

# **Đánh giá Chính sách**

*Bài giảng 8: Thiết kế Hồi quy Gián đoạn*

Edmund Malesky, Ph.D.

July 11, 2018

Duke University

## Nội dung

---

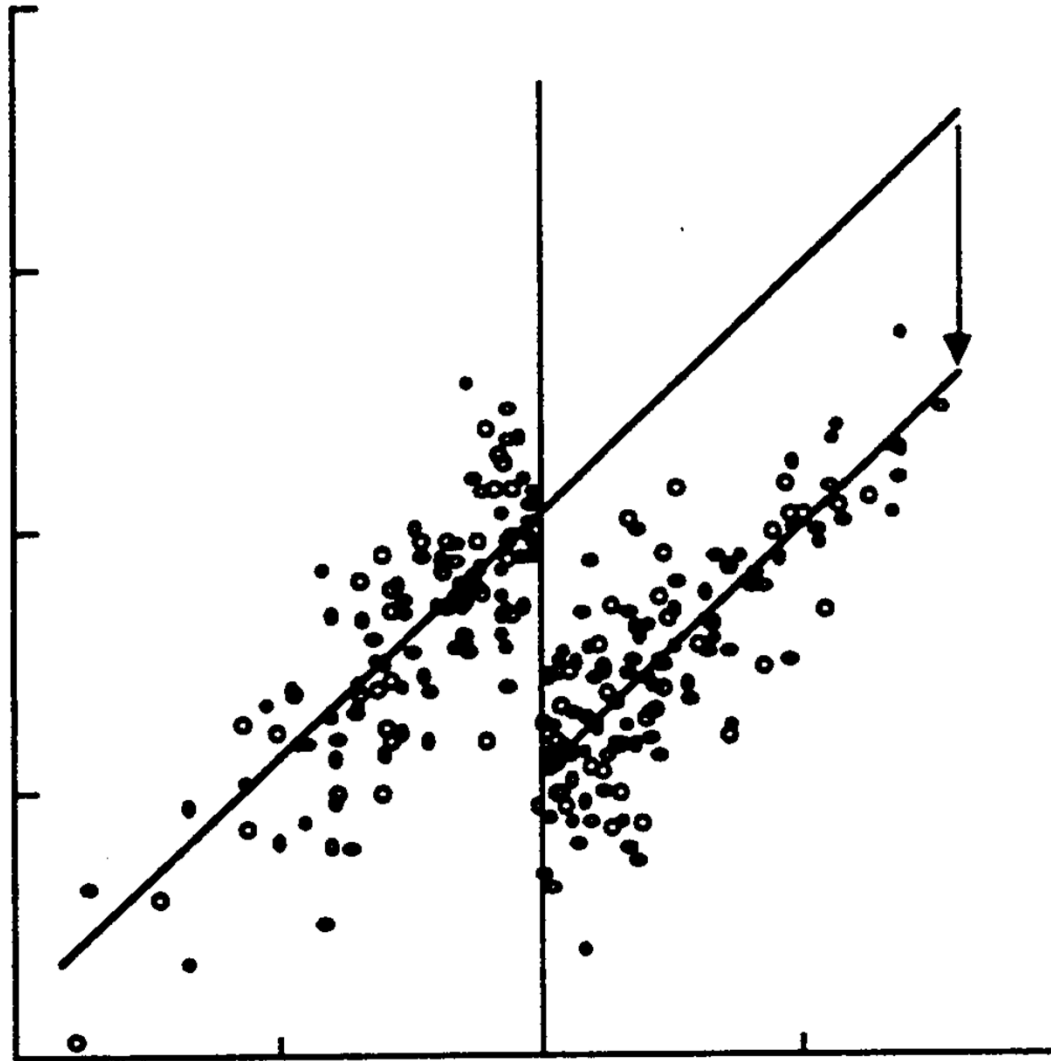
**Tổng quan về RDD**

**Phúc lợi và Việc làm**

**Mua phiếu bầu cử**

## Thiết kế hồi quy gián đoạn (RDD)

---



## Thiết kế hồi quy gián đoạn (RDD)

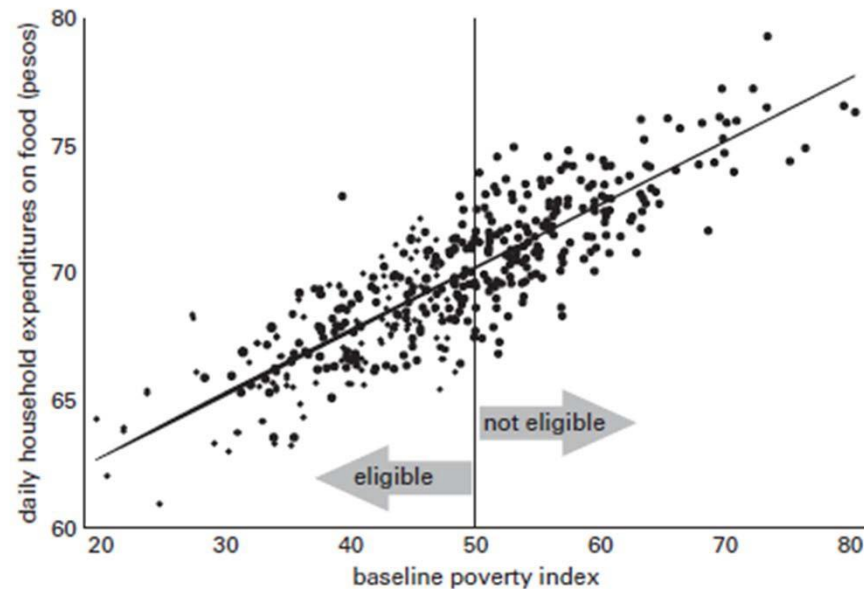
- Nhiều chính sách sử dụng các chỉ số với tiêu chí cụ thể để lựa chọn người thụ hưởng
- 
- Thiết kế RDD xác lập liệu có sự gián đoạn nhảy cóc về kết quả tại tiêu chí lựa chọn
  - Các quan sát tại mức dưới sát và cận trên của ngưỡng tiêu chí thường giống nhau, nhưng chỉ khác về tình trạng thụ hưởng
  - Các quan sát tại mức dưới sát tiêu chí nên không được tham gia chương trình được sử dụng làm nhóm đối chứng cho nhóm đủ tiêu chuẩn tham gia ngay ở cận trên của ngưỡng tiêu chuẩn
  - Đánh giá tác động sử dụng thiết kế RDD yêu cầu:
    - Chỉ số đánh giá liên tục
    - Ngưỡng lựa chọn tham gia được xác lập rõ ràng
  - Có thể ước lượng tác động mà không loại trừ toàn bộ quần thể các quan sát.
  - Vấn đề hiệu lực ngoại vi: ước lượng chỉ có hiệu lực tại ngưỡng tiêu chí được sử dụng để xác lập tình trạng thụ hưởng:
    - Tác động can thiệp trung bình nội tại, không áp dụng cho tất cả quần thể
    - Có thể ứng dụng tốt cho việc quyết định liệu có nên mở rộng chương trình ở mức cận biên

