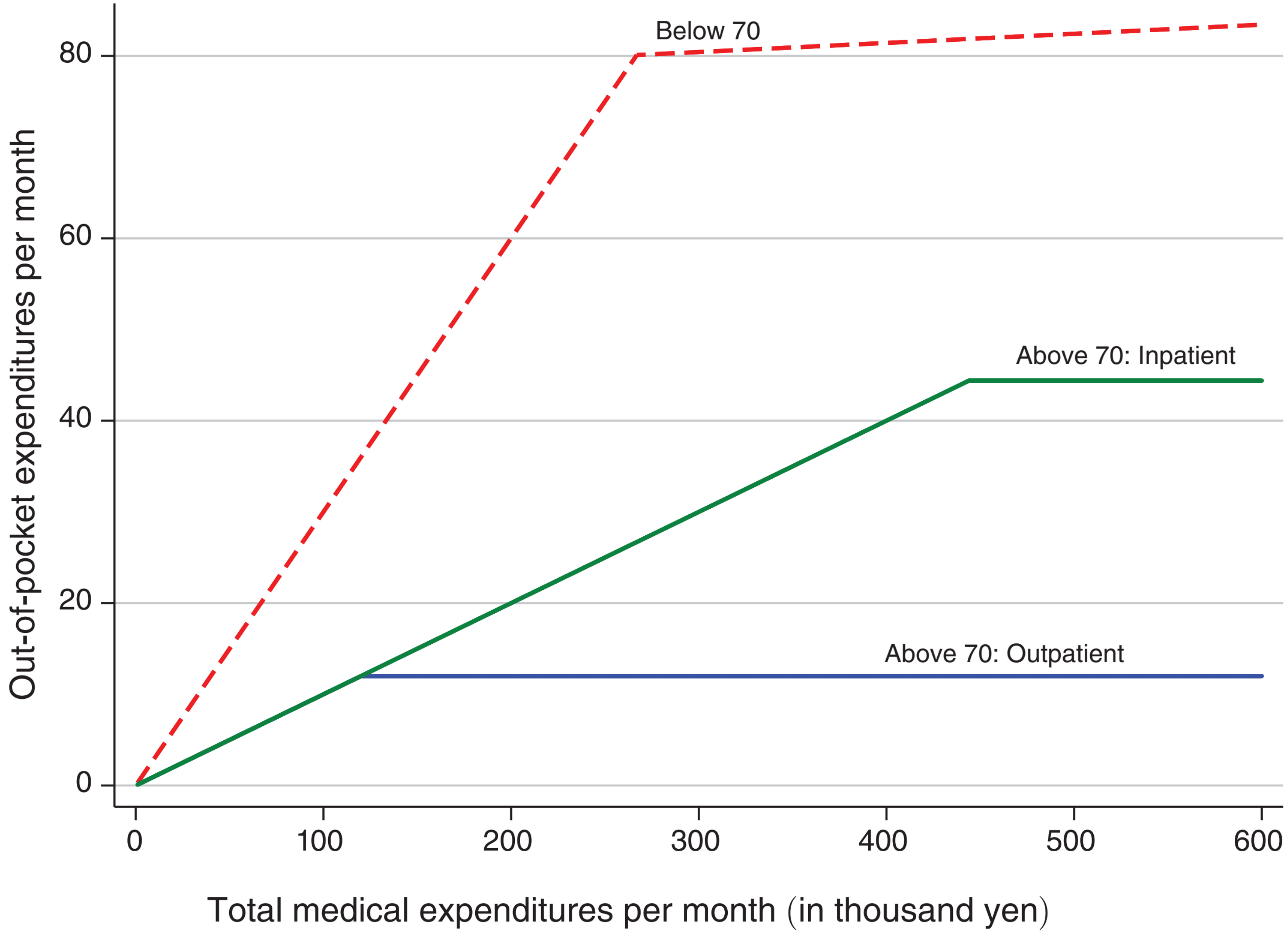
- Question: how does patient cost sharing affect
  - utilization (demand elasticity)?
  - health?
  - risk protection (out-of-pocket expenditures)?

# Background and Cross-sectional Data

- All Japanese citizens are mandatorily covered by health insurance.
- Use a sharp reduction in cost sharing for patients aged over 70 in Japan.
- The sources are the Patient Survey and the Comprehensive Survey of Living Conditions (CSLC). 1984-2008.
- Advantages
  - There are no confounding factors at age 70. We can isolate the effect of patient cost sharing.
  - Medical providers do not have incentive to differentiate prices by the patients' insurance type.
  - We can separate inpatient and outpatient.

