

La Evaluación de Impacto en la Inversión Pública: Desafíos y Oportunidades para el Chile Actual

Matias D. Cattaneo

Princeton University

May 23, 2020

Overview

- Inferencia Causal, Evaluación de Impacto, RD.
- RD Plots (Gráficos) y Falsificación de Diseño.
- Estimación e Inferencia.
 - ▶ Local Polynomial Approach.
 - ▶ Local Randomization Approach.

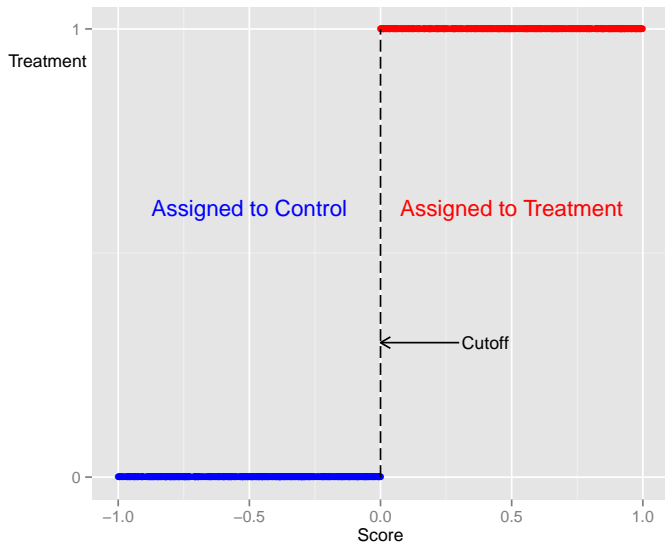
Evidencia Causal & Evaluación de Impacto

- Objetivo: conocer el efecto de tratamiento de una política o intervención.
- Si el tratamiento es asignado aleatoriamente, es fácil estimar efectos causales.
- Si el tratamiento no es asignado aleatoriamente, entonces debemos usar una estrategia de evaluación.
- **Diseño de Regression Discontinua (RD).**
 - ▶ Simple y objetiva. Requiere poco información adicional.
 - ▶ Puede ser vista como un experimento controlado “local”.
 - ▶ Es fácil de falsificar, y fácil de interpretar.
 - ▶ *Cuidado*: puede ser muy local!

Diseño de Regression Discontinua

- Tres características: puntaje (score), punto de corte (cutoff), tratamiento (treatment).
 - ▶ Cada unidad de análisis recibe un puntaje.
 - ▶ Un tratamiento es asignado basado en el puntaje y punto de corte.
 - ▶ El tratamiento es ofrecido si la unidad tiene un puntaje igual o mayor al punto de corte.
 - ▶ El tratamiento no es ofrecido si la unidad tiene un puntaje igual o menor al punto de corte.
 - ▶ Bajo ciertas supuestos, un cambio abrupto en la probabilidad de asignación del tratamiento permite aprender el efecto de dicho tratamiento.

Asignación de Tratamiento: Sharp RD



Diseño de Regression Discontinua

- n unidades, $i = 1, 2, \dots, n$.
- Puntaje es X_i , indicador de tratamiento $T_i = \mathbf{1}(X_i \geq \bar{x})$.
- Cada unidad tiene dos *potential outcomes*:

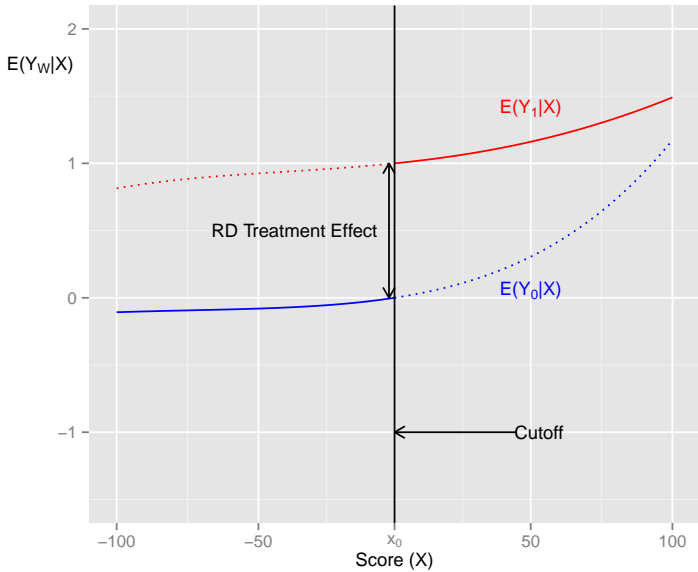
$Y_i(1)$: resultado si unidad i cuando tuvo tratamiento

$Y_i(0)$: resultado si unidad i cuando no tuvo tratamiento

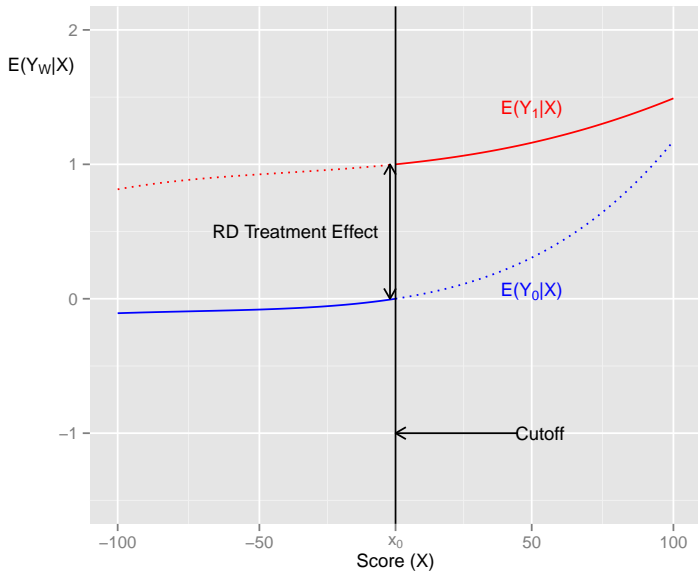
- Variable de resultado *observada* es:

$$Y_i = \begin{cases} Y_i(0) & \text{if } X_i < \bar{x}, \\ Y_i(1) & \text{if } X_i \geq \bar{x}. \end{cases}$$

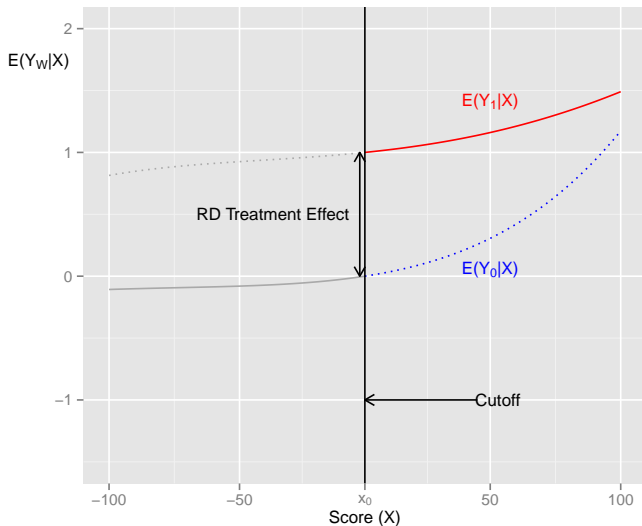
- Problema fundamental de inferencia causal.