## Cơ chế lựa chọn theo thiết kế RDD

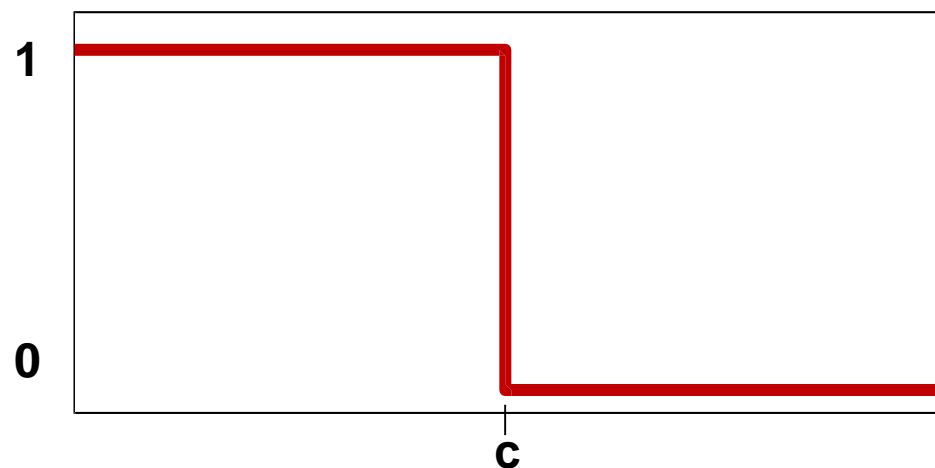
- Việc lựa chọn tham gia chương trình tùy vào việc đạt được tiêu chí cho trước, không tùy tiện
  - Thiết kế RDD sắc vs RDD nhiều/mờ (sharp vs fuzzy)
- Phân bổ vào nhóm tham gia ( $D_i=1$ ) hay đối chứng ( $D_i=0$ ) phụ thuộc vào biến chỉ số  $X$  (forcing variable)
- Được tham gia chương trình nếu như biến chỉ số  $X \leq$  ngưỡng ( $c$ )

$$D_i = \begin{cases} D_i = 1 & \text{if } X_i \leq c \\ D_i = 0 & \text{if } X_i > c \end{cases}$$

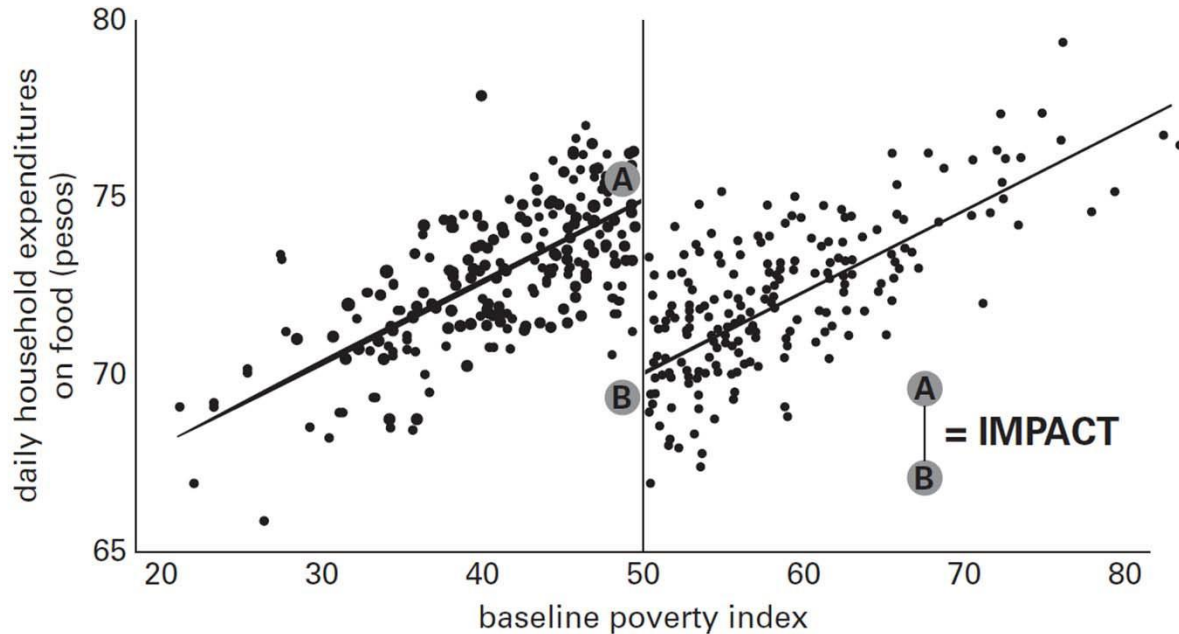
- $X$  thường tương quan với biến kết quả  $Y$  -> Nhóm không tham gia không phải là phản thực hợp lệ với nhóm tham gia



Probability of Treatment



## Ước lượng RDD



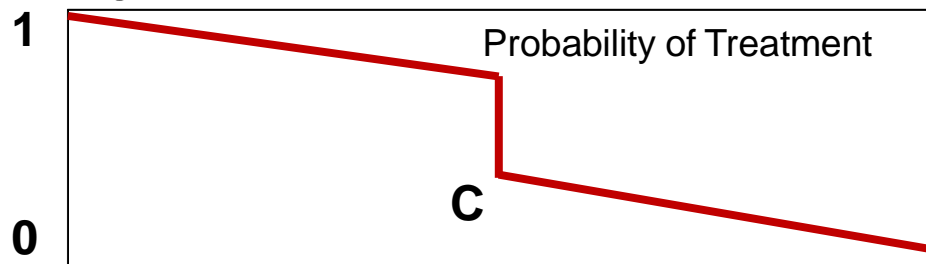
- Không tham gia nhưng gần ngưỡng c: nhóm so sánh để ước lượng phản thực
- Cận ngưỡng: cùng đặc tính với nhóm đối chứng, nhưng được tham gia chương trình
- Nếu quan hệ giữa X và Y là liên tục thì có thể quy kết sự gián đoạn về kết quả là do tình trạng tham gia chương trình:

Source: Figure from

$$y_i = \beta_0 + \beta_1 D_i + \delta(\text{index}_i) + \varepsilon_i$$

## Thiết kế RDD nhiều

- Ví dụ trước: thiết kế RDD sắc
  - Xác suất tham gia chương trình nhảy từ 0 đến 1 khi qua ngưỡng tiêu chuẩn
- Tuy nhiên có nhiều chương trình có ngưỡng tiêu chí không hoàn toàn rõ ràng
- *Thiết kế RDD nhiều*: ngưỡng tiêu chí tạo ra gián đoạn trong xác suất tham gia chương trình



- Đối với quan sát gần ngưỡng tiêu chí, có thể sử dụng biến  $Z$  làm biến công cụ cho tình trạng tham gia:

$$Z_i = \begin{cases} 1 & \text{if } X_i \leq c \\ 0 & \text{if } X_i > c \end{cases}$$

