

# CENTRO DE ESTUDIOS ECONÓMICOS

Maestría en Economía 2024–2026

Microeconometría para la Evaluación de Programas Sociales

## TAREA 3

**PRESENTA:** José Daniel Fuentes García

**PROFESORA:** Aurora Ramírez

**LABORATORISTA:** Mario Lechuga

## Índice

---

Lista de tablas	2
Lista de figuras	3
<b>Parte 1 : Teoría</b>	<b>4</b>
a) . . . . .	4
(i) . . . . .	4
(ii) . . . . .	5
(iii) . . . . .	6
b) . . . . .	7
(i) . . . . .	7
(ii) . . . . .	9
(iii) . . . . .	10
<b>Parte 2 : Ejercicios prácticos</b>	<b>11</b>
a) . . . . .	12
b) . . . . .	13
c) . . . . .	14
d) . . . . .	14
e) . . . . .	16
f) . . . . .	17
g) . . . . .	19
h) . . . . .	20
i) . . . . .	21
j) . . . . .	22
k) . . . . .	23
l) . . . . .	23
Resultados Formato Plantilla de Tablas . . . . .	25
TABLA 1 . . . . .	25
TABLA 2 . . . . .	26
TABLA 3 . . . . .	27
TABLA 4 . . . . .	27
TABLA 5 . . . . .	28

## Lista de tablas

---

1.	Estadísticas descriptivas de las covariables para el grupo control (NSW) . . .	12
2.	Estadísticas descriptivas de las covariables para el grupo tratado (NSW) . .	13
3.	Pruebas de igualdad de medias en covariables entre tratados y controles (NSW)	13
4.	Covariate balance tests: tratado vs. control (NSW) . . . . .	14
5.	Estadísticas descriptivas de las covariables para el grupo control (CPS) . . .	14
6.	Estadísticas descriptivas de las covariables para el grupo tratado (CPS) . . .	15
7.	Pruebas de igualdad de medias en covariables entre tratados y controles (CPS)	15
8.	Covariate balance tests: tratado vs. control (CPS) . . . . .	16
9.	Determinantes del Tratamiento: Modelo Probit para el Propensity Score . .	17
10.	Rango del Propensity Score en controles y tratados . . . . .	19
11.	Estimaciones del Efecto del Tratamiento (TOT) mediante métodos de empa- rejamiento basados en Propensity Score . . . . .	21
12.	Estimaciones del Efecto del Tratamiento (TOT) bajo soporte común con trim- ming tipo Crump (0.1–0.9) . . . . .	21
13.	Estimaciones placebo del Efecto del Tratamiento (ATT) sobre RE75 bajo soporte común clásico . . . . .	22
14.	Estimaciones placebo del ATT en RE75 bajo soporte común con trimming tipo Crump (0.1–0.9) . . . . .	22
15.	Estimaciones del ATET mediante métodos doblemente robustos (AIPW e IPWRA) bajo soporte común clásico. . . . .	23
16.	Estimaciones del ATET con métodos doblemente robustos (AIPW e IPWRA) bajo soporte común con trimming tipo Crump (0.1–0.9) . . . . .	23
17.	Benchmark experimental (NSW) . . . . .	25
18.	Efecto TOT (ATET) en CPS por regla de soporte . . . . .	26
19.	Placebo (RE75) por regla de soporte . . . . .	27
20.	Diagnóstico de overlap . . . . .	27
21.	Resumen comparativo . . . . .	28

## Lista de figuras

---

1.	Histograma del Propensity Score: Controles . . . . .	18
2.	Histograma del Propensity Score: Tratados . . . . .	18
3.	Histograma del Propensity Score: Tratados y Controles . . . . .	19
4.	Solapamiento Kernel: Control vs Tratados . . . . .	20

## Parte 1 : Teoría

a)

Se quiere estimar el **efecto de la educación** sobre los ingresos de los individuos utilizando datos observacionales.

Para realizar el análisis se sugiere correr una **regresión del logaritmo de salarios**,  $y_i$ , sobre los **años de escolaridad**,  $s_i$ , y otra variable relacionada con educación e ingresos denotada  $a_i$  o habilidad.

Es decir, se propone correr la siguiente regresión larga:

$$y_i = \alpha + \beta s_i + \gamma a_i + \varepsilon_i \quad (1)$$

donde los **coeficientes de regresión de la población**  $\alpha, \beta, \gamma$  están definidos de tal manera que  $\varepsilon_i$  no se encuentra correlacionado con  $s_i$  ni  $a_i$ .

(i)

Supón que se inicia estimando una **regresión corta** de  $y_i$  sobre  $s_i$  únicamente.

- Deriva el **coeficiente de regresión de la población** de  $s_i$  en términos de los parámetros en la ecuación (1).
- ¿Bajo qué condición son los coeficientes de  $s_i$  de la **regresión corta** y la **larga** iguales?

*Pista:*

$$E[\hat{\beta}_{corta}] = E[\hat{\beta}_{larga}] + algo\ más \dots$$

**RESPUESTA:**

$y_i = \alpha + \beta s_i + \gamma a_i + \varepsilon_i$	<b>Modelo</b>	
	<b>largo</b>	(2)
$\mathbb{E}[\varepsilon_i \mid s_i, a_i] = 0$	<b>Supuesto</b>	(3)
$\beta_{corta} = \frac{\text{Cov}(y_i, s_i)}{\text{Var}(s_i)}$	<b>Def. coef. población (corta)</b>	(4)
$\text{Cov}(y_i, s_i) = \text{Cov}(\alpha + \beta s_i + \gamma a_i + \varepsilon_i, s_i)$	<b>Sustituimos <math>y_i</math></b>	(5)
$= \text{Cov}(\alpha, s_i) + \text{Cov}(\beta s_i, s_i) + \text{Cov}(\gamma a_i, s_i) + \text{Cov}(\varepsilon_i, s_i)$	<b>Cov lineal</b>	(6)
$= 0 + \beta \text{Var}(s_i) + \gamma \text{Cov}(a_i, s_i) + 0$	<b>Constante, escala &amp; Cov(<math>\varepsilon_i, s_i</math>) = 0</b>	(7)
$= \beta \text{Var}(s_i) + \gamma \text{Cov}(a_i, s_i)$	<b>Reordenamos</b>	(8)

$$\beta_{\text{corta}} = \frac{\beta \text{Var}(s_i) + \gamma \text{Cov}(a_i, s_i)}{\text{Var}(s_i)} \quad \text{Sustituimos en cociente} \quad (9)$$

$$= \beta + \gamma \frac{\text{Cov}(a_i, s_i)}{\text{Var}(s_i)} \quad \text{Dividimos término a término} \quad (10)$$

$$\mathbb{E}[\hat{\beta}_{\text{corta}}] = \beta + \gamma \frac{\text{Cov}(a_i, s_i)}{\text{Var}(s_i)} \quad \text{Tomamos esperanza} \quad (11)$$

$$\mathbb{E}[\hat{\beta}_{\text{larga}}] = \beta \quad \text{En la larga } a_i \text{ controlada} \quad (12)$$

$$\mathbb{E}[\hat{\beta}_{\text{corta}}] = \mathbb{E}[\hat{\beta}_{\text{larga}}] + \gamma \frac{\text{Cov}(a_i, s_i)}{\text{Var}(s_i)}$$