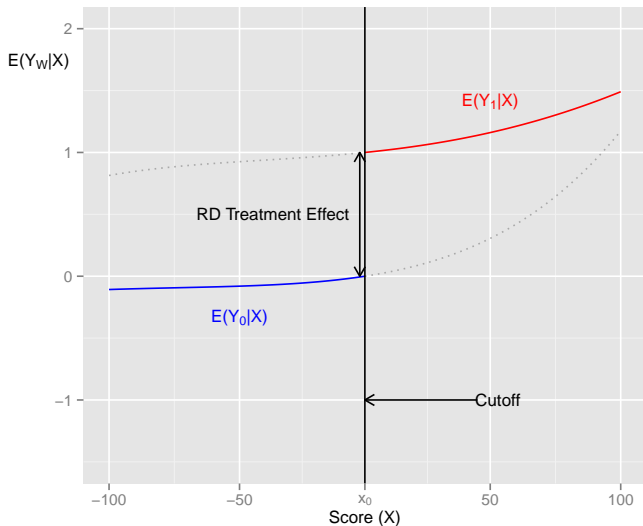
$$\tau_{\text{SRD}} = \mathbb{E}[Y_i(1) - Y_i(0) | X_i = \bar{x}]$$



$$\tau_{\text{SRD}} = \underbrace{\mathbb{E}[Y_i(1) - Y_i(0)|X_i = \bar{x}]}_{\text{Unobservable}} = \lim_{x \downarrow \bar{x}} \mathbb{E}[Y_i|X_i = x] - \lim_{x \uparrow \bar{x}} \mathbb{E}[Y_i|X_i = x]$$



$$\tau_{\text{SRD}} = \underbrace{\mathbb{E}[Y_i(1) - Y_i(0)|X_i = \bar{x}]}_{\text{Unobservable}} = \underbrace{\lim_{x \downarrow \bar{x}} \mathbb{E}[Y_i|X_i = x]}_{\text{Observable}} - \underbrace{\lim_{x \uparrow \bar{x}} \mathbb{E}[Y_i|X_i = x]}_{\text{Observable}}$$



Experimentos Controlados

- **Notación:** $(Y_i(0), Y_i(1), X_i)$, $i = 1, 2, \dots, n$.
- **Tratamiento:** $T_i \in \{0, 1\}$, T_i independent of $(Y_i(0), Y_i(1), X_i)$.
- **Datos:** (Y_i, T_i, X_i) , $i = 1, 2, \dots, n$, with

$$Y_i = \begin{cases} Y_i(0) & \text{if } T_i = 0 \\ Y_i(1) & \text{if } T_i = 1 \end{cases}$$

- **Average Treatment Effect:**

$$\tau_{\text{ATE}} = \mathbb{E}[Y_i(1) - Y_i(0)] = \mathbb{E}[Y_i|T = 1] - \mathbb{E}[Y_i|T = 0]$$

- **Diseño experimental.**

Sharp RD Design

- **Notación:** $(Y_i(0), Y_i(1), X_i)$, $i = 1, 2, \dots, n$, X_i continuous
- **Tratamiento:** $T_i \in \{0, 1\}$, $T_i = \mathbf{1}(X_i \geq \bar{x})$.
- **Datos:** (Y_i, T_i, X_i) , $i = 1, 2, \dots, n$, with

$$Y_i = \begin{cases} Y_i(0) & \text{if } T_i = 0 \\ Y_i(1) & \text{if } T_i = 1 \end{cases}$$

- **Average Treatment Effect at the cutoff:**

$$\tau_{\text{SRD}} = \mathbb{E}[Y_i(1) - Y_i(0) | X_i = \bar{x}] = \lim_{x \downarrow \bar{x}} \mathbb{E}[Y_i | X_i = x] - \lim_{x \uparrow \bar{x}} \mathbb{E}[Y_i | X_i = x]$$

- **Diseño Quasi-experimental:** “local randomization”.

RD Packages

<https://sites.google.com/site/rdpackages>

- **rdr** package: `rdr`, `rdbwselect`, `rdplot`.
- **rdlocrand** package: `rdrandinf`, `rdwinselect`, and more.
- **rddensity** package: `rddensity`.
- **rdmulti** package.
- **rdpower** package.

Ilustración Empírica: Ludwig and Miller (2007, QJE)

- **Problem:** impact of Head Start on Infant Mortality

- **Data:**

Y_i = child mortality 5 to 9 years old

T_i = whether county received Head Start assistance

X_i = 1960 poverty index ($\bar{x} = 59.1984$)

Z_i = see database.

- **Potential outcomes:**

$Y_i(0)$ = child mortality if **had not received** Head Start

$Y_i(1)$ = child mortality if **had received** Head Start

Overview

- Inferencia Causal, Evaluación de Impacto, RD.
- RD Plots (Gráficos) y Falsificación de Diseño.
- Estimación e Inferencia.
 - ▶ Local Polynomial Approach.
 - ▶ Local Randomization Approach.

Presentación Gráfica and Métodos de Falsificación

- Siempre usar RD Plots (Gráficos): una gran ventaja de RD!
- Graficar $\mathbb{E}[Y_i|X_i = x]$ para entender el diseño y su validez.
- Graficar $\mathbb{E}[Z_i|X_i = x]$, donde Z_i representa otras variables determinadas antes del tratamiento.
- Graficar la densidad de X_i para determinar validez del diseño: test de continuidad.

RD Plots (Gráficos)

- Ingredientes principales:
 - ▶ Estimación polinómica global (smooth fit).
 - ▶ Estimación local via bloques (bins) desconectados.
- Dos usos principales:
 - ▶ Detección de discontinuidades.
 - ▶ Representación de variabilidad.
- Parametros a seleccionar:
 - 1 Estimación polinómica: orden de polynomio (k_n)
 - 2 Estimación local: posición (ES or QS) y número de bloques (J_n)

RD Packages

<https://sites.google.com/site/rdpackages>

- **rdrobust** package: rdrobust, rdbwselect, rdplot.
- **rdlocrand** package: rdrandinf, rdwinselect, and more.
- **rddensity** package: rddensity.
- **rdmulti** package.
- **rdpower** package.