- Có thể ước lượng tác động can thiệp đối với nhóm tuân thủ (tình trạng tham gia  $D$  phụ thuộc vào  $Z$ )

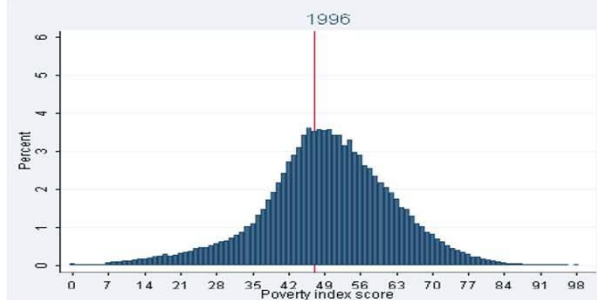
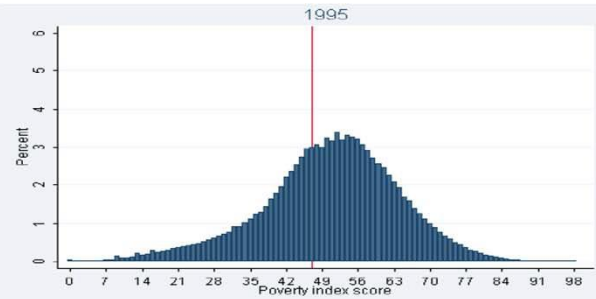
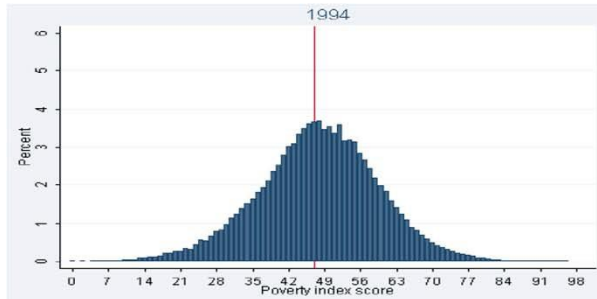
## Tính tin cậy của ước lượng RDD

---

- Bắt buộc phải kiểm tra tính tin cậy của ước lượng RDD
- 4 vấn đề chính liên quan đến tác động lựa chọn (sorting), cân bằng, tính vững đối với cấu trúc hàm, và kiểm định đối chứng
- *Lựa chọn (Sorting)*: Xác định liệu có chuyện các quan sát lựa chọn xung quanh ngưỡng tiêu chí tam gia chương trình
- *Cân bằng*: Kiểm tra liệu các biến kiểm soát khác có sự khác biệt lớn tại ngưỡng tiêu chí
- *Tính vững*: Đảm bảo các ước lượng không phụ thuộc vào cấu trúc hàm
- *Kiểm định đối chứng giả (placebo test)*: Để chứng minh là không có sự gián đoạn tại các ngưỡng tiêu chí giả lập khác
- Chúng ta sẽ xem xét các ví dụ sau



## Lựa chọn

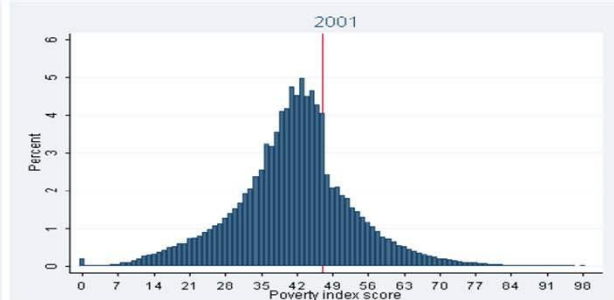
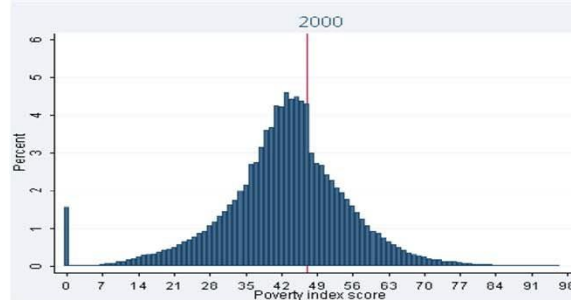
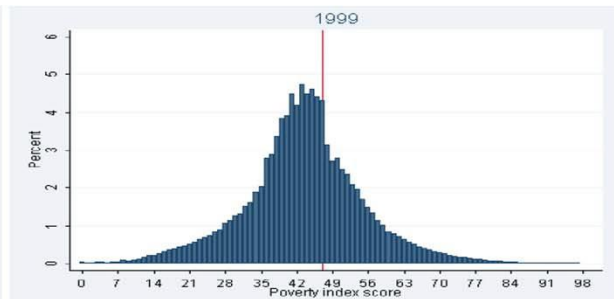
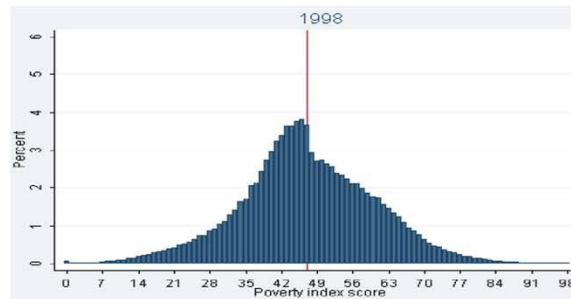
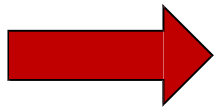


- Colombia: bắt đầu từ đầu thập kỷ 1990s, áp dụng chỉ số nghèo để tham gia chương trình xã hội