# Cost Sharing and Out-of-Pocket Medical Expenditure

- In sum, the proportion is 30% for $< 69$ and 10% for $70 \leq$.
- Out-of-pocket medical expenditure for impatient admissions can reach 27% for a 69-year-old.
- However, for 70, it would be reduced to 8.6%.
- We need to take the stop-loss into account.


Below 70
Above 70: Inpatient
Above 70: Outpatient
Out-of-pocket expenditures per month
Total medical expenditures per month (in thousand yen)
0
20
40
60
80
0
100
200
300
400
500
600

## Estimated Out-of-Pocket Medical Expenditure per Month

| | Out-of-pocket medical expenditure (thousand yen) | | |
|---|---|---|---|
| Type of service | Below 70 (1) | Above 70 (2) | Percent reduction ((1)–(2))/(3) |
| Outpatient visits | 4.0 | 1.1 | 73 |
| Inpatient admissions | 41.7 | 13.0 | 69 |

# Identification Strategy

- Standard RD designs.
- Basic estimation equation for the CSLC is

$$Y_{iat} = f(a) + \beta Post70_{iat} + X'_{iat}\gamma + \varepsilon_{iat}.$$

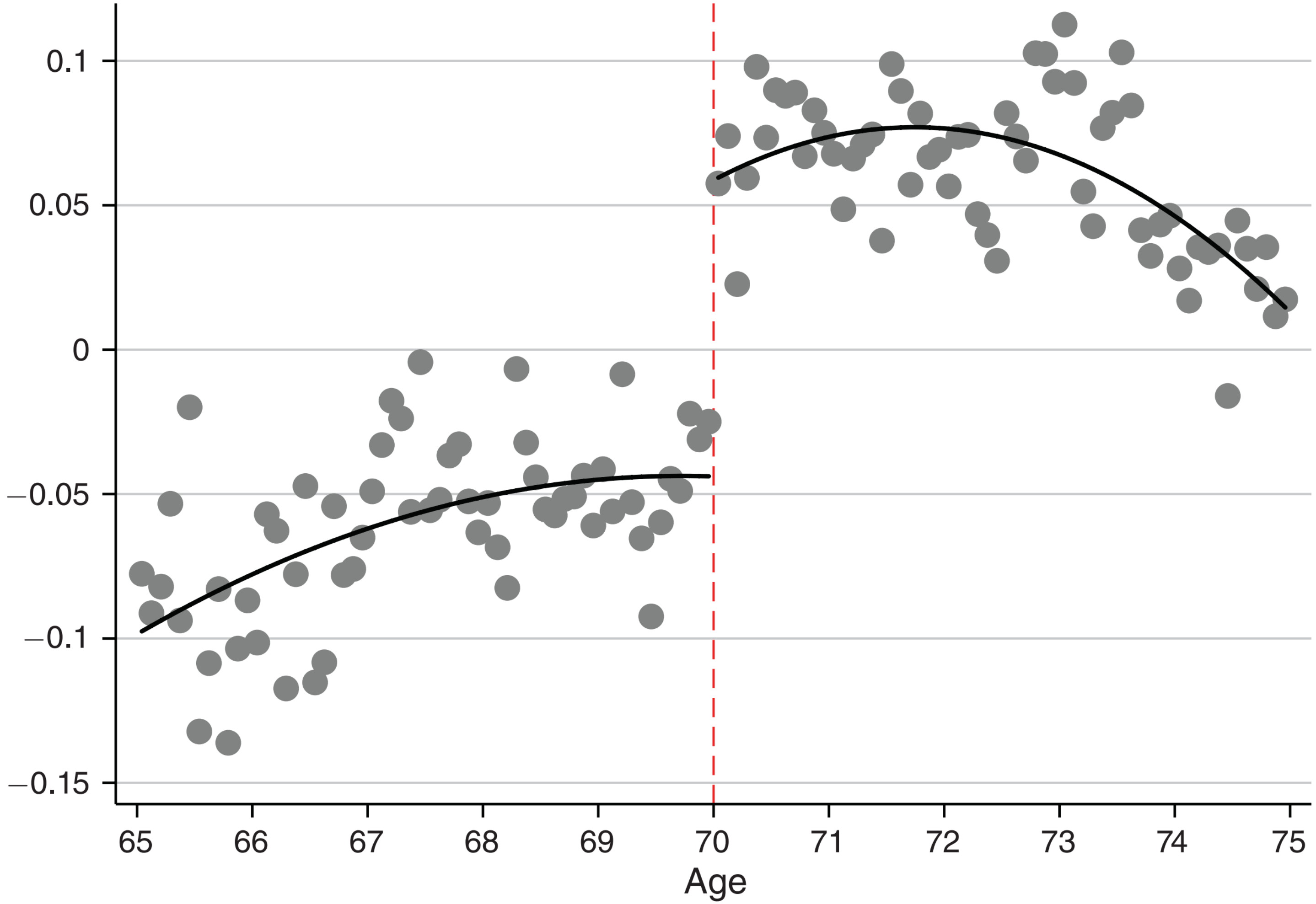
  - $Y_{iat}$: a measure of morbidity or out-of-pocket medical expenditure
  - $f(a)$: a smooth function of age.
  - $X_{iat}$: a set of individual covariates
  - $Post70_{iat}: = 0$ if individual $i$ is over 70.
- Patient Survey/mortality data represents individuals who are present in the medical institutions/deceased.
- As in Card, Dobkin, and Maestas (2004), basic estimation equation for the Patient Survey and mortality data is