Ilustración Empírica: Ludwig and Miller (2007, QJE)

- **Problem:** impact of Head Start on Infant Mortality

- **Data:**

Y_i = child mortality 5 to 9 years old

T_i = whether county received Head Start assistance

X_i = 1960 poverty index ($\bar{x} = 59.1984$)

Z_i = see database.

- **Potential outcomes:**

$Y_i(0)$ = child mortality if **had not received** Head Start

$Y_i(1)$ = child mortality if **had received** Head Start

Continuidad de Densidad del Puntaje

- **Idea:** la distribución del puntaje debería ser similar (continua) de ambos lados del punto de corte.
- **Método 1:** Histograms & test Binomial.
- **Método 2:** Estimación polynomica local.

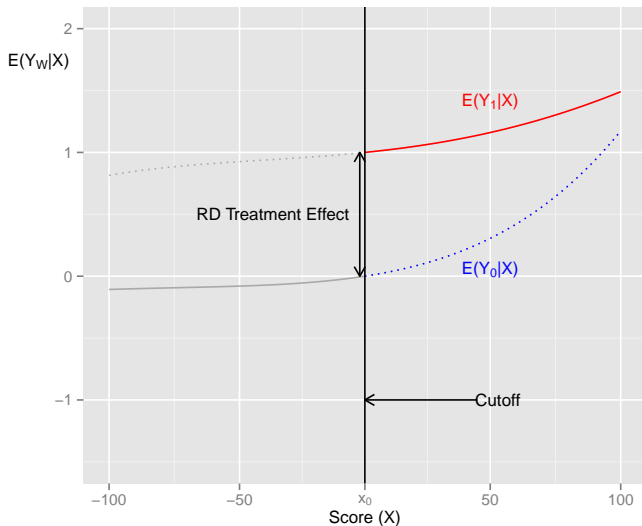
Covariate Balance and Placebo Tests

- Test placebo en covariados predeterminados.
 - ▶ **Idea:** el efecto de tratamiento debería ser cero.
 - ▶ Métodos: local polynomial, randomization-based.
- Test placebo en variables de resultado no afectadas.
 - ▶ **Idea:** el efecto de tratamiento debería ser cero.
 - ▶ Métodos: local polynomial, randomization-based.
- Necesitamos primero discutir estimación e inferencia en RD.

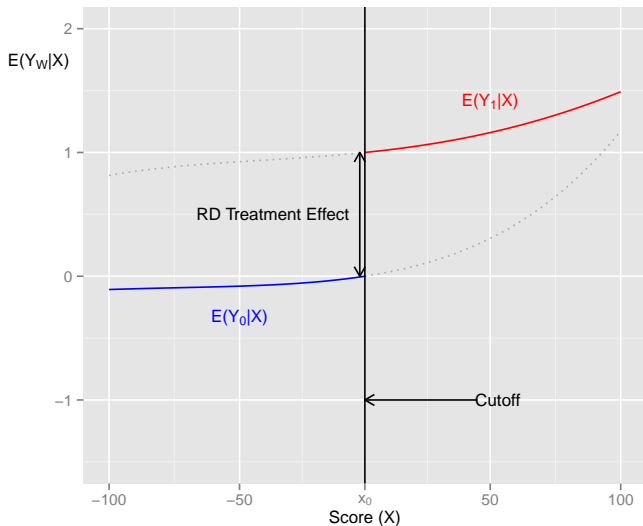
Overview

- Inferencia Causal, Evaluación de Impacto, RD.
- RD Plots (Gráficos) y Falsificación de Diseño.
- Estimación e Inferencia.
 - ▶ Local Polynomial Approach.
 - ▶ Local Randomization Approach.

$$\tau_{\text{SRD}} = \underbrace{\mathbb{E}[Y_{1i} - Y_{0i} | X_i = x_0]}_{\text{Unobservable}} = \lim_{x \downarrow x_0} \mathbb{E}[Y_i | X_i = x] - \lim_{x \uparrow x_0} \mathbb{E}[Y_i | X_i = x]$$



$$\tau_{\text{SRD}} = \underbrace{\mathbb{E}[Y_{1i} - Y_{0i} | X_i = x_0]}_{\text{Unobservable}} = \underbrace{\lim_{x \downarrow x_0} \mathbb{E}[Y_i | X_i = x]}_{\text{Observable}} - \underbrace{\lim_{x \uparrow x_0} \mathbb{E}[Y_i | X_i = x]}_{\text{Observable}}$$



Métodos de Estimación e Inferencia

- Local polynomial methods.
 - ▶ Métodos básicos.
 - ▶ Selección de bandwidth.
 - ▶ Estimación e inferencia.
- Local randomization methods.
 - ▶ Interpretación de experimento “local” cerca del punto de corte.
 - ▶ Selección de Window (ventana) cerca del punto de corte.
 - ▶ Estimación e inferencia.