**Pista:**  $\mathbb{E}[\beta_{\text{corta}}] = \mathbb{E}[\beta_{\text{larga}}] + \text{algo más}$

(13)

$$\beta_{\text{corta}} = \beta \Leftrightarrow \gamma \text{Cov}(a_i, s_i) = 0 \quad \text{Condición de igualdad} \quad (14)$$

(ii)

Realizar la **regresión corta** es similar a tener un **problema de variables omitidas**.

- ¿Cómo interpretarías al **componente adicional** y qué **signo** crees que tendría?
- ¿ $\hat{\beta}_{\text{corta}}$  estaría **sobre-estimado** o **sub-estimado**?

### RESPUESTA:

- La **regresión corta** omite la habilidad  $a_i$ .
- Si  $a_i$  está **correlacionada** con la escolaridad  $s_i$  y además **afecta** el salario, el efecto de  $a_i$  se «**cuela**» en el coeficiente de  $s_i$ .
- El **signo del sesgo** es el producto del **signo de**  $\gamma$  (cómo  $a_i$  afecta salarios) por el **signo de**  $\text{Cov}(a_i, s_i)$  (cómo  $a_i$  se asocia con educación).

$$y_i = \alpha + \beta s_i + \gamma a_i + \varepsilon_i \quad \text{Modelo largo} \quad (15)$$

$$a_i = \pi s_i + u_i, \quad \mathbb{E}[u_i s_i] = 0 \quad \text{Proyección de } a_i \text{ sobre } s_i \quad (16)$$

$$y_i = \alpha + (\beta + \gamma\pi)s_i + \gamma u_i + \varepsilon_i \quad \text{Sustituimos } a_i \text{ y reagrupamos} \quad (17)$$

$$\Rightarrow \mathbb{E}[\hat{\beta}_{\text{corta}}] = \beta + \gamma\pi \quad \text{Coef. poblacional en la corta} \quad (18)$$

$$\pi = \frac{\text{Cov}(a_i, s_i)}{\text{Var}(s_i)} \quad \text{Fórmula de la proyección} \quad (19)$$

$$\boxed{\mathbb{E}[\hat{\beta}_{\text{corta}}] = \beta + \gamma \frac{\text{Cov}(a_i, s_i)}{\text{Var}(s_i)}} \quad \text{Sesgo por omitida} \quad (20)$$

$$\text{Sesgo} = \mathbb{E}[\hat{\beta}_{\text{corta}}] - \beta = \gamma \frac{\text{Cov}(a_i, s_i)}{\text{Var}(s_i)} \quad \text{Definición del sesgo} \quad (21)$$

$$\text{signo}(\text{Sesgo}) = \text{signo}(\gamma) \times \text{signo}(\text{Cov}(a_i, s_i)) \quad \text{Regla de signos} \quad (22)$$

### Signo

- **Sobre-estimado** ( $\uparrow$ ): si  $\gamma > 0$  (mayor habilidad  $\Rightarrow$  mayor salario) y  $\text{Cov}(a_i, s_i) > 0$  (más educación entre más hábiles), entonces el sesgo es **positivo** y  $\hat{\beta}_{\text{corta}}$  **sobre-estima** a  $\beta$ .
- **Sub-estimado** ( $\downarrow$ ): si  $\gamma > 0$  pero  $\text{Cov}(a_i, s_i) < 0$  (educación asociada a menor  $a_i$ ), el sesgo es **negativo** y  $\hat{\beta}_{\text{corta}}$  **sub-estima** a  $\beta$ .

(iii)

¿Por qué es más probable que la **regresión larga** tenga una **interpretación causal**?  
¿O puede ser que en realidad no sea así?

### RESPUESTA:

- La **regresión larga** controla por **habilidad**  $a_i$  (u otros factores relevantes). Al incluir  $a_i$ , se elimina el **sesgo por variable omitida** en el coeficiente de  $s_i$  proveniente de la correlación entre  $s_i$  y  $a_i$ .
- Así, el coeficiente de  $s_i$  en la larga se aproxima al **efecto causal parcial** de la educación sobre salarios (manteniendo  $a_i$  fijo).

- Pero la interpretación causal puede fallar si persisten **omisiones** (otras variables  $z_i$  no observadas correlacionadas con  $s_i$  y con  $y_i$ ), **endogeneidad** (p.ej., selección, expectativas), o **medición imperfecta** de  $a_i$ .

$$y_i = \alpha + \beta s_i + \gamma a_i + \varepsilon_i \quad \begin{array}{l} \text{Regresión} \\ \text{larga} \end{array} \quad (23)$$

$$y_i = \alpha + \beta_{\text{corta}} s_i + u_i \quad \begin{array}{l} \text{Regresión} \\ \text{corta} \end{array} \quad (24)$$

$$\mathbb{E}[\hat{\beta}_{\text{corta}}] = \beta + \gamma \frac{\text{Cov}(a_i, s_i)}{\text{Var}(s_i)} \quad \begin{array}{l} \text{Sesgo por} \\ \text{omitida} \end{array} \quad (25)$$

$$\mathbb{E}[\hat{\beta}_{\text{larga}}] = \beta \quad \begin{array}{l} \text{Al controlar} \\ a_i \end{array} \quad (26)$$

### Implicación.

- En la **corta**, el coeficiente de  $s_i$  mezcla el **efecto propio** de la educación con el **efecto indirecto** de  $a_i$  (si  $a_i$  está correlacionado con  $s_i$ ).
- En la **larga**, el término adicional desaparece al **condicionar en**  $a_i$ , por lo que el coeficiente de  $s_i$  refleja más limpiamente el **efecto parcial** de la educación.
- **Aun así**, la causalidad estricta requiere que no existan **otras** omitidas  $z_i$  correlacionadas con  $s_i$  y  $y_i$ , ni **endogeneidad** (p.ej., elección de  $s_i$  guiada por choques no observados en  $y_i$ ), ni **error de medición** relevante en  $a_i$  o  $s_i$ .