- Các thông tin về nhà ở, nhân khẩu học, việc làm

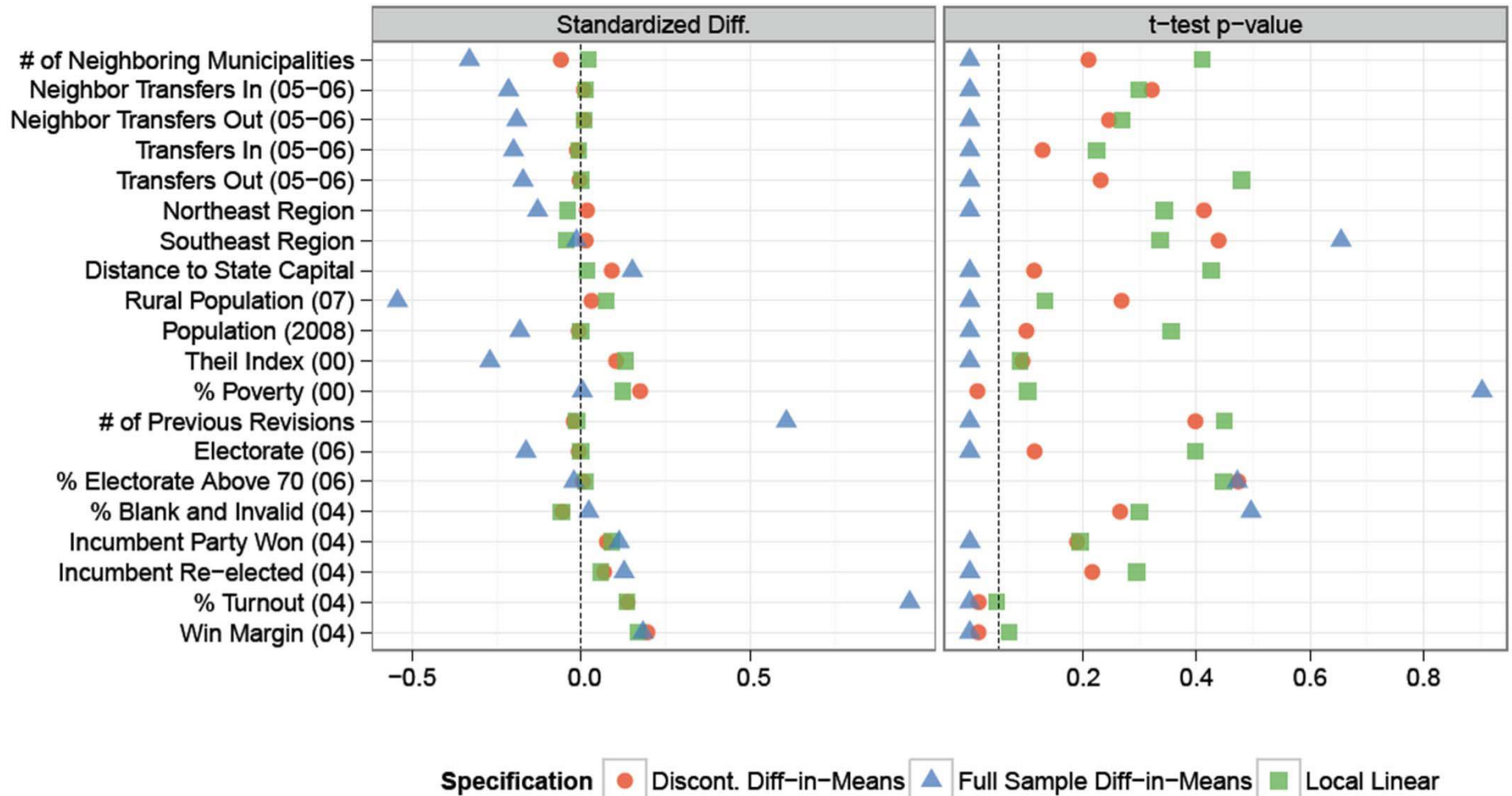
- $\leq 47 \rightarrow$  bảo hiểm y tế

- Thuật toán để tính chỉ số nghèo được cấp cho quan chức địa phương vào 7/1997



Source: Camacho & Conover, 2010.