$$\log(Y_{at}) = f(a) + \beta Post70_{at} + \mu_{at}.$$

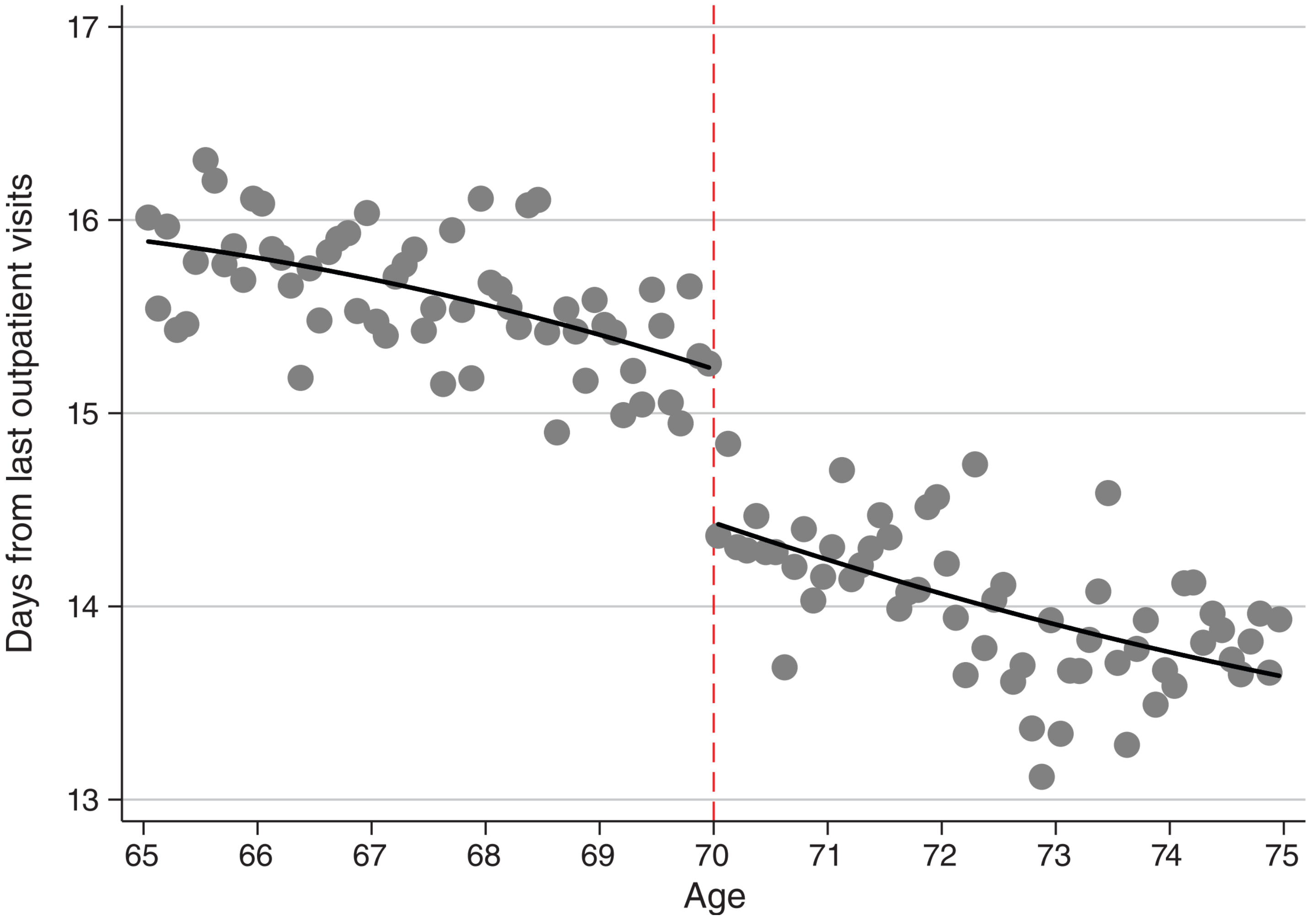
# Results: Outpatient Visits

- 10.3% increase in overall visits. The implied elasticity is $-0.18$.
- Sharp drop in the duration from the last visit by one day.
- The effect is heterogeneous across institutions, genders, and diagnoses.