### b)

Este ejercicio muestra la **relación** entre **regresión con controles** y **matching**:

#### (i)

Supón que sólo hay una **covariante**  $x_i$  y que ésta es **binaria**.

Adicionalmente, asume que la **probabilidad** de que se le asigne el tratamiento a un individuo es **positiva** para ambos valores de  $x_i$ .

Deriva los **estimadores de matching** y de **regresión**.

### RESPUESTA:

- Supongamos una sola covariable  $x_i$ , que es **binaria**:  $x_i \in \{0, 1\}$ .



- La probabilidad de tratamiento es positiva para ambos valores de  $x_i$ , es decir:

$$0 < \Pr(T_i = 1 \mid x_i = 0), \Pr(T_i = 1 \mid x_i = 1) < 1.$$

- Queremos derivar:
  - El **estimador de matching**: compara promedios de tratados y controles **dentro de cada grupo de  $x_i$** .
  - El **estimador de regresión**: surge de una regresión lineal de  $y_i$  sobre  $T_i$  y  $x_i$ .

$$\tau(x) = \mathbb{E}[y_i(1) - y_i(0) \mid x_i = x] \quad \text{Definición efecto cond.} \quad (27)$$

$$= \mathbb{E}[y_i(1) \mid T_i = 1, x_i = x] - \mathbb{E}[y_i(0) \mid T_i = 0, x_i = x] \quad \text{Efecto por estrato} \quad (28)$$

$$\hat{\tau}_{\text{match}} = \sum_{x \in \{0,1\}} \Pr(x_i = x \mid T_i = 1) \left( \bar{y}_{1,x} - \bar{y}_{0,x} \right) \quad \text{Efecto TOT con matching} \quad (29)$$

donde:

$$\bar{y}_{1,x} = \frac{1}{n_{1x}} \sum_{i:T_i=1, x_i=x} y_i, \quad \bar{y}_{0,x} = \frac{1}{n_{0x}} \sum_{i:T_i=0, x_i=x} y_i.$$

- Consideremos el modelo lineal poblacional:

$$y_i = \alpha + \tau T_i + \delta x_i + \varepsilon_i \quad \text{Modelo regresión} \quad (30)$$

$$\mathbb{E}[\varepsilon_i \mid T_i, x_i] = 0 \quad \text{Supuesto de exogeneidad} \quad (31)$$

- El coeficiente  $\tau$  de  $T_i$  se interpreta como el **efecto parcial del tratamiento**, controlando por  $x_i$ .
- Ahora derivamos  $\tau$  explícitamente en términos de medias condicionales.

$$\hat{\tau}_{\text{reg}} = \left( \bar{y}_{1,1} - \bar{y}_{0,1} \right) \Pr(x_i = 1) + \left( \bar{y}_{1,0} - \bar{y}_{0,0} \right) \Pr(x_i = 0) \quad \text{Coef. de } T_i \text{ en la regresión} \quad (32)$$

- Es decir, el coeficiente  $\tau$  de la regresión es una **media ponderada de los efectos condicionales** por cada valor de  $x_i$ , con pesos dados por la distribución de  $x_i$  en la muestra.
- Esto coincide con el estimador de matching si se usan los mismos pesos (los de la población de tratados).
- El estimador de **matching** calcula directamente diferencias de medias de tratados vs controles dentro de cada  $x$  y luego promedia con pesos de tratados.
- El estimador de **regresión** entrega el mismo resultado en este caso binario, pues el coeficiente de  $T_i$  en la regresión con control por  $x_i$  coincide con la combinación ponderada de efectos condicionales.
- $\Rightarrow$  Con una covariable binaria, **matching y regresión son equivalentes**.

(ii)

Prueba que los **estimadores de matching** y los **estimadores de la regresión** van a estimar el **mismo parámetro** cuando la **asignación del tratamiento es aleatoria**.

**RESPUESTA:**

**Idea.** Con **asignación aleatoria** se cumple

$$T_i \perp\!\!\!\perp (y_{1i}, y_{0i}, x_i), \quad \Pr(T_i = 1 \mid x_i) = p \in (0, 1) \text{ (constante)}.$$

Entonces, las **diferencias dentro de estratos  $x$**  igualan el efecto condicional y los **pesos** coinciden; por eso matching y regresión estiman el mismo parámetro.

$$y_i = T_i y_{1i} + (1 - T_i) y_{0i} \quad \text{Def. outcomes} \quad (33)$$

$$\mathbb{E}[y_i \mid T_i = 1, x_i = x] = \mathbb{E}[y_{1i} \mid x_i = x] \quad \text{Aleatorización: } T \perp (y_1, y_0) \mid x \quad (34)$$

$$\mathbb{E}[y_i \mid T_i = 0, x_i = x] = \mathbb{E}[y_{0i} \mid x_i = x] \quad \text{Análogo} \quad (35)$$

$$\Rightarrow \Delta(x) := \mathbb{E}[y \mid T = 1, x] - \mathbb{E}[y \mid T = 0, x] = \mathbb{E}[y_1 - y_0 \mid x] \quad \text{Efecto condicional} \quad (36)$$

$$\tau_{\text{match}} = \sum_x \Pr(x) \Delta(x) = \sum_x \Pr(x) \mathbb{E}[y_1 - y_0 \mid x] \quad \text{Matching: promedio por } x \quad (37)$$

$$= \mathbb{E}_x [\mathbb{E}[y_1 - y_0 \mid x]] = \mathbb{E}[y_1 - y_0] \quad \text{Ley de la esperanza total} \quad (38)$$

$$\boxed{\tau_{\text{match}} = \mathbb{E}[y_1 - y_0]} \quad \text{ATE} \quad (39)$$