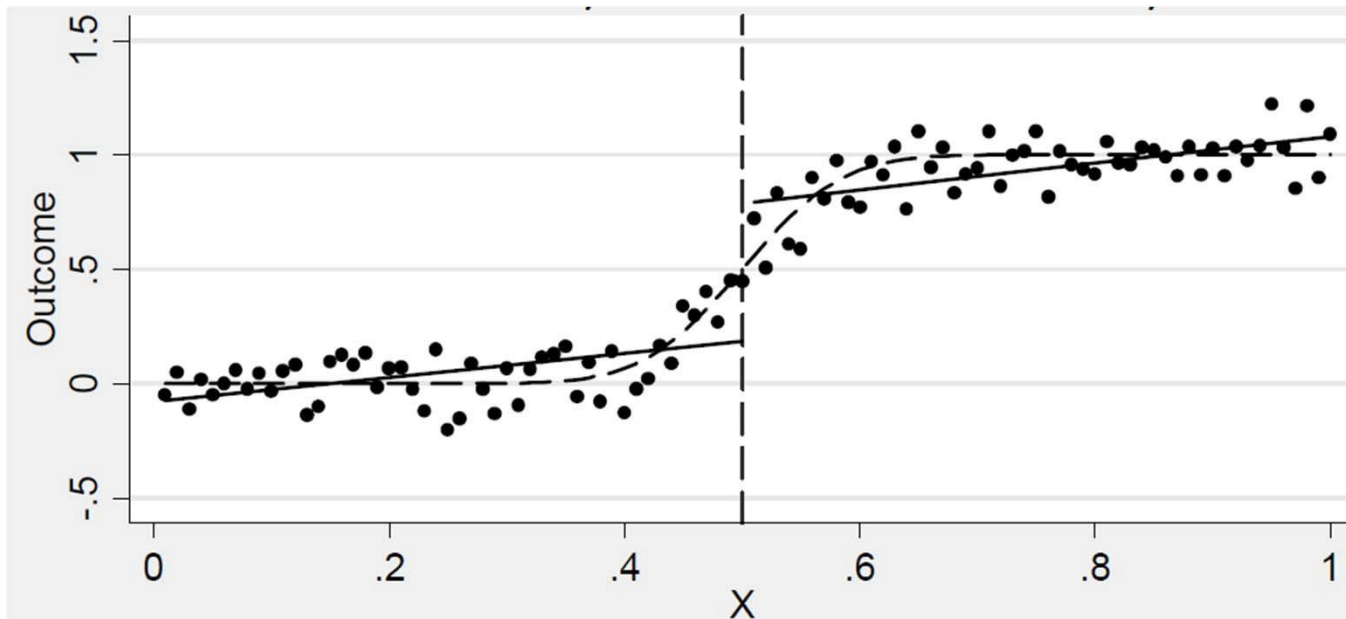
# Cân bằng

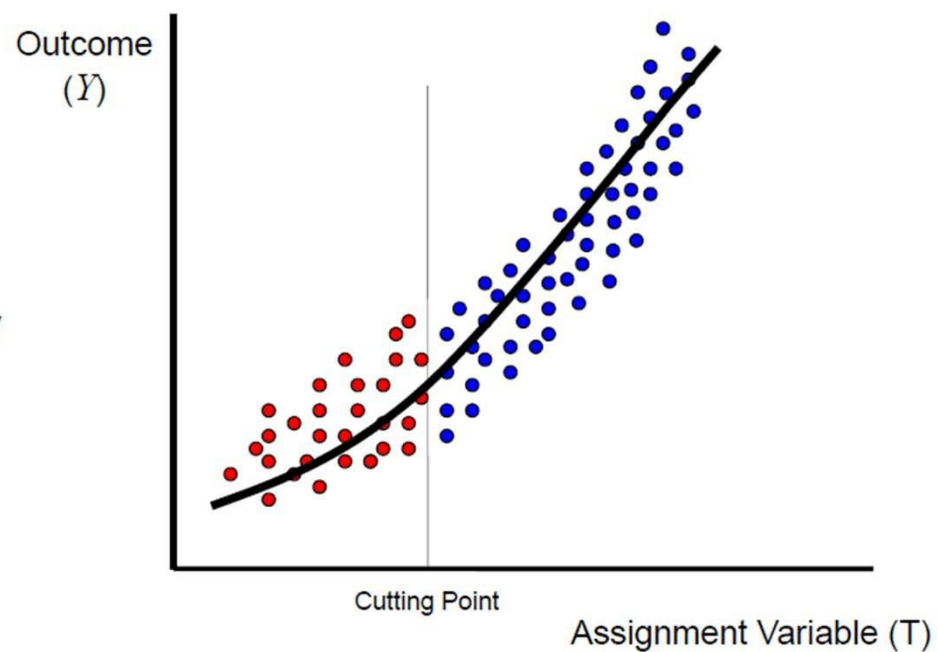
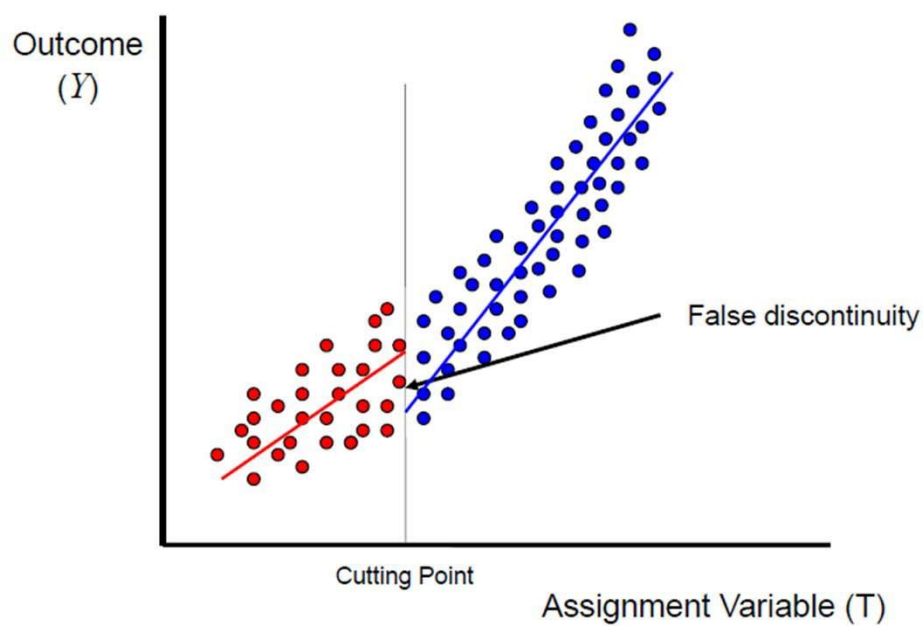


Source: Hidalgo & Nichter, 2012.

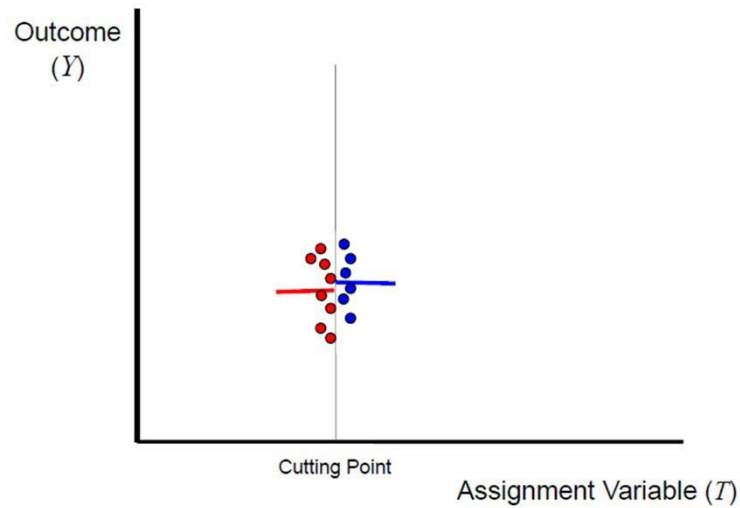
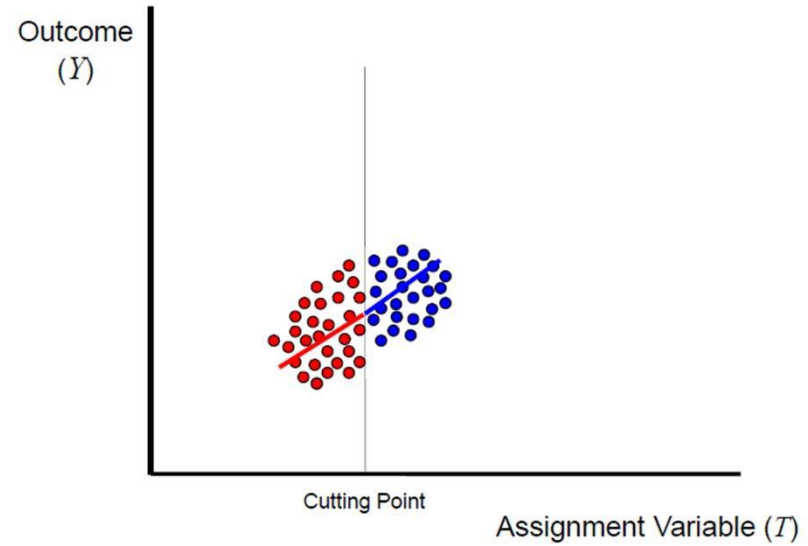
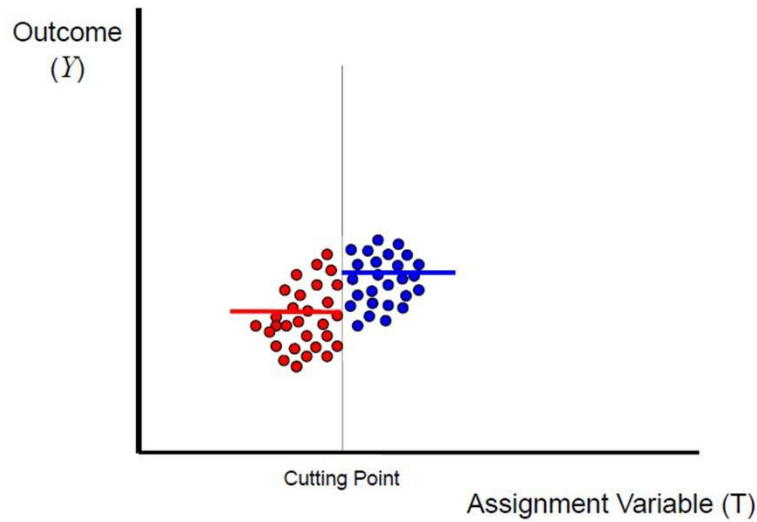
## Tính vững

- Đảm bảo các ước lượng không nhạy cảm đối với các cấu trúc hàm khác nhau
- Ví dụ tính phi tuyến có thể bị lẫn lộn với gián đoạn
- Thử nghiệm các cấu trúc hàm đa dạng để chứng minh tính vững





# Tính vững đối với khác biệt về trung bình so tuyến tính nội tại (local linear); độ rộng của khoảng ước lượng (bandwidth)