Figure: Danger in Global Approximations

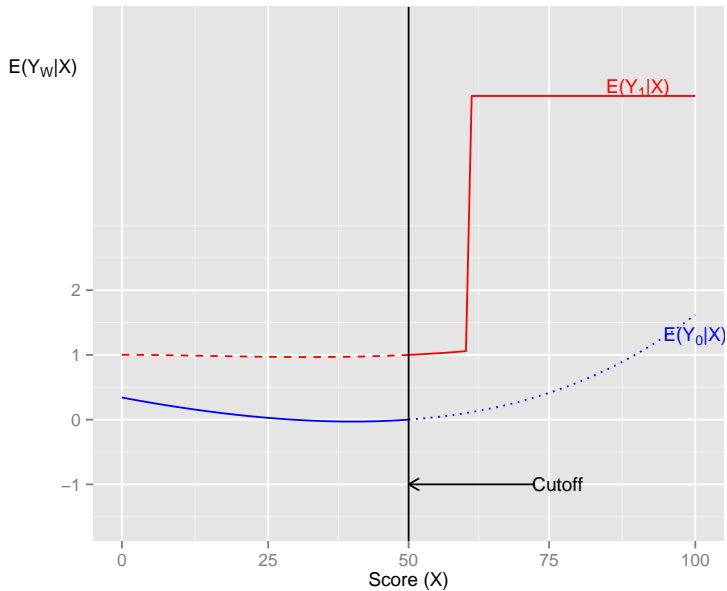


Figure: Danger in Global Approximations

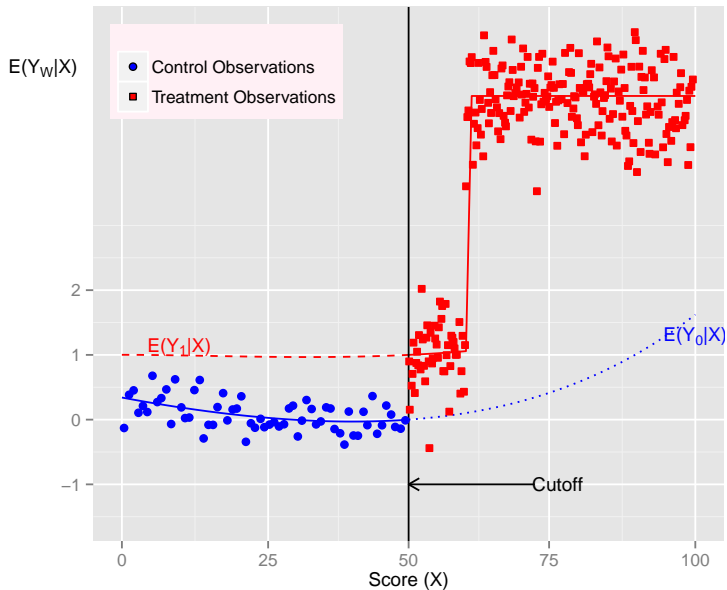


Figure: Danger in Global Approximations

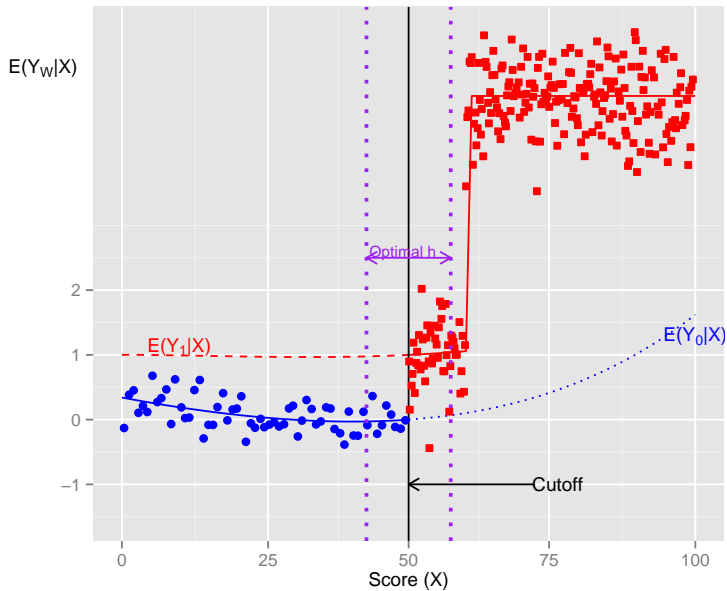


Figure: Danger in Global Approximations

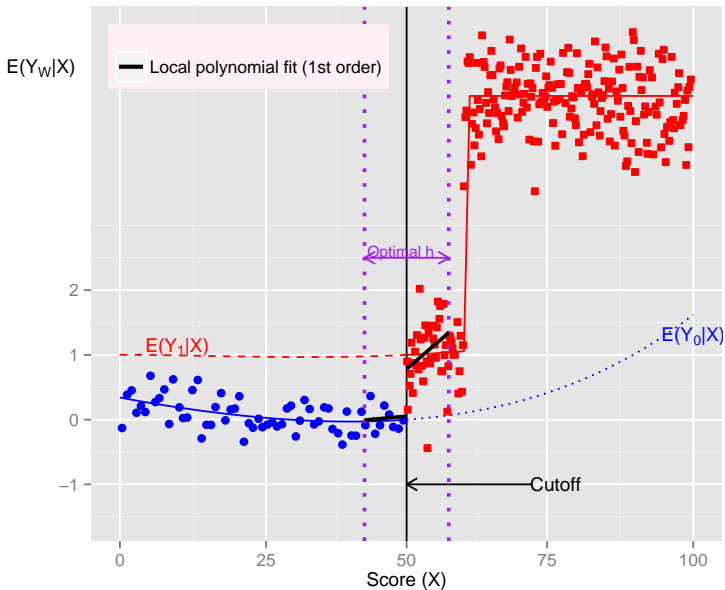
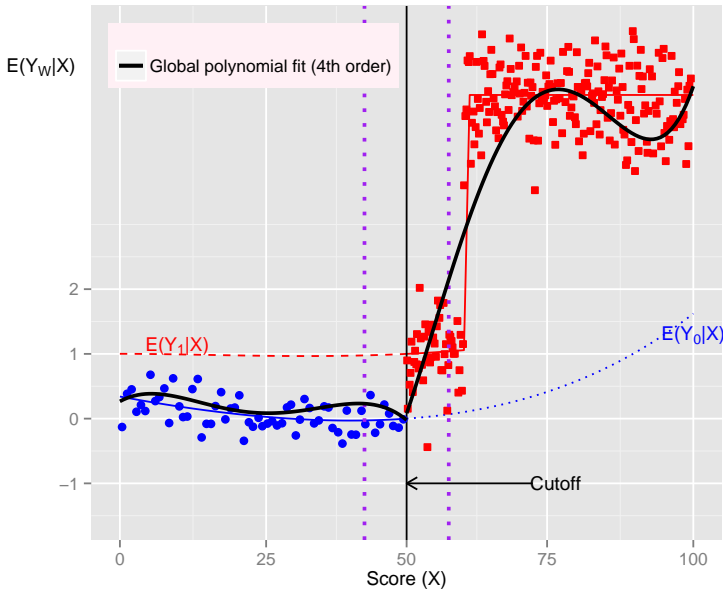
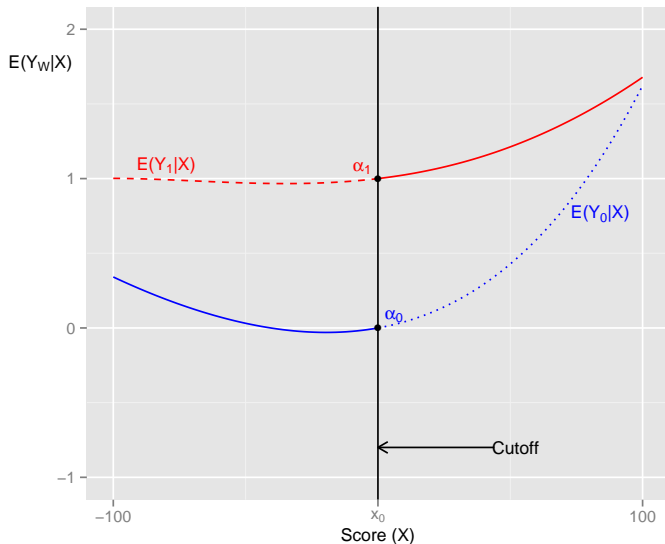