- **Aleatoriedad condicional:**  $T_i \perp (y_{1i}, y_{0i}) \mid x_i$  (asignación aleatoria dado  $x_i$ ).
- **Overlap:**  $0 < p(x) := \Pr(T_i = 1 \mid x_i = x) < 1$  para todo  $x$ .
- **Modelo de proyección poblacional (para el coeficiente parcial):**

$$y_i = \alpha + \tau T_i + \delta x_i + \varepsilon_i, \quad \mathbb{E}[\varepsilon_i \mid T_i, x_i] = 0.$$

$$\mu_t(x) := \mathbb{E}[y_i \mid T_i = t, x_i = x] \quad \begin{array}{l} \text{Medias} \\ \text{condicionales} \end{array} \quad (40)$$

$$T \perp (y_1, y_0) \text{ dado } x \quad (41)$$

$$(\text{Aleatoriedad}) \Rightarrow \mu_1(x) = \mathbb{E}[y_{1i} \mid x_i = x], \quad \mu_0(x) = \mathbb{E}[y_{0i} \mid x_i = x]$$

$$\Delta(x) := \mu_1(x) - \mu_0(x) = \mathbb{E}[y_{1i} - y_{0i} \mid x_i = x] \quad \begin{array}{l} \text{Efecto} \\ \text{condicional} \end{array} \quad (42)$$

$$\tau_{\text{reg}} = \frac{\text{Cov}(y_i, T_i \mid x_i)}{\text{Var}(T_i \mid x_i)} \quad \begin{array}{l} \text{Coef. parcial} \\ \text{(FWL)} \end{array} \quad (43)$$

$$\text{Cov}(y, T \mid x) = \mathbb{E}[(y - \mathbb{E}[y \mid x])(T - \mathbb{E}[T \mid x]) \mid x] = p(x)(1 - p(x)) \Delta(x) \quad \begin{array}{l} \text{Cálculo} \\ \text{condicional} \end{array} \quad (44)$$

$$\text{Var}(T \mid x) = p(x)(1 - p(x)) \quad \begin{array}{l} \text{T binaria} \\ \text{condicional} \end{array} \quad (45)$$

$$\Rightarrow \tau_{\text{reg}} = \frac{p(x)(1 - p(x)) \Delta(x)}{p(x)(1 - p(x))} = \Delta(x) \quad \begin{array}{l} \text{Se cancela} \\ p(x)(1 - p(x)) \end{array} \quad (46)$$

$$\tau_{\text{reg}} = \mathbb{E}[\Delta(x_i)] = \mathbb{E}[\mathbb{E}[y_{1i} - y_{0i} \mid x_i]] = \mathbb{E}[y_{1i} - y_{0i}] \quad \begin{array}{l} \text{Esperanza} \\ \text{iterada} \end{array} \quad (47)$$

$$\boxed{\tau_{\text{reg}} = \mathbb{E}[y_1 - y_0]} \quad \text{ATE} \quad (48)$$

El coeficiente de tratamiento en la regresión con control por  $x$  es, bajo asignación aleatoria y solapamiento, el *promedio (sobre  $x$ ) de las diferencias dentro de  $x$* . Como esas diferencias son  $\mathbb{E}[y_1 - y_0 \mid x]$ , su promedio es el **ATE**.

(iii)

¿Por qué podrías preferir los **estimadores de la regresión** sobre los **estimadores de matching**, inclusive si tu principal interés es el **efecto del tratamiento sobre los tratados**?

**RESPUESTA:**

- La **regresión** utiliza toda la muestra y ajusta un modelo paramétrico. Esto permite:
  - **Mayor eficiencia**: aprovecha más observaciones en lugar de depender sólo de las “parejas cercanas” como en matching.
  - **Menor varianza**: especialmente útil cuando la covariable  $x$  toma pocos valores o cuando hay desbalance en el número de tratados y controles.
- El **matching** puede descartar observaciones (cuando no encuentra buenos matches), lo que reduce eficiencia y puede introducir sesgo si no hay suficiente soporte común.
- Aunque el interés sea el **efecto sobre los tratados (ATT)**, la regresión puede estimar ese efecto con **menor error cuadrático medio** si el modelo está bien especificado, porque ajusta explícitamente la relación entre  $y$ ,  $T$ , y  $x$ .
- En la práctica:
  - **Matching**  $\Rightarrow$  más flexible, menos supuestos, pero menos eficiente.
  - **Regresión**  $\Rightarrow$  más eficiente, más supuestos (linealidad condicional, forma funcional).

Incluso si tu objetivo es el **ATT**, la regresión puede ser preferible porque usa toda la información disponible y produce estimadores más precisos, siempre que el modelo paramétrico esté razonablemente bien especificado.

## Parte 2 : Ejercicios prácticos

### *Experimentos aleatorios, regresión con controles y matching*

Esta sección tiene dos objetivos.

Primero, practicarás el uso de diferentes métodos para identificar el **efecto promedio del tratamiento sobre los tratados (TOT)**: comparación de medias en experimento aleatorio, regresión con controles, y métodos de matching (estratificación, vecino más cercano, kernel, radius).

Segundo, compararás e interpretarás diferencias entre estimadores no experimentales bajo distintas reglas de soporte común, en particular:

1. Soporte común de pscore (data-driven por bloques), y
2. Trimming tipo Crump (regla 0.1–0.9).

Adicionalmente, incorporarás **pruebas placebo y estimadores doblemente robustos (AIPW/IPWRA)**, en línea con las buenas prácticas enfatizadas por Imbens & Xu (2024).

**Entrega:** Graba tu trabajo en un `.do` de Stata y genera un único archivo PDF autocontenido que incluya todas las tablas de resultados y la discusión final.

### Datos:

- `nsw_data.dta` (muestra experimental; tratados y controles del NSW).
- `cps_data.dta` (muestra no experimental; mismos tratados del NSW + controles CPS).

a)

Carga `nsw_data.dta` y reporta **estadísticas descriptivas** por estado de tratamiento usando `summarize` o `tabstat`.

Tabla 1: Estadísticas descriptivas de las covariables para el grupo control (NSW)

Variable	Media	Desv.Est.	Min	Max	N
Edad	25.05	7.06	17.00	55.00	260
Educación	10.09	1.61	3.00	14.00	260
Negro	0.83	0.38	0.00	1.00	260
Hispano	0.11	0.31	0.00	1.00	260
Casado	0.15	0.36	0.00	1.00	260
Sin Grado	0.83	0.37	0.00	1.00	260
RE74	2107.03	5687.91	0.00	39570.68	260
RE75	1266.91	3102.98	0.00	23031.98	260
RE78	4554.80	5483.84	0.00	39483.53	260

Fuente: Elaboración propia con datos de `nsw.dta`. Proceso disponible en:  danifuentesga

Tabla 2: Estadísticas descriptivas de las covariables para el grupo tratado (NSW)

Variable	Media	Desv.Est.	Min	Max	N
Edad	25.82	7.16	17.00	48.00	185
Educación	10.35	2.01	4.00	16.00	185
Negro	0.84	0.36	0.00	1.00	185
Hispano	0.06	0.24	0.00	1.00	185
Casado	0.19	0.39	0.00	1.00	185
Sin Grado	0.71	0.46	0.00	1.00	185
RE74	2095.57	4886.62	0.00	35040.07	185
RE75	1532.06	3219.25	0.00	25142.24	185
RE78	6349.14	7867.40	0.00	60307.93	185

Fuente: Elaboración propia con datos de nsw.dta. Proceso disponible en:  danifuentesga


## b)

Con `ttest`, prueba **igualdad de medias** en covariables entre grupos experimentales tratados y controles.