Panel A. Overall outpatient visits (log scale)



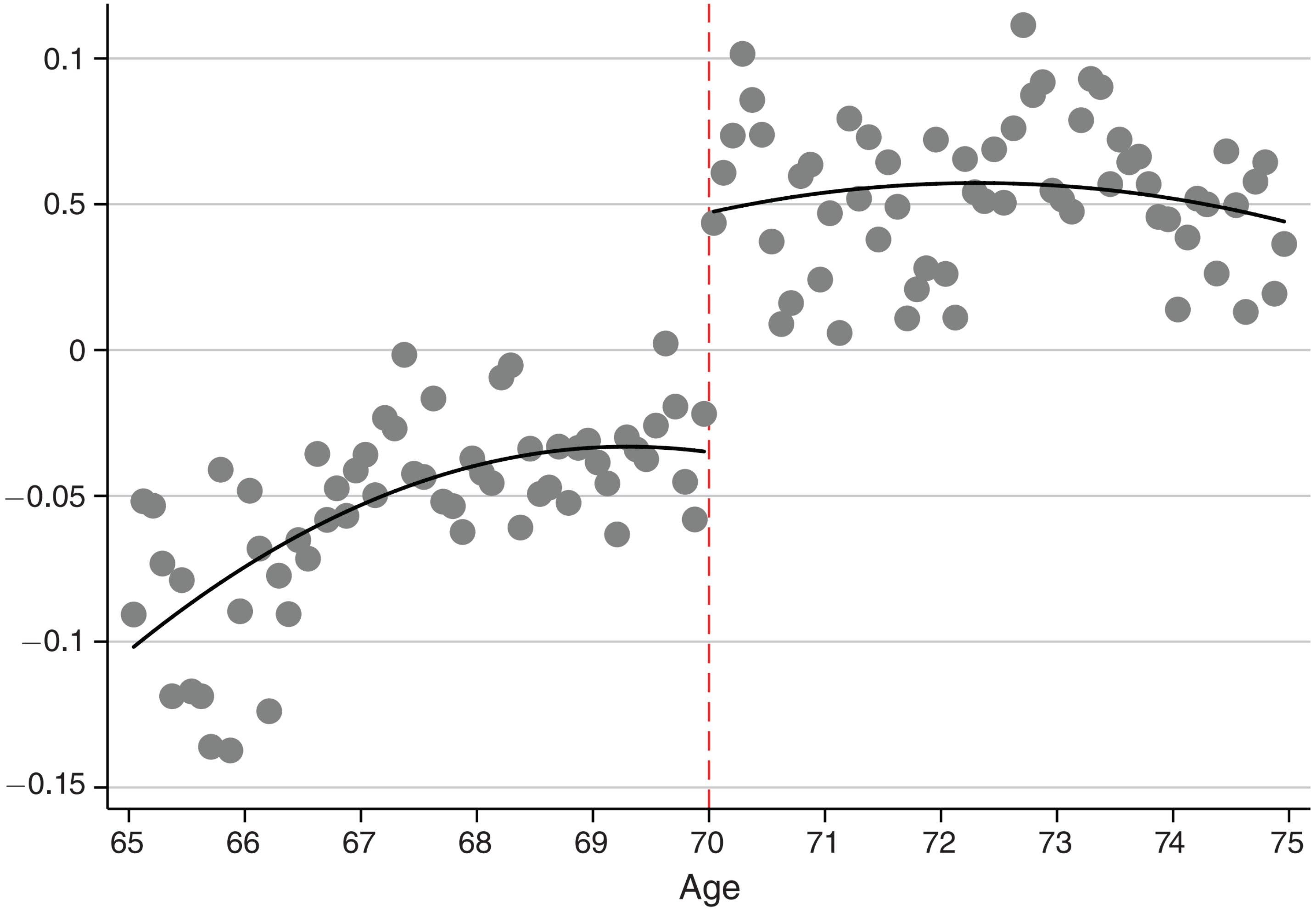
Panel B. Days from last outpatient visit for repeat patients