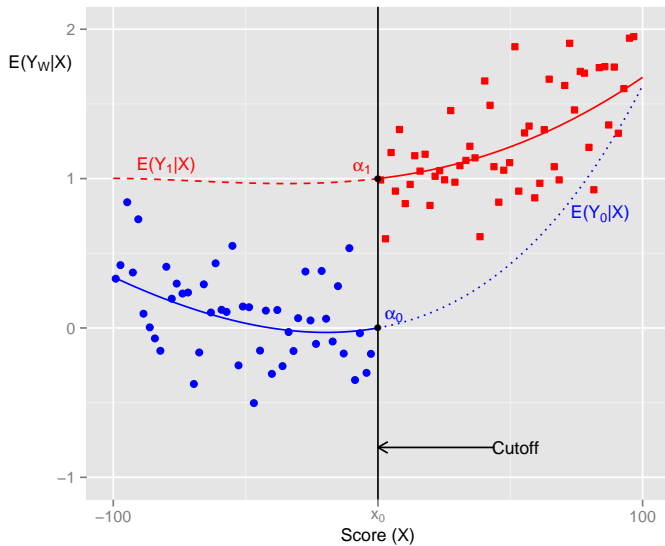
Figure: Danger in Global Approximations



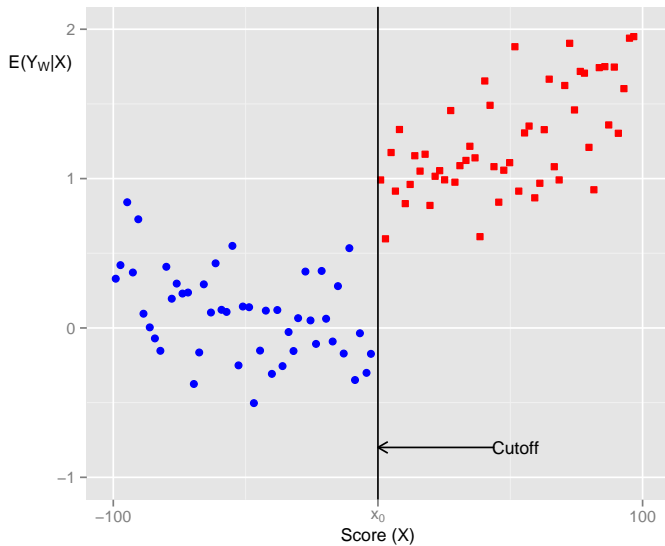
Local Polynomial Estimation



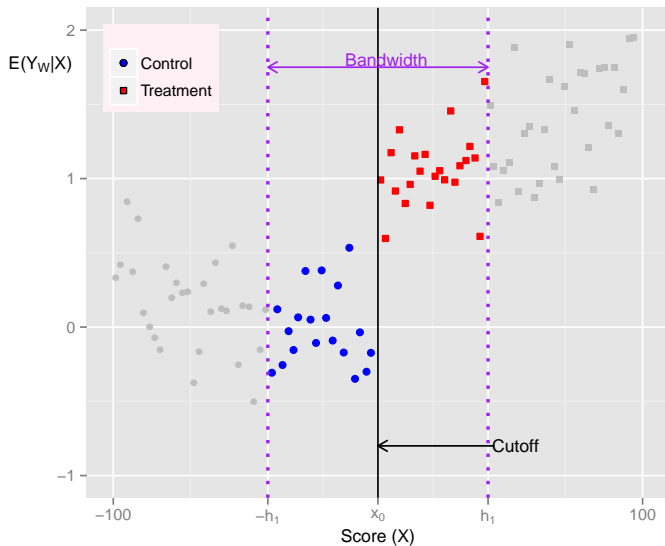
Local Polynomial Estimation



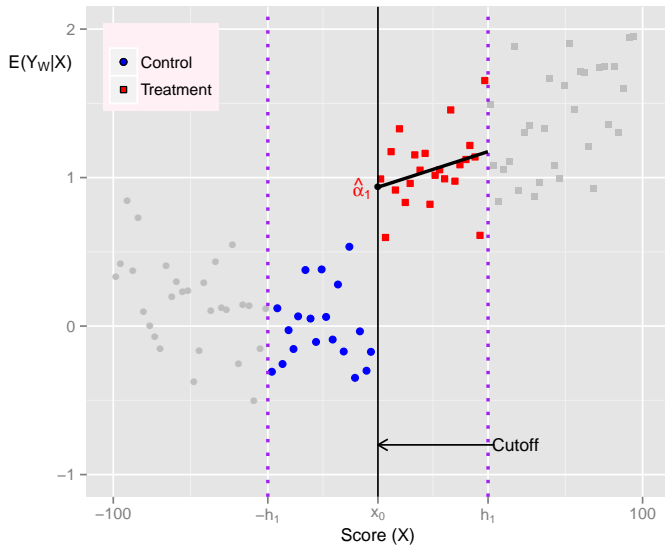
Local Polynomial Estimation



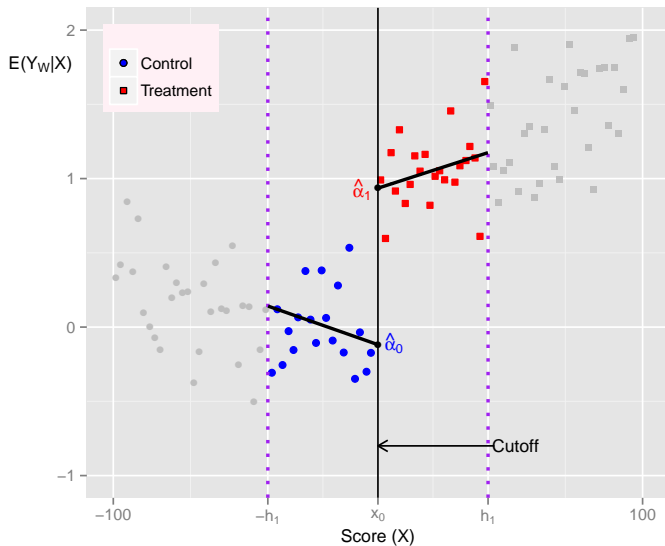
Local Polynomial Estimation



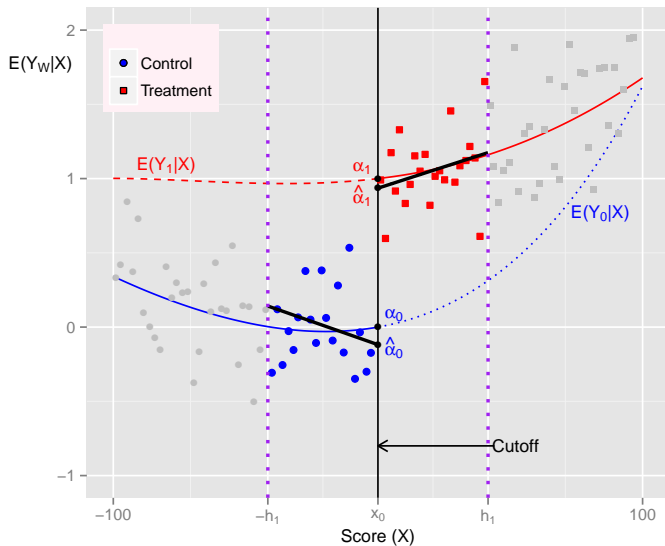
Local Polynomial Estimation



Local Polynomial Estimation



Local Polynomial Estimation



Local Polynomial Approach: Ingredientes Básicos

- $\mathbb{E}[Y_i|X_i = x]$ es aproximada cerca del punto de corte \bar{x} .
- Estimación polinómica local:
 - 1 Seleccionar bandwidth h : solo observaciones con $X_i \in [\bar{x} - h, \bar{x} + h]$.
 - 2 Seleccionar pesos via kernel function: $w_i = K\left(\frac{x_i - \bar{x}}{h}\right)$
 - 3 Seleccionar orden del polinomio: $p = 1$
- **Crítico:** selección de bandwidth h .