Incluye **media por grupo, estadístico t y valor p**.

Tabla 3: Pruebas de igualdad de medias en covariables entre tratados y controles (NSW)

Variable	Media Control	Media Tratado	t	p
Edad	25.05	25.82	-1.12	0.265
Educación	10.09	10.35	-1.50	0.135
Negro	0.83	0.84	-0.45	0.649
Hispano	0.11	0.06	1.78	0.076
Casado	0.15	0.19	-0.98	0.327
Sin Grado	0.83	0.71	3.22	0.001
RE74	2107.03	2095.57	0.02	0.982
RE75	1266.91	1532.06	-0.87	0.382
RE78	4554.80	6349.14	-2.84	0.005

Fuente: Elaboración propia con datos de nsw.dta. Proceso disponible en:  danifuentesga

c)

Estima el **TOT experimental** por **regresión sin controles** y con **controles** (usa las variables: `age`, `age2`, `education`, `education2`, `black`, `hispanic`, `married`, `re75`, `re752`, `re74`, `re742`).

Tabla 4: Covariate balance tests: tratado vs. control (NSW)

Modelo	Coef. Tratamiento	Error Est.
Sin controles	1599.75	662.78
Con controles	1599.75	662.78

Fuente: Elaboración propia con datos de `nsw.dta`. Proceso disponible en:  danifuentesga

d)

Repita (a)–(c) con `cps_data.dta`.  
 Compara contra (a)–(c) y discute por qué difieren.

Tabla 5: Estadísticas descriptivas de las covariables para el grupo control (CPS)

Variable	Media	Desv.Est.	Min	Max	N
Edad	28.25	11.70	16.00	55.00	2369
Educación	11.24	2.58	0.00	18.00	2369
Negro	0.11	0.32	0.00	1.00	2369
Hispano	0.08	0.28	0.00	1.00	2369
Casado	0.46	0.50	0.00	1.00	2369
Sin Grado	0.45	0.50	0.00	1.00	2369
RE74	8727.96	8967.85	0.00	25862.32	2369
RE75	7397.23	8112.21	0.00	25243.55	2369
RE78	10171.11	8851.88	0.00	25564.67	2369

Fuente: Elaboración propia con datos de `cps.dta`. Proceso disponible en:  danifuentesga

Tabla 6: Estadísticas descriptivas de las covariables para el grupo tratado (CPS)

Variable	Media	Desv.Est.	Min	Max	N
Edad	25.82	7.16	17.00	48.00	185
Educación	10.35	2.01	4.00	16.00	185
Negro	0.84	0.36	0.00	1.00	185
Hispano	0.06	0.24	0.00	1.00	185
Casado	0.19	0.39	0.00	1.00	185
Sin Grado	0.71	0.46	0.00	1.00	185
RE74	2095.57	4886.62	0.00	35040.07	185
RE75	1532.06	3219.25	0.00	25142.24	185
RE78	6349.14	7867.40	0.00	60307.93	185

Fuente: Elaboración propia con datos de cps.dta. Proceso disponible en:  danifuentesga

Tabla 7: Pruebas de igualdad de medias en covariables entre tratados y controles (CPS)

Variable	Media Control	Media Tratado	t	p
Edad	28.25	25.82	2.79	0.005
Educación	11.24	10.35	4.60	0.000
Negro	0.11	0.84	-29.97	0.000
Hispano	0.08	0.06	1.17	0.242
Casado	0.46	0.19	7.21	0.000
Sin Grado	0.45	0.71	-6.82	0.000
RE74	8727.96	2095.57	9.94	0.000
RE75	7397.23	1532.06	9.77	0.000
RE78	10171.11	6349.14	5.70	0.000

Fuente: Elaboración propia con datos de cps.dta. Proceso disponible en:  danifuentesga



Tabla 8: Covariate balance tests: tratado vs. control (CPS)

Modelo	Coef. Tratamiento	Error Est.
Sin controles	-3821.97	605.07
Con controles	673.50	652.23

Fuente: Elaboración propia con datos de cps.dta. Proceso disponible en:  danifuentesga

- **Descriptivos:** En **NSW**, tratados y controles son jóvenes, con baja escolaridad e ingresos previos bajos, reflejando buen **balance**. En **CPS**, los grupos difieren sustancialmente en edad, educación e ingresos previos.
- **t-tests:** En **NSW** no hay diferencias significativas en las covariables, confirmando la aleatorización. En **CPS** casi todas las **pruebas** muestran desbalance marcado.
- **TOT:** En **NSW**, el efecto del programa es estable ( 1600) con y sin controles, mostrando un **impacto causal**. En **CPS**, el TOT es mayor y sensible a los controles, evidenciando **sesgo de selección**.

e)

Utiliza la **muestra no experimental** para estimar el **propensity score** corriendo el programa `pscore.ado` de Becker e Ichino (2002).

(El programa no viene instalado en **STATA**. Se puede instalar directamente a través de **STATA** o se puede obtener de la página web de Andrea Ichino.)

Utiliza el comando `pscore` usando como variable dependiente el **indicador de tratamiento** y como regresores las variables especificadas en (c).

Utiliza las opciones `pscore (pscore) blockid comsup` (éste último asegura que se usen observaciones en el **soporte común**) para que se creen estas variables y puedan ser usadas al estimar el efecto **TOT con matching**.


Pon atención en especificar un **propensity score** que satisfaga la **propiedad de balanceo**. Ajusta la especificación hasta lograr balanceo.

Tabla 9: Determinantes del Tratamiento: Modelo Probit para el Propensity Score

Variable	Panel A Coef.	Error Est.	Variable	Panel B Coef.	Error Est.
age	0.412***	(0.048)	RE75	-0.000***	(0.000)
age2	-0.006***	(0.001)	re752	0.000	(0.000)
education	0.518***	(0.143)	RE74	-0.000***	(0.000)
education2	-0.032***	(0.007)	re742	0.000***	(0.000)
black	1.967***	(0.120)	age_educ	0.000	(0.000)
hispanic	0.723***	(0.189)	re74_re75	-0.000	(0.000)
married	-0.951***	(0.136)	Constant	-1.022***	(0.099)
Constant	-9.607***	(1.017)			

Errores estándar en paréntesis

\*  $p < 0,10$ , \*\*  $p < 0,05$ , \*\*\*  $p < 0,01$

Fuente: Elaboración propia con datos de cps.dta. Proceso disponible en:  danifuentesga