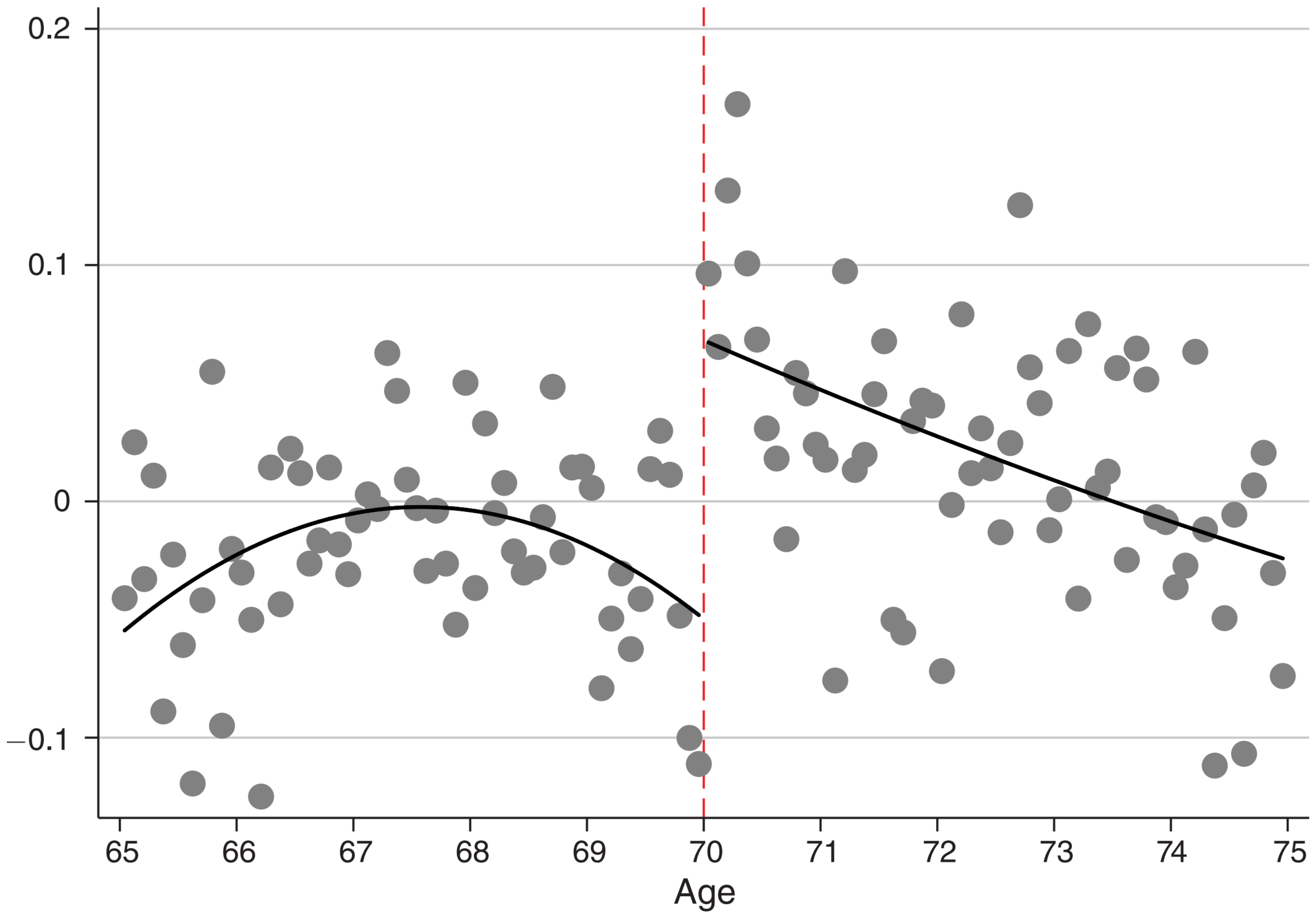
# Results: Inpatient Admissions

- Left: 8.2% increase in overall admissions. The implied elasticity is $-0.16$.
- Right: Surge (increase by 12.0%) in admissions with surgery.
- From robustness checks, the implied elasticity is around $-0.2$.

Panel A. Overall



Panel B. With surgery



# Benefits: Health Outcomes

- We cannot find significant discontinuity in mortality.
- This result is expected because health is stock (Grossman 1972).
- There is no discontinuity in morbidity (self-reported health).
- The available health measures here are limited, so we would underestimate the benefit.