# Kiểm định đối chứng giả

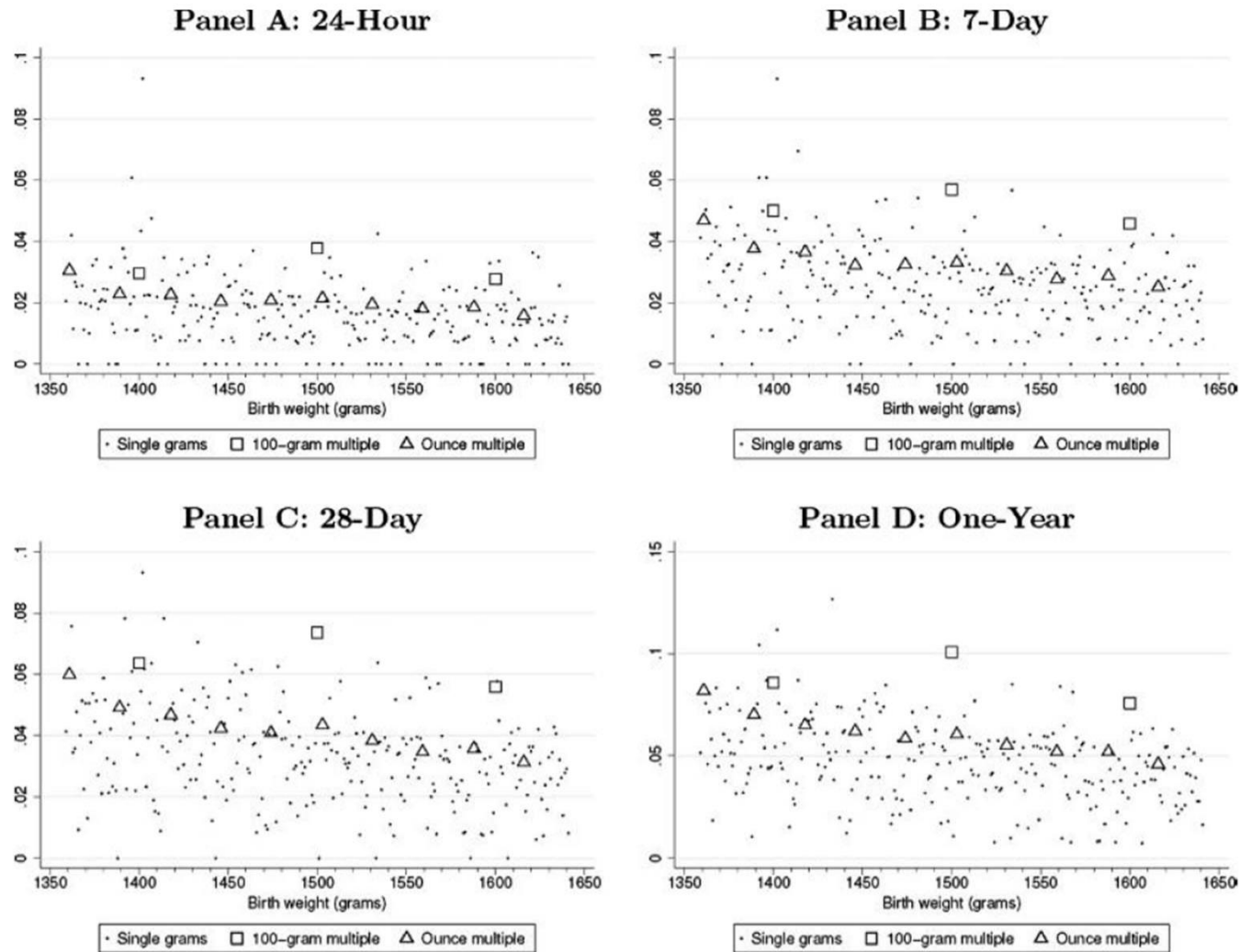


FIGURE I

Means of Mortality Rates

Source: Barecca et al, 2011.