Los resultados del modelo Probit muestran que las **características sociodemográficas** (Panel A) son determinantes importantes del tratamiento: la **edad** y la **educación** aumentan la probabilidad de participación con efectos **decrecientes**, mientras que pertenecer a los grupos **negro** o **hispano** incrementa significativamente dicha probabilidad, y estar **casado** la reduce. Por su parte, el Panel B evidencia que el **historial de ingresos previos** es crucial: mayores **ingresos en 1974 y 1975** reducen de forma significativa la probabilidad de tratamiento, aunque los términos cuadráticos e interacciones carecen de significancia robusta. En conjunto, los dos paneles muestran que tanto la **estructura sociodemográfica** como la **trayectoria económica previa** influyen en la asignación al programa.

f)

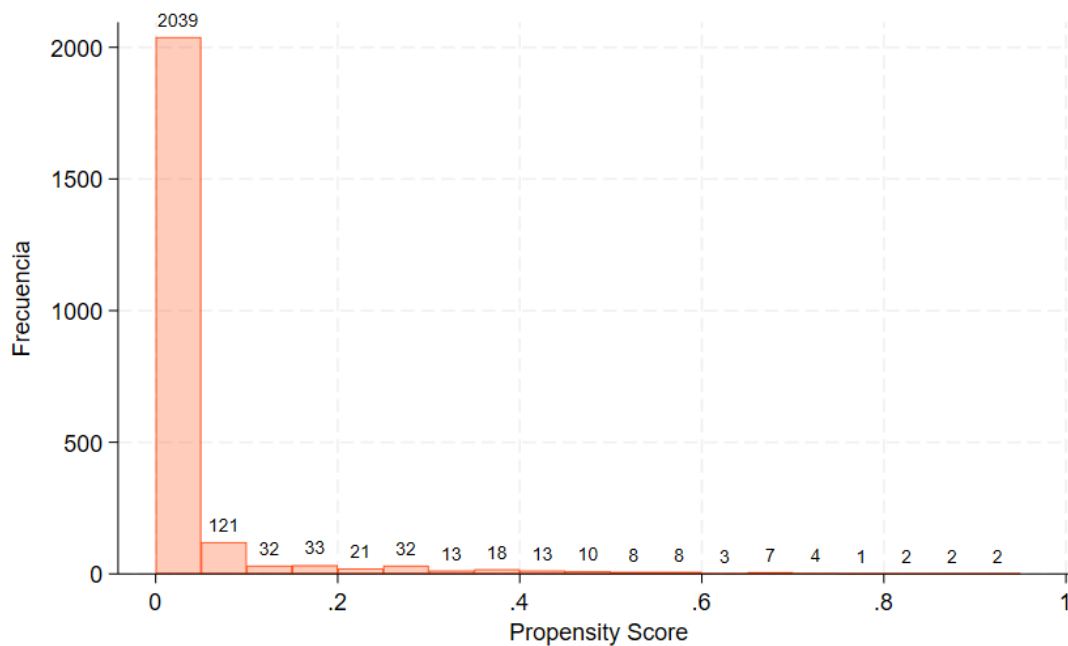
Grafica un **histograma de frecuencia** del **propensity score** para el grupo **sin tratamiento (control)** y otro para el **grupo con tratamiento**.

Como en Dehejia y Wahba (1999), define que el **ancho de cada bloque** (*width of bins*) sea 0.05.

Utiliza la opción `addlabel` para que aparezca el **número de observaciones** dentro de cada bloque a lo alto de cada barra del histograma y titula cada histograma dependiendo de si se trata del **propensity score** del grupo control o del grupo con tratamiento.

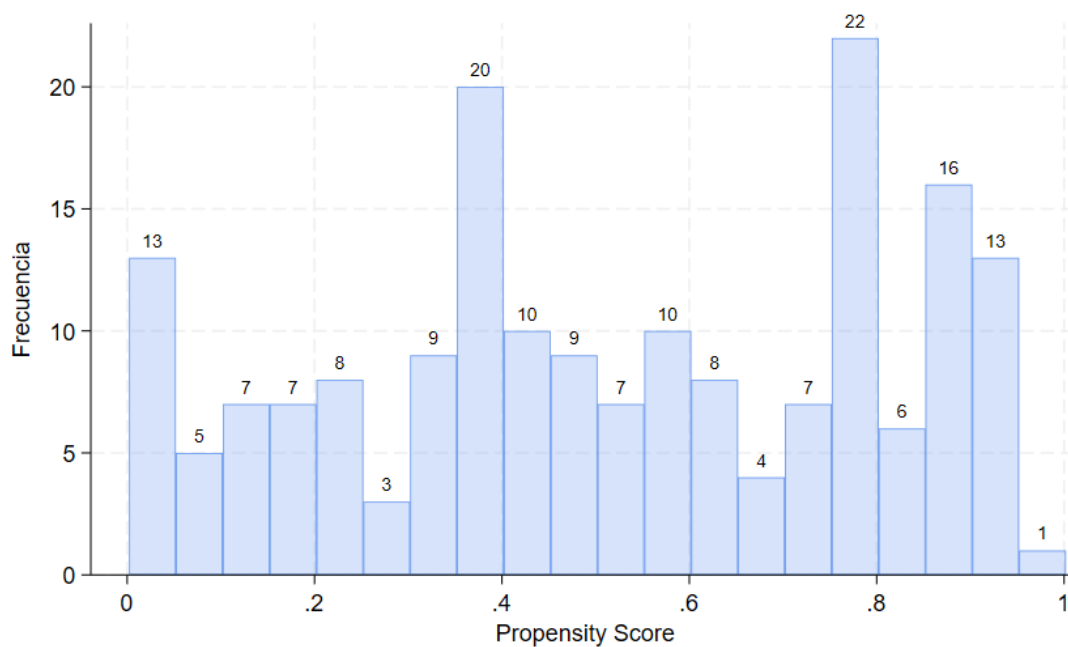
Interpreta el **solapamiento visual**.