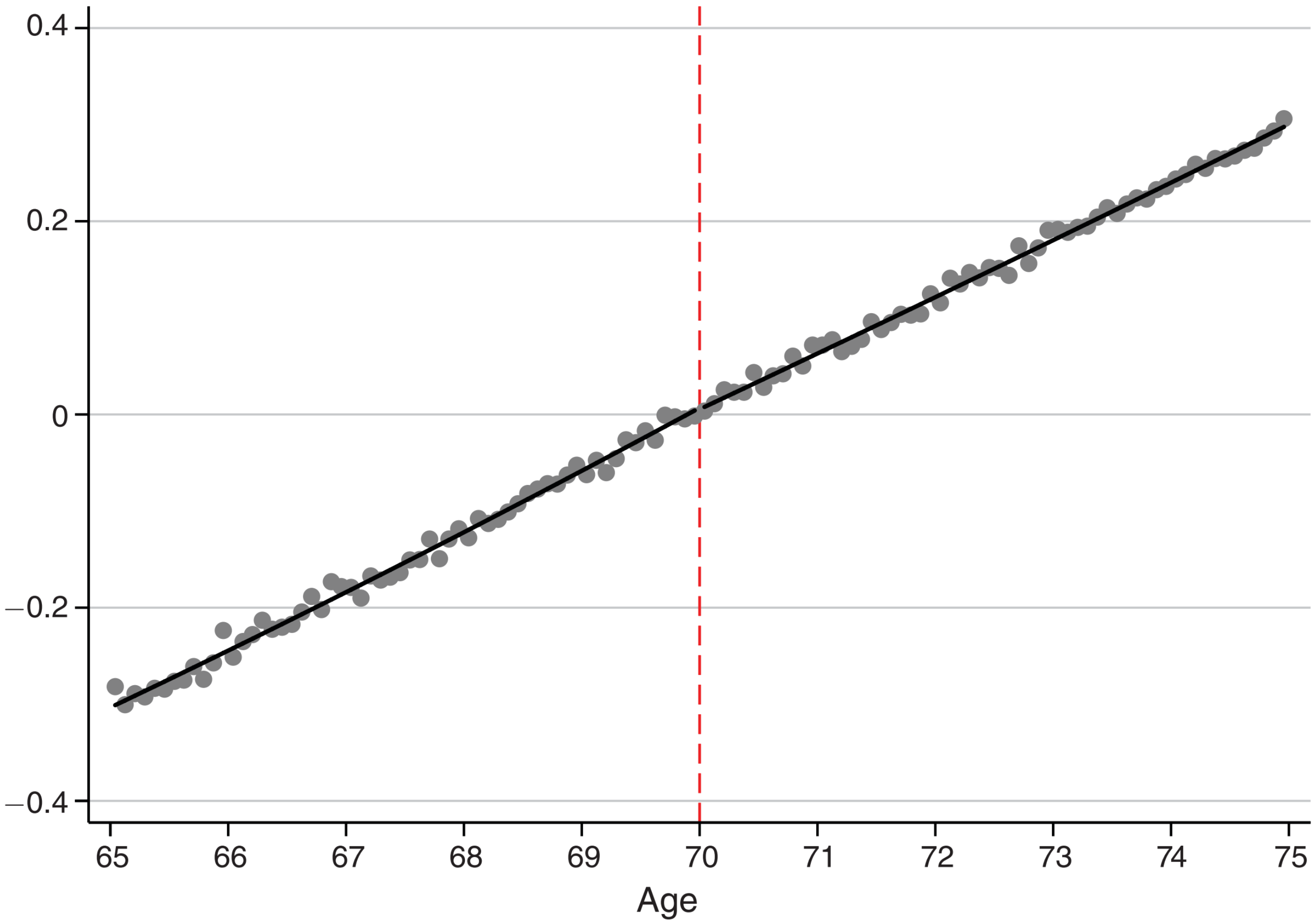