## Nội dung

---

**Tổng quan về RDD**

**Phúc lợi và Việc làm**

**Mua phiếu bầu cử**

## Tác động của phúc lợi lên việc làm

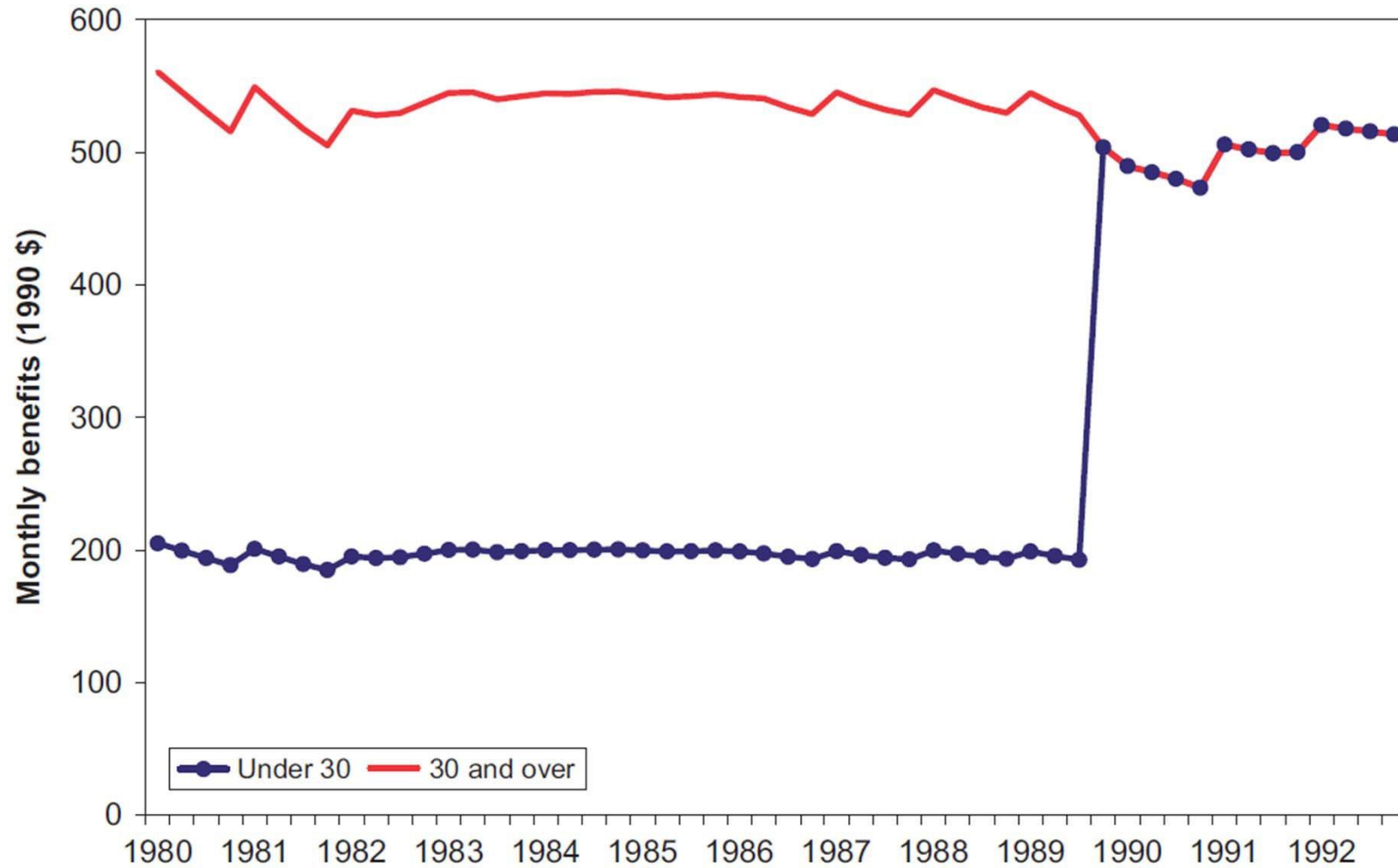
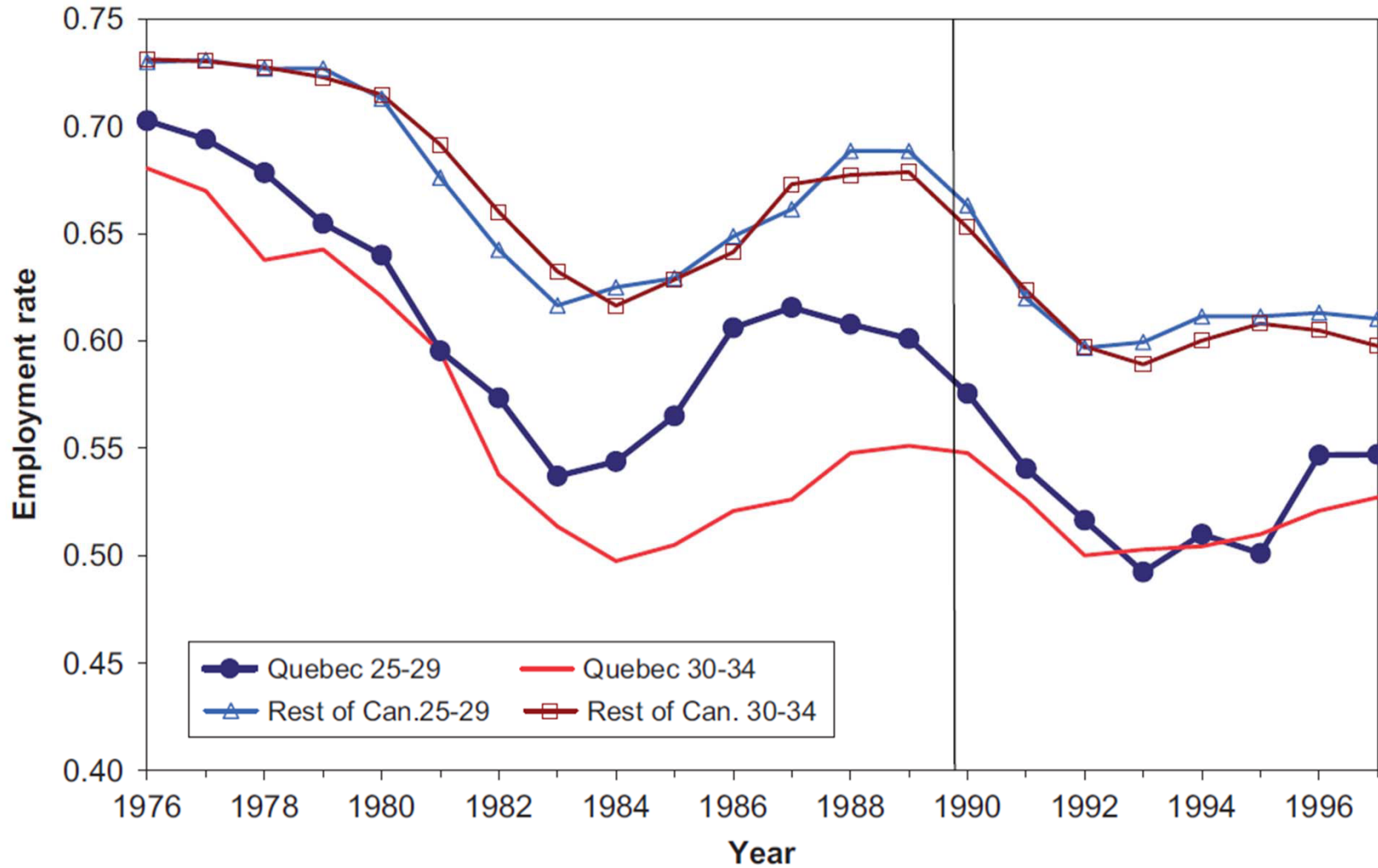


Fig. 1. Social assistance benefits, single individual.