Selección de bandwidth h

- Mean Square Error Optimal (MSE-optimal) Choice:

$$\hat{h}_{\text{mse}} = \hat{C}_{\text{mse}} \cdot n^{-1/5}$$

- Coverage Error Rate Optimal (CER-optimal) Choice:

$$\hat{h}_{\text{cer}} = n^{-1/20} \cdot \hat{h}_{\text{mse}}$$

- **Idea:** trade-off sesgo y varianza de $\hat{\tau}_{\text{SRD}}(h_n)$,

$$\uparrow \text{Bias}(\hat{\tau}_{\text{SRD}}) \quad \implies \quad \downarrow \hat{h}$$

$$\uparrow \text{Var}(\hat{\tau}_{\text{SRD}}) \quad \implies \quad \uparrow \hat{h}$$

Selección de bandwidth h

- MSE-Optimal:

$$h_{\text{mse}} = C_{\text{mse}}^{1/5} \cdot n^{-1/5}$$

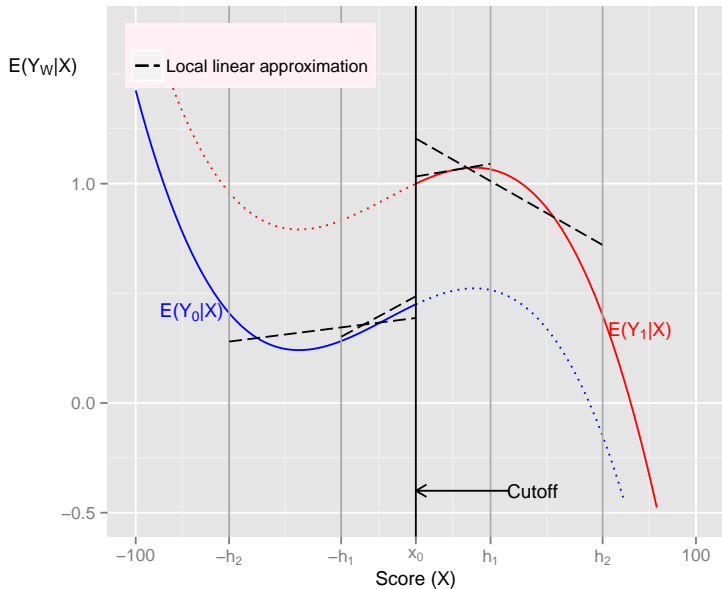
$$C_{\text{mse}} = C(K) \cdot \frac{\text{Var}(\hat{\tau}_{\text{SRD}})}{\text{Bias}(\hat{\tau}_{\text{SRD}})^2}$$

- CER-optimal:

$$h_{\text{cer}} = C_{\text{cer}}^{1/4} \cdot n^{-1/4}$$

$$C_{\text{cer}} \propto C(K) \cdot \frac{\text{Var}(\hat{\tau}_{\text{SRD}})}{|\text{Bias}(\hat{\tau}_{\text{SRD}})|}$$

Figure: Bias in Local Approximations—Choosing bandwidth



Local Polynomial Methods: Inferencia Robusta

- Approach 1: *Undersmoothing*/“*Small Bias*”.

$$\hat{I}(h_n) = \left[\hat{\tau}_{\text{SRD}} \pm 1.96 \cdot \sqrt{\hat{V}_n} \right]$$

- Approach 2: *Bias correction* (**not recommended**).

$$\hat{I}^{\text{bc}}(h_n, b_n) = \left[\left(\hat{\tau}_{\text{SRD}} - \hat{B}_n \right) \pm 1.96 \cdot \sqrt{\hat{V}_n} \right]$$

- Approach 3: *Robust Bias correction* (**recommended**).

$$\hat{I}^{\text{rbc}}(h_n, b_n) = \left[\left(\hat{\tau}_{\text{SRD}} - \hat{B}_n \right) \pm 1.96 \cdot \sqrt{\hat{V}_n + \hat{W}_n} \right]$$

RD Packages

<https://sites.google.com/site/rdpackages>

- **rdr** package: `rdr`, `rdbwselect`, `rdplot`.
- **rdlocrand** package: `rdrandinf`, `rdwinselect`, and more.
- **rddensity** package: `rddensity`.
- **rdmulti** package.
- **rdpower** package.

Ilustración Empírica: Ludwig and Miller (2007, QJE)

- **Problem:** impact of Head Start on Infant Mortality

- **Data:**

Y_i = child mortality 5 to 9 years old

T_i = whether county received Head Start assistance

X_i = 1960 poverty index ($\bar{x} = 59.1984$)

Z_i = see database.

- **Potential outcomes:**

$Y_i(0)$ = child mortality if **had not received** Head Start

$Y_i(1)$ = child mortality if **had received** Head Start

TABLE III
REGRESSION DISCONTINUITY ESTIMATES OF THE EFFECT OF HEAD START ASSISTANCE ON MORTALITY

Variable	Control mean	Nonparametric estimator			Parametric	
					Flexible linear	Flexible quadratic
Bandwidth or poverty range		9	18	36	8	16
Number of observations (counties) with nonzero weight		527	961	2,177	484	863
Main results						
Ages 5–9, Head Start-related causes, 1973–1983	3.238	-1.895** (0.980) [0.036]	-1.198* (0.796) [0.081]	-1.114** (0.544) [0.027]	-2.201** (1.004) [0.022]	-2.558** (1.261) [0.021]
Specification checks						
Ages 5–9, injuries, 1973–1983	22.303	0.195 (3.472) [0.924]	2.426 (2.476) [0.345]	0.679 (1.785) [0.755]	-0.164 (3.380) [0.998]	0.775 (3.401) [0.835]
Ages 5–9, all causes, 1973–1983	40.232	-3.416 (4.311) [0.415]	0.053 (3.098) [0.982]	-1.537 (2.253) [0.558]	-3.896 (4.268) [0.317]	-2.927 (4.295) [0.505]
Ages 25+, Head Start-related causes, 1973–1983	131.825	2.204 (5.719) [0.700]	6.016 (4.349) [0.147]	5.872 (3.338) [0.114]	2.091 (5.581) [0.749]	2.574 (6.415) [0.689]

Métodos de Estimación e Inferencia

- Local polynomial methods.
 - ▶ Métodos básicos.
 - ▶ Selección de bandwidth.
 - ▶ Estimación e inferencia.
- Local randomization methods.
 - ▶ Interpretación de experimento “local” cerca del punto de corte.
 - ▶ Selección de Window (ventana) cerca del punto de corte.
 - ▶ Estimación e inferencia.