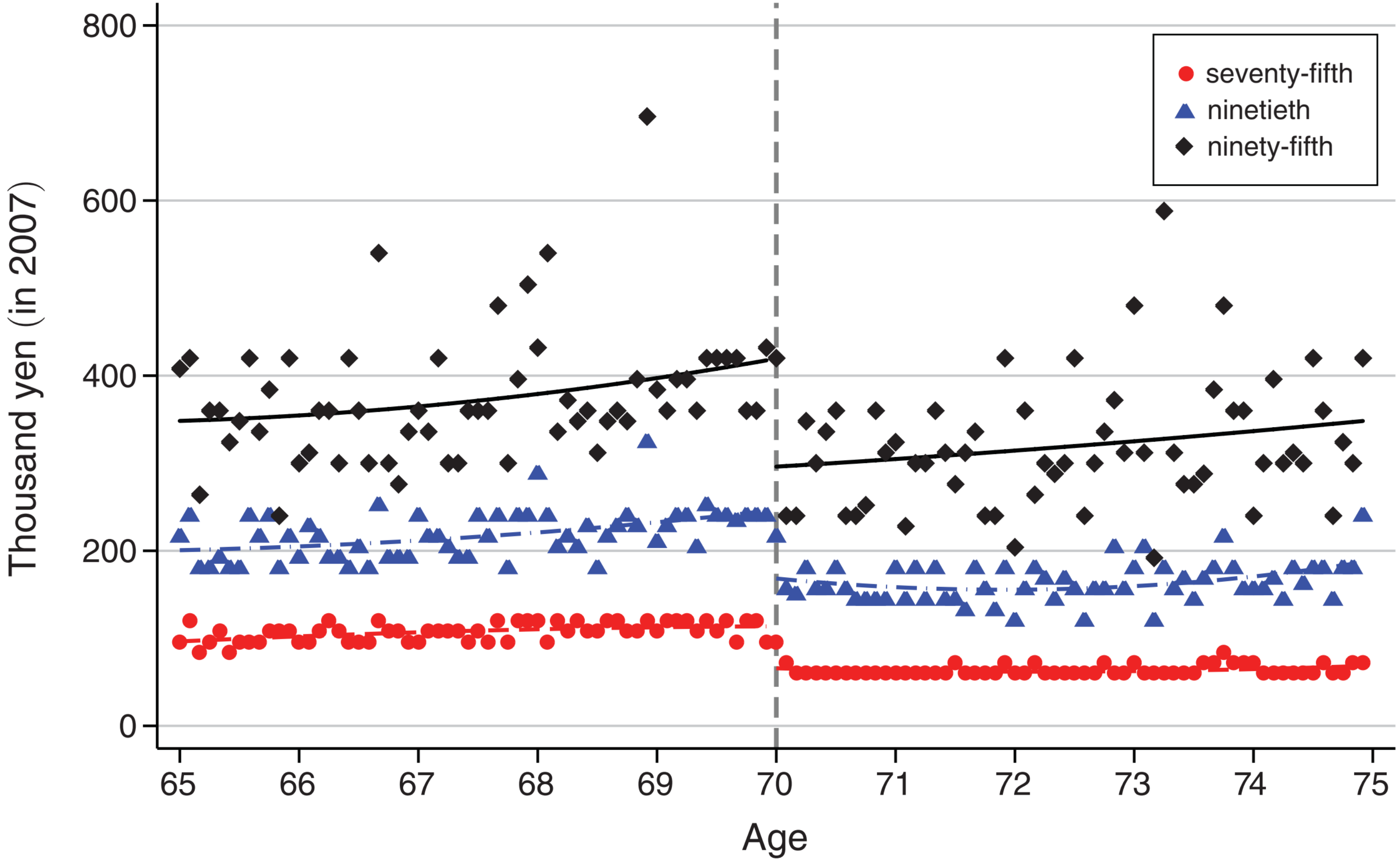