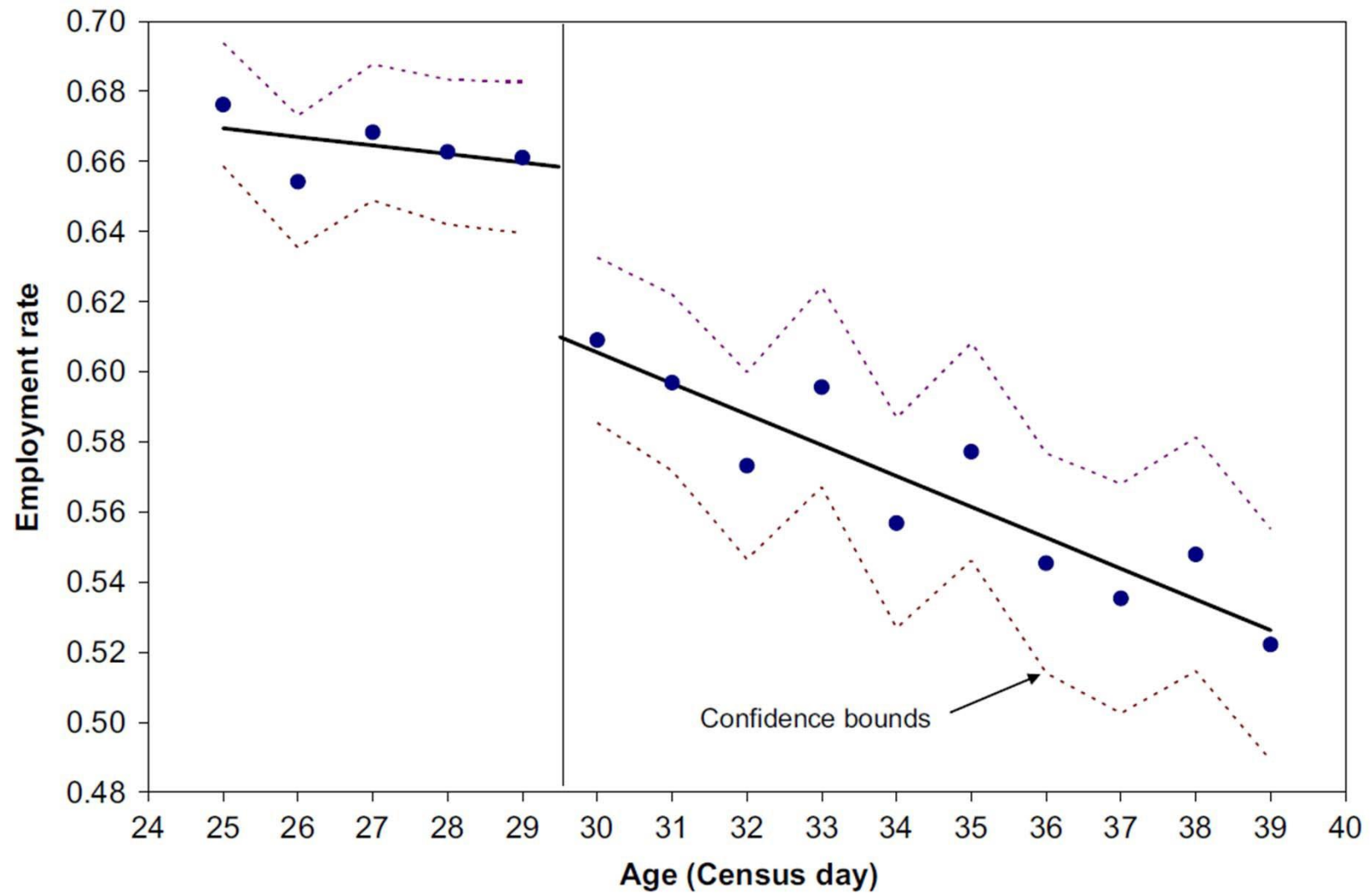


## Tác động của phúc lợi lên việc làm



Source: Lemieux & Milligan, 2008.

## Tác động của phúc lợi lên việc làm



Source: Lemieux & Milligan, 2008.

# Tác động của phúc lợi lên việc làm

## Ước lượng RD

$$Y_{ia} = \beta_0 + \beta_1 TREAT_{ia} + \delta(a) + \varepsilon_{ia}, \quad TREAT_{ia} = \begin{cases} 0 & \text{if } a < 30, \\ 1 & \text{if } a \geq 30. \end{cases}$$

Regression discontinuity estimates of the effect of higher social assistance benefits on labour supply in Quebec, 1986

Specification for age	Empl. rate last year	Empl. rate at Census	Difference in empl. rate	Weekly hours
<i>Mean of the dependent variable</i>				
	0.562	0.618	0.056	24.39
<i>Regression discontinuity estimates</i>				
Linear	-0.045 <sup>***</sup> (0.012)	-0.041 <sup>***</sup> (0.012)	-0.029 <sup>**</sup> (0.011)	-1.45 <sup>**</sup> (0.54)
Quadratic	-0.048 <sup>***</sup> (0.013)	-0.051 <sup>***</sup> (0.012)	-0.031 <sup>**</sup> (0.012)	-1.75 <sup>**</sup> (0.61)
Cubic	-0.043 <sup>**</sup> (0.018)	-0.048 <sup>***</sup> (0.014)	-0.030 <sup>**</sup> (0.013)	-1.47 <sup>*</sup> (0.70)
Linear spline	-0.047 <sup>***</sup> (0.013)	-0.049 <sup>***</sup> (0.011)	-0.032 <sup>**</sup> (0.013)	-1.72 <sup>***</sup> (0.55)
Quadratic spline	-0.038 (0.024)	-0.056 <sup>**</sup> (0.018)	-0.035 <sup>*</sup> (0.016)	-1.66 (0.94)

*Note:*

<sup>\*\*\*</sup>Indicate statistical significance at the 1% level.

<sup>\*\*</sup>For the 5% level.

<sup>\*</sup>For the 10% level.

# Tác động của phúc lợi lên việc làm

## Tính vững đối với các khoảng ước lượng khác nhau

Linear spline regression discontinuity estimates with different age windows in Quebec, 1986

Window width	Empl. rate last year	Empl. rate at Census	Difference in empl. rate	Weekly hours
All ages 25–39	–0.047 <sup>***</sup> (0.013)	–0.049 <sup>***</sup> (0.011)	–0.032 <sup>**</sup> (0.013)	–1.72 <sup>***</sup> (0.55)
±5 years	–0.056 <sup>***</sup> (0.014)	–0.046 <sup>**</sup> (0.014)	–0.037 <sup>**</sup> (0.015)	–1.49 <sup>**</sup> (0.66)
±4 years	–0.042 <sup>**</sup> (0.013)	–0.057 <sup>**</sup> (0.015)	–0.038 <sup>**</sup> (0.010)	–2.09 <sup>**</sup> (0.62)
±3 years	–0.050 <sup>*</sup> (0.014)	–0.039 <sup>**</sup> (0.006)	–0.034 <sup>*</sup> (0.012)	–1.37 <sup>*</sup> (0.34)
±2 years	–0.033 (–)	–0.045 (–)	–0.044 (–)	–1.60 (–)

*Note:*

<sup>\*\*\*</sup> Indicate statistical significance at the 1% level.

<sup>\*\*</sup> For the 5% level.

<sup>\*</sup> For the 10% level.

# Tác động của phúc lợi lên việc làm

## Kiểm định giả mạo (Falsification Tests)

Falsification test: comparing labour supply response in Quebec and rest of Canada in 1986 and 1991

Specification for age	Quebec, 1986	Rest of Canada, 1986	Quebec, 1991	Rest of Canada, 1991
<i>Regression discontinuity estimates: employment rate on Census week</i>				
Linear	-0.041*** (0.012)	-0.013** (0.006)	0.041* (0.022)	0.005 (0.011)
Quadratic	-0.051*** (0.012)	-0.013* (0.007)	0.012 (0.023)	-0.017*** (0.006)
Cubic	-0.048*** (0.014)	-0.009 (0.007)	0.037** (0.015)	-0.016** (0.007)
Linear spline	-0.049*** (0.011)	-0.014* (0.006)	0.010 (0.017)	-0.010 (0.007)
Quadratic spline	-0.056** (0.018)	-0.007 (0.010)	0.042* (0.022)	-0.007 (0.007)
<i>Regression discontinuity estimates: difference in employment rate</i>				
Linear	-0.029** (0.011)	-0.009 (0.007)	0.022* (0.011)	-0.007 (0.006)
Quadratic	-0.031** (0.012)	-0.006 (0.007)	0.022 (0.013)	-0.005 (0.006)
Cubic	-0.030** (0.013)	-0.004 (0.006)	0.020 (0.014)	-0.002 (0.006)
Linear spline	-0.032** (0.013)	-0.004 (0.008)	0.021 (0.014)	-0.003 (0.006)
Quadratic spline	-0.035* (0.016)	0.001 (0.009)	0.012 (0.016)	-0.005 (0.008)

Note:

\*\*\* Indicate statistical significance at the 1% level.

\*\* For the 5% level.

\* For the 10% level.