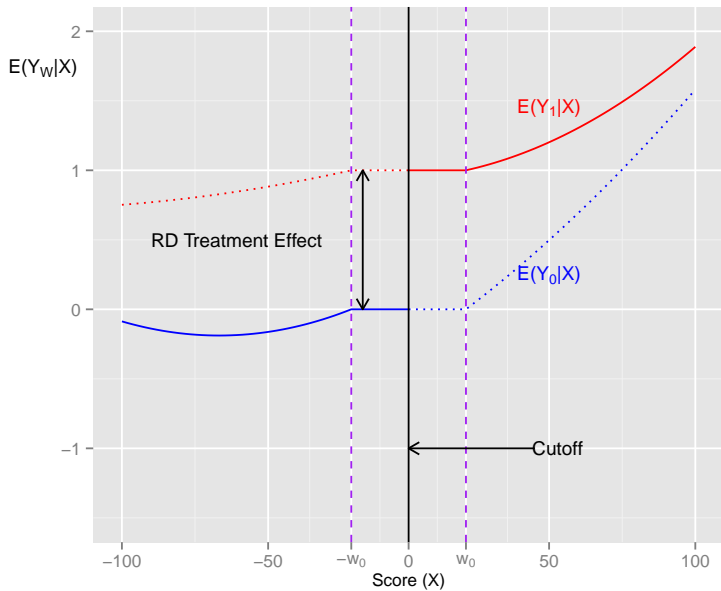
Local Randomization Approach

- **Idea:** cerca del punto de corte, RD es como si fuera un experimento controlado.
- Utilizamos métodos para el análisis de experimentos en una pequeña ventana alrededor del punto de corte:
 - ▶ Asume existe una $W = [-h_n, h_n]$, with $-h_n < \bar{x} < h_n$, tal que:

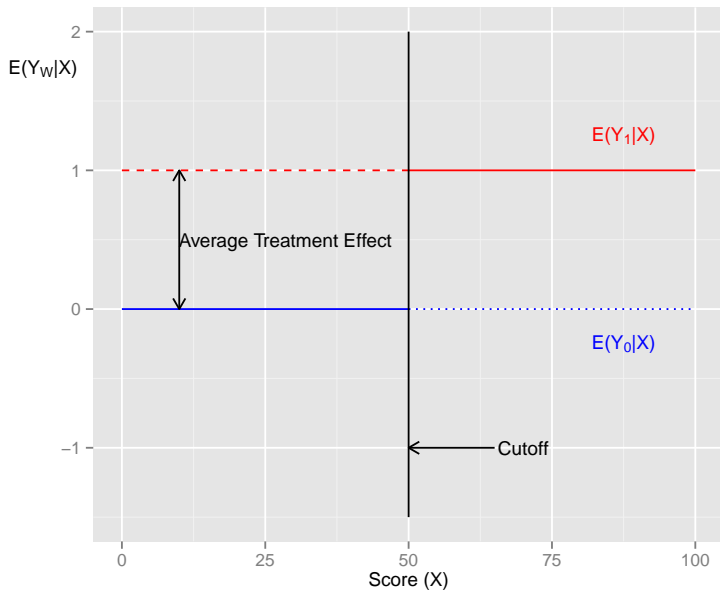
para todo $X_i \in W \implies T_i$ independiente de $(Y_i(0), Y_i(1))$

- **Cuidado:** este supuesto puede ser una buena aproximación *solamente muy cerca* del punto de corte!
- Una vez que la ventana W fue seleccionada, estimación e inferencia continua como en cualquier experimento.