Figure 6. Age Profile of Overall Mortality

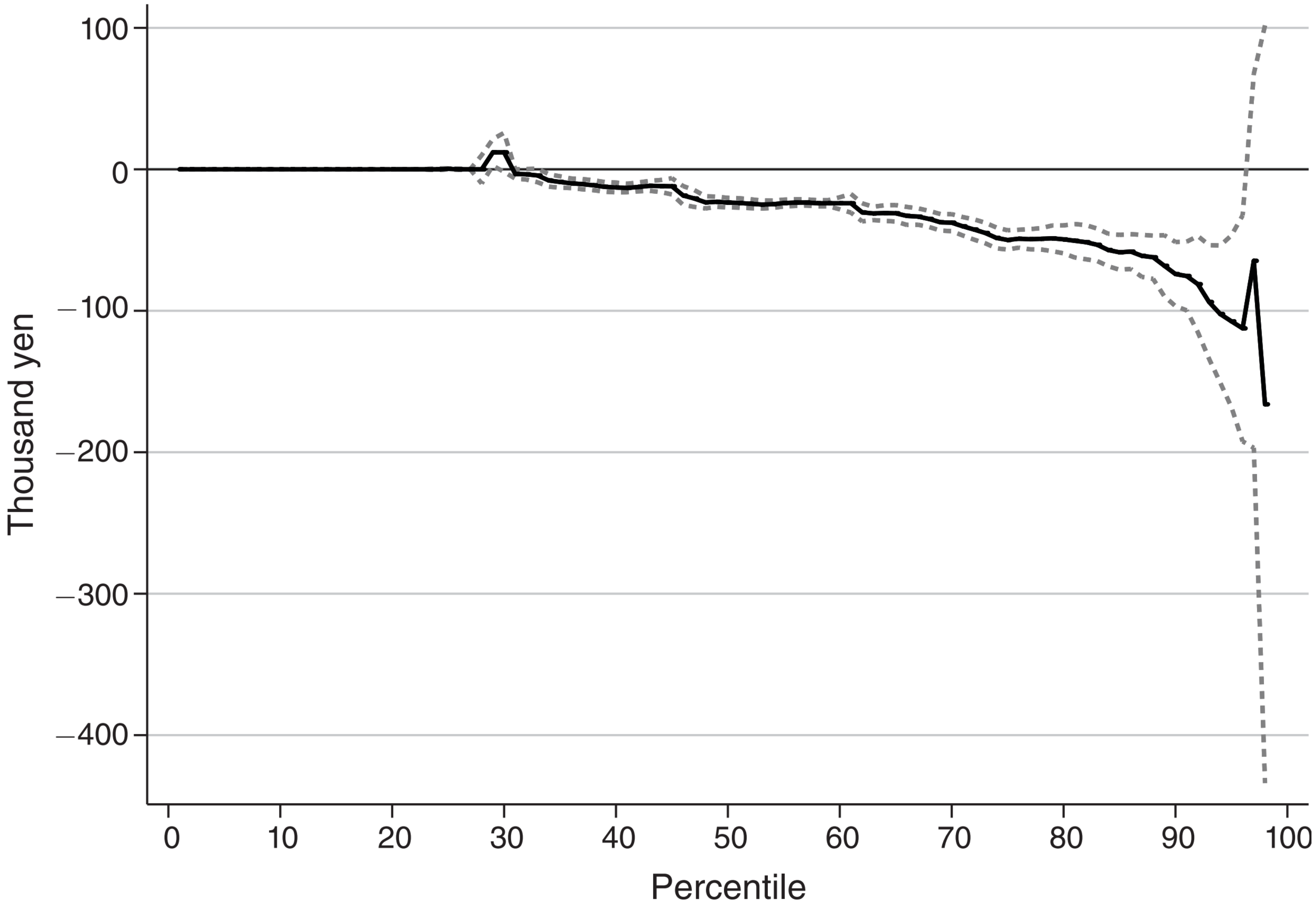
# Benefits: Risk Reduction

- Another benefit is a lower risk of unexpected out-of-pocket medical spending.
- We use a nonparametric estimator for quantile treatment effects.
- Patients at the right tail of the distribution in particular are substantially benefited.

Panel A. At the seventy-fifth, ninetieth, and ninety-fifth percentile



Panel B. RD estimates and each quantile



# Discussion

- Price Elasticities
  - We cannot distinguish own- from cross-price effects.
  - However, for some diagnosis groups, cross-price effects should be nearly zero.
  - The overall effect of the price change for the groups is an approximately 10 percent increase in visits.
- Cost-Benefit Analysis
  - Imposing many assumptions, we speculate that the welfare gain of risk protection from lower patient cost sharing is comparable to the total social cost.
  - We cannot include welfare gains from health improvements.

# Appendix: Formal Analysis

# Formal Identification Analysis

- Key: continuity assumptions: Both $E[Y_{1i}|W_i = w]$ and $E[Y_{0i}|W_i = w]$ are continuous at the threshold $w = \bar{W}$.
  - This is not directly testable assumption (because we cannot observe $Y_{1i}$ below the threshold).
  - Will discuss several validating approaches.
- To see how this works, notice that

$$\begin{aligned} E[Y_i|W_i = w] =& E[Y_{0i}|W_i = w] \\ & + \mathbf{1}\{w \geq \bar{W}\} \left(E[Y_{1i}|W_i = w] - E[Y_{0i}|W_i = w]\right) \end{aligned}$$

- Taking the limit of $w$ to $\bar{W}$ from above and below

$$\lim_{w \uparrow \bar{W}} E[Y_i|W_i = w] = \lim_{w \uparrow \bar{W}} E[Y_{0i}|W_i = w] = E[Y_{0i}|W_i = \bar{W}]$$
$$\lim_{w \downarrow \bar{W}} E[Y_i|W_i = w] = \lim_{w \downarrow \bar{W}} E[Y_{1i}|W_i = w] = E[Y_{1i}|W_i = \bar{W}]$$

- Notice that we use continuity in the second equalities!

- Remember that

$$\lim_{w \uparrow \bar{W}} E[Y_i|W_i = w] = \lim_{w \uparrow \bar{W}} E[Y_{0i}|W_i = w] = E[Y_{0i}|W_i = \bar{W}]$$
$$\lim_{w \downarrow \bar{W}} E[Y_i|W_i = w] = \lim_{w \downarrow \bar{W}} E[Y_{1i}|W_i = w] = E[Y_{1i}|W_i = \bar{W}]$$

- So, we have

$$E[Y_{1i} - Y_{0i}|W_i = \bar{W}] = \lim_{w \downarrow \bar{W}} E[Y_i|W_i = w] - \lim_{w \uparrow \bar{W}} E[Y_i|W_i = w]$$

    - LHS: Average treatment effect at the threshold
    - RHS: We can observe from the data.
        - Conditional expectation near the threshold.