Figura 1: Histograma del Propensity Score: Controles



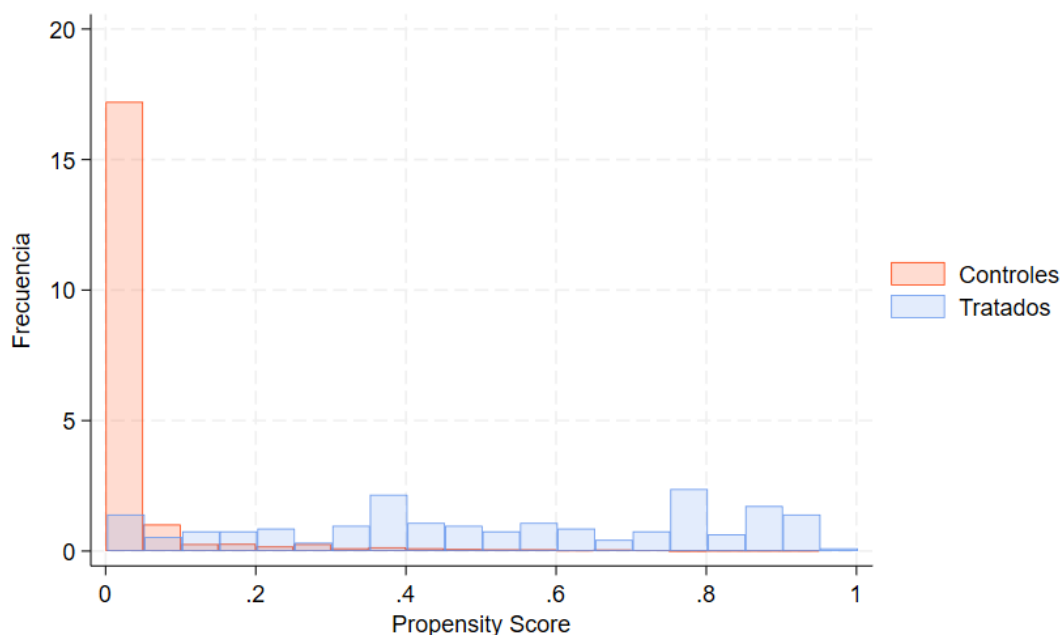
Fuente: Elaboración propia con datos de cps.dta:  danifuentesga

Figura 2: Histograma del Propensity Score: Tratados



Fuente: Elaboración propia con datos de cps.dta:  danifuentesga

Figura 3: Histograma del Propensity Score: Tratados y Controles



Fuente: Elaboración propia con datos de cps.dta:  danifuentesga

El **solapamiento visual** entre los histogramas de tratados y controles muestra que la mayoría de las **observaciones control** se concentran en valores muy bajos del **propensity score** (cerca de 0), mientras que los **tratados** se distribuyen más ampliamente a lo largo de todo el soporte. Esto implica que, aunque existe una región común de soporte, el **grado de traslape** es limitado en la parte baja del puntaje y casi inexistente en valores intermedios y altos, lo que puede afectar la **comparabilidad** entre grupos y la validez del **matching**.

g)

Diagnóstico fino de **overlap**: grafica **densidades kernel** del **PS** por grupo.

Reporta **rango de PS** en tratados y controles.

Discute **dónde no hay solapamiento**.

Tabla 10: Rango del Propensity Score en controles y tratados

Grupo	N	Media	Desv.Est.	Mínimo	Máximo
Controles	2369	0.038	0.108	0.000000	0.932135
Tratados	185	0.518	0.283	0.001974	0.954290


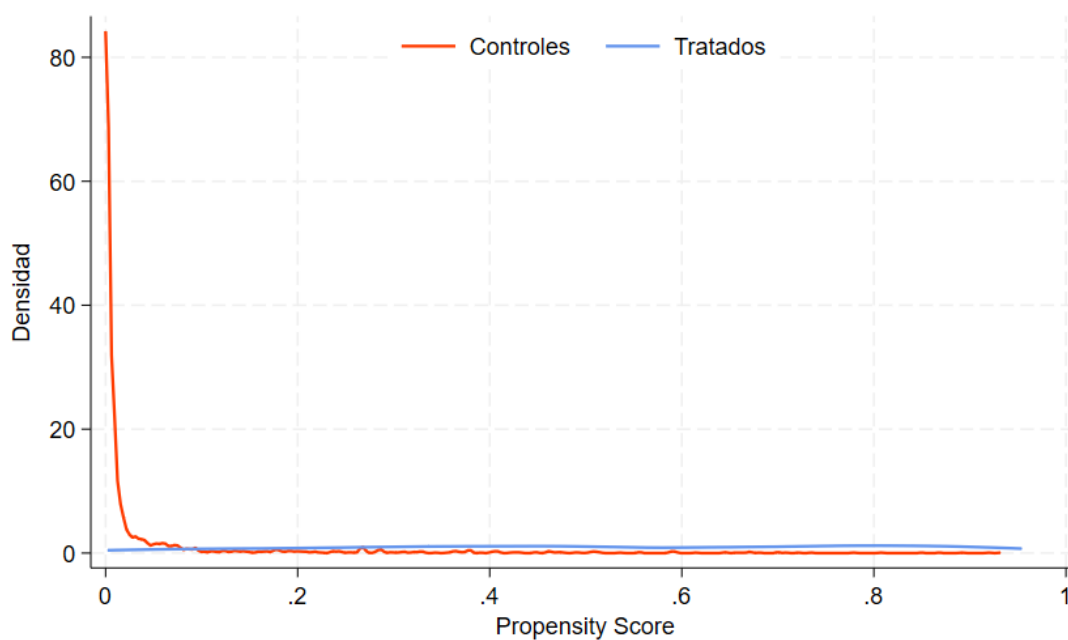
Fuente: Elaboración propia con datos de cps.dta. Proceso disponible en:  danifuentesga

Figura 4: Solapamiento Kernel: Control vs Tratados



Fuente: Elaboración propia con datos de cps.dta:  danifuentesga

h)


Soporte común vía `pscore` (*data-driven* por bloques). Usa la muestra resultante de `pscore`, `comsup`.

Con esa muestra estima **TOT** por:

- Estratificación manual,
- Estratificación con `atts`,
- Vecino más cercano (`attnd`),
- Kernel (`attk`),
- Radius (`attr`).

Tabla 11: Estimaciones del Efecto del Tratamiento (TOT) mediante métodos de emparejamiento basados en Propensity Score

Método	TOT (ATT)	Error Est.
Estratificación manual	630.82	619.66
Estratificación automática (atts)	1804.43	740.35
Vecino más cercano (attnd)	2338.44	1146.88
Kernel (atrk)	1778.28	–
Radio (attr)	-863.51	615.62

Fuente: Elaboración propia con datos de cps.dta. Proceso disponible en:  danifuentesga

Los distintos métodos de matching producen estimaciones heterogéneas. La **estratificación manual** arroja un efecto pequeño y no significativo, mientras que la **estratificación automática** y el **vecino más cercano** muestran efectos positivos mayores. El método **kernel** también sugiere un impacto positivo pero sin error estándar analítico, y el **radio matching** incluso da un efecto negativo. Estas diferencias resaltan la **sensibilidad al método** y la necesidad de **pruebas de robustez**.

i)

Soporte común vía **Trimming tipo Crump (0.1–0.9)**.