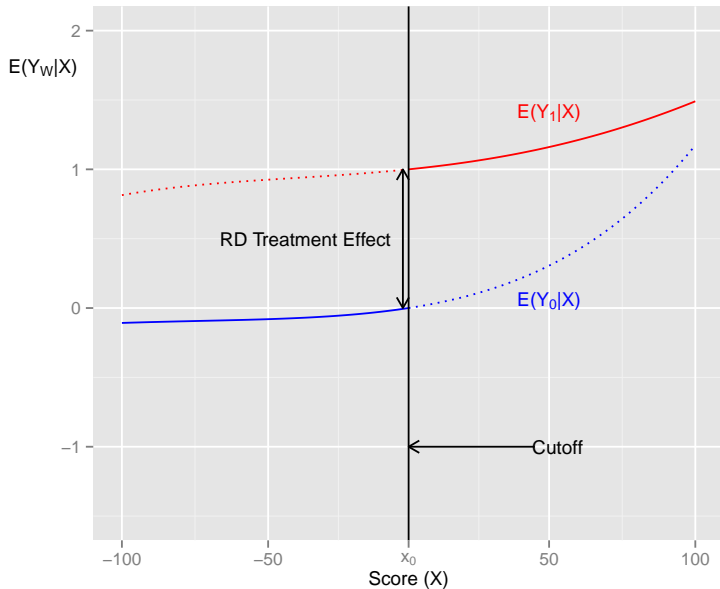
Figure: RD as a Local Experiment



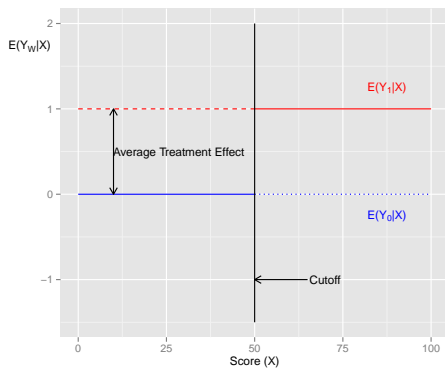
Experiments



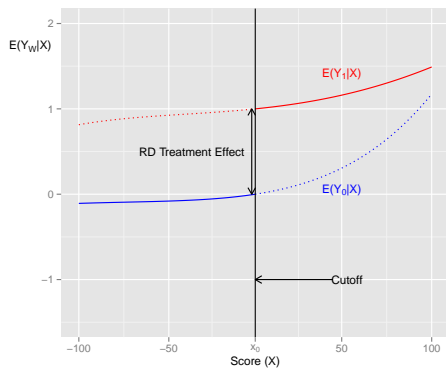
RD Designs



Experiment versus RD Design

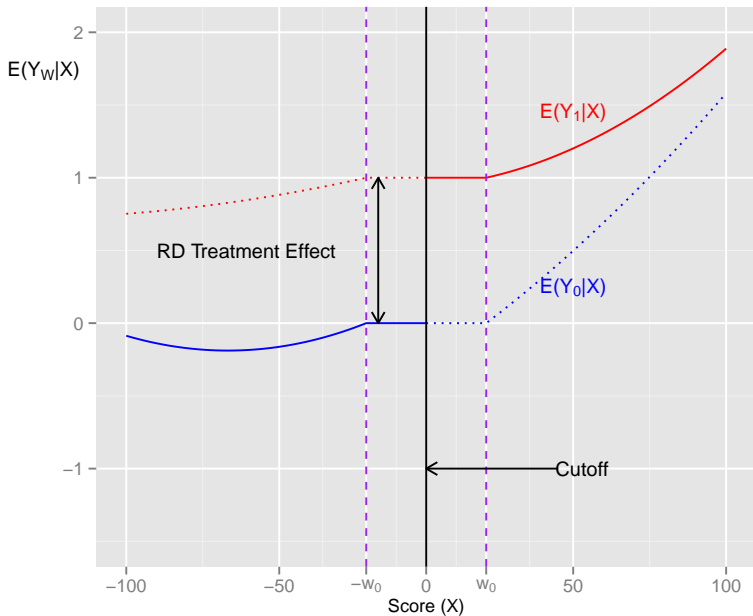


(a) Randomized Experiment



(b) RD Design

If as-if random interpretation is true: Locally Random



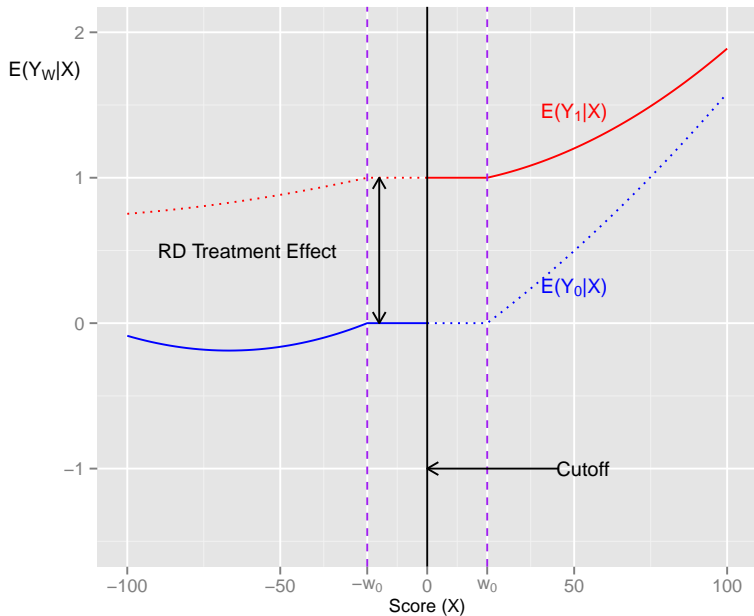
Muestras chicas: Randomization Inference

- Provee una alternativa de inferencia.
- Supuesto: existe una $W = [-h_n, h_n]$, with $-h_n < \bar{x} < h_n$, tal que:

para todo $X_i \in W \implies T_i$ independiente de $(Y_i(0), Y_i(1))$

- Inferencia exacta para toda muestra via randomization inference.
- Dos pasos:
 - ▶ Paso 1: Seleccionar ventana (window) y mecanismo de asignación de tratamiento.
 - ▶ Paso 2: Implementar randomization inference tools.

Remember: Locally Random RD



Choosing W_0 using predetermined covariate Z

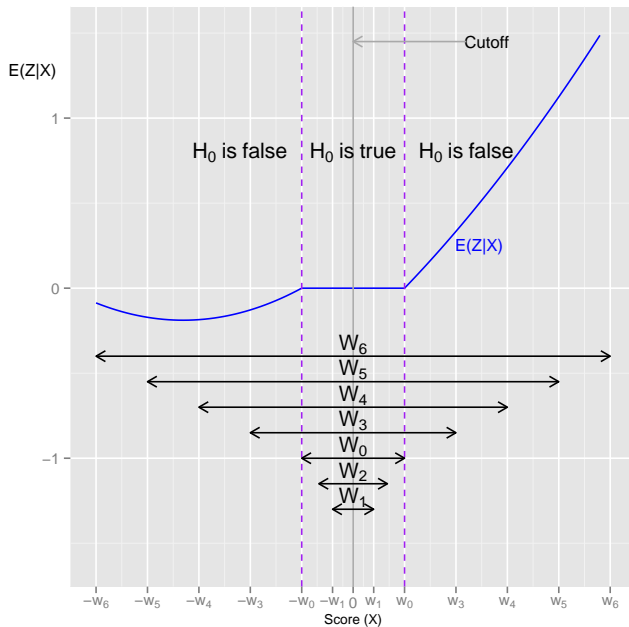


Ilustración Empírica: Ludwig and Miller (2007, QJE)

- **Problem:** impact of Head Start on Infant Mortality

- **Data:**

Y_i = child mortality 5 to 9 years old

T_i = whether county received Head Start assistance

X_i = 1960 poverty index ($\bar{x} = 59.1984$)

Z_i = see database.

- **Potential outcomes:**

$Y_i(0)$ = child mortality if **had not received** Head Start

$Y_i(1)$ = child mortality if **had received** Head Start

TABLE III
REGRESSION DISCONTINUITY ESTIMATES OF THE EFFECT OF HEAD START ASSISTANCE ON MORTALITY

Variable	Control mean	Nonparametric estimator			Parametric	
					Flexible linear	Flexible quadratic
Bandwidth or poverty range		9	18	36	8	16
Number of observations (counties) with nonzero weight		527	961	2,177	484	863
Main results						
Ages 5–9, Head Start-related causes, 1973–1983	3.238	-1.895** (0.980) [0.036]	-1.198* (0.796) [0.081]	-1.114** (0.544) [0.027]	-2.201** (1.004) [0.022]	-2.558** (1.261) [0.021]
Specification checks						
Ages 5–9, injuries, 1973–1983	22.303	0.195 (3.472) [0.924]	2.426 (2.476) [0.345]	0.679 (1.785) [0.755]	-0.164 (3.380) [0.998]	0.775 (3.401) [0.835]
Ages 5–9, all causes, 1973–1983	40.232	-3.416 (4.311) [0.415]	0.053 (3.098) [0.982]	-1.537 (2.253) [0.558]	-3.896 (4.268) [0.317]	-2.927 (4.295) [0.505]
Ages 25+, Head Start-related causes, 1973–1983	131.825	2.204 (5.719) [0.700]	6.016 (4.349) [0.147]	5.872 (3.338) [0.114]	2.091 (5.581) [0.749]	2.574 (6.415) [0.689]