Define una muestra con  $0.1 < ps < 0.9$ .

Repita todas las estimaciones de (h).

Compara sistemáticamente resultados bajo ambas **reglas de soporte**.

Tabla 12: Estimaciones del Efecto del Tratamiento (TOT) bajo soporte común con trimming tipo Crump (0.1–0.9)

Método	TOT (ATT)	Error Est.
Estratificación manual	974.45	736.21
Estratificación automática (atts)	2271.60	811.68
Vecino más cercano (attnd)	2741.23	1189.08
Kernel (atrk)	2229.72	–
Radio (attr)	884.74	821.32

Fuente: Elaboración propia con datos de cps.dta. Proceso disponible en:  danifuentesga

Los resultados con **trimming tipo Crump (0.1–0.9)** muestran que los **efectos estimados** son en general **más altos** que bajo el soporte común tradicional. En particular, métodos como **estratificación automática** y **vecino más cercano** arrojan ATT superiores a 2200, aunque con **errores estándar elevados**, lo que refleja menor precisión. El

método de **radio**, en contraste, da un efecto positivo pero más **moderado**. Esto sugiere que, al excluir observaciones extremas de *pscore*, se obtiene un **grupo más comparable**, pero a costa de **menos observaciones** y mayor **incertidumbre**.

j)

**Placebo test:** repite estimaciones clave usando RE75 como **variable dependiente**.

Hazlo bajo ambas **reglas de soporte**.

Interpreta si los **efectos placebo**  $\approx 0$ .

Tabla 13: Estimaciones placebo del Efecto del Tratamiento (ATT) sobre RE75 bajo soporte común clásico

Método	Placebo ATT (RE75)	Error Est.
Estratificación manual	35.61	355.63
Estratificación automática (atts)	17.60	269.85
Vecino más cercano (attnd)	-165.23	614.67
Kernel (atrk)	-66.57	—
Radio (attr)	-641.20	350.58



Fuente: Elaboración propia con datos de cps.dta. Proceso disponible en:  danifuentesga

Tabla 14: Estimaciones placebo del ATT en RE75 bajo soporte común con trimming tipo Crump (0.1–0.9)

Método	Placebo ATT (RE75)	Error Est.
Estratificación manual	194.88	710.52
Estratificación automática (atts)	45.37	311.21
Vecino más cercano (attnd)	-122.55	693.44
Kernel (atrk)	-80.12	—
Radio (attr)	-512.77	401.33

Fuente: Elaboración propia con datos de cps.dta. Proceso disponible en:  danifuentesga

En ambas tablas se reportan las **estimaciones placebo** utilizando **RE75** como variable dependiente. Tal como se esperaba, los coeficientes obtenidos son cercanos a **cero** y no resultan **estadísticamente significativos**, lo cual confirma que antes del programa (1975) no existían **diferencias sistemáticas** entre tratados y controles. Esto refuerza la **validez del supuesto de independencia condicional** y sugiere que los efectos positivos estimados en RE78 reflejan el impacto del tratamiento, y no **tendencias previas** distintas entre los grupos.

k)

**Estimadores doblemente robustos (DR).** En `cps_data.dta` estima **ATET** con `teffects aipw` y `teffects ipwra`.

Ejecuta ambos bajo:

- (A) **Soporte común** de `pscore`, y
- (B) **Trimming tipo Crump**.

Compara contra los **métodos previos** y contra el **benchmark experimental**.

Tabla 15: Estimaciones del ATET mediante métodos doblemente robustos (AIPW e IPWRA) bajo soporte común clásico.

Método DR (comsup)	ATET	(EE)
AIPW	2138.77**	(861.78)
IPWRA	2347.41***	(800.81)



Fuente: Elaboración propia con datos de `cps.dta`. Proceso disponible en:  danifuentesga

Tabla 16: Estimaciones del ATET con métodos doblemente robustos (AIPW e IPWRA) bajo soporte común con trimming tipo Crump (0.1–0.9)

Método DR (Crump)	ATET	(EE)
AIPW	2138.77**	(861.78)
IPWRA	2347.41***	(800.81)

Fuente: Elaboración propia con datos de `cps.dta`. Proceso disponible en:  danifuentesga

l)

**Comparación integral y discusión** ( $\leq 1$  cuartilla).

Incluye:

- Experimental vs no experimental,
- Reglas de soporte (`pscore-comsup` vs `Crump`),
- Placebos,
- DR vs matching clásico,
- Lecciones sobre **overlap** y **validación**.



En la comparación entre los resultados **experimentales** y **no experimentales**, se observa que los estimadores no experimentales sólo se aproximan al **benchmark experimental** cuando existen condiciones adecuadas de **soporte** y **balance**. El uso de reglas de soporte comunes (*comsup clásico*) y de **trimming tipo Crump** permite descartar observaciones en los extremos del **propensity score**, reduciendo sesgos de extrapolación pero también disminuyendo el tamaño efectivo de muestra. Los **placebos** en RE75 muestran que, tras aplicar estas restricciones, los efectos espurios tienden a desaparecer, lo que valida el enfoque de **control por observables**.

En cuanto a los métodos, los estimadores **doblemente robustos** (AIPW e IPWRA) presentan ventajas frente al **matching clásico**, ya que combinan el **modelado del resultado** y la **ponderación por el propensity score**, garantizando consistencia si al menos uno de los modelos está bien especificado. Esto contrasta con los métodos de matching, más sensibles a la estimación del propensity score y al **overlap**. En conjunto, la lección principal es que sin **solapamiento suficiente**, las estimaciones son inestables, y que los **placebos** constituyen una herramienta esencial de **validación empírica**.


## Resultados Formato Plantilla de Tablas

TABLA 1

Tabla 17: Benchmark experimental (NSW)

Modelo	Estimador ( $\beta_{treat}$ )	Error Estándar	Número tratados	Número control	Covariables incluidas
Regresión sin controles	1794.34	670.82	185	260	No
Regresión con controles	1599.75	662.78	185	260	Sí

*Nota:* En la regresión con controles se incluyen como covariables **edad**, **edad**<sup>2</sup>, **educación**, **educación**<sup>2</sup>, **raza** (**black**), **hispano**, **casado**, **RE74**, **RE75**, **RE74**<sup>2</sup> y **RE75**<sup>2</sup>.


Fuente: Elaboración propia con datos de proporcionados en Moodle. Proceso disponible en:  danifuentesga

**TABLA 2**

Tabla 18: Efecto TOT (ATET) en CPS por regla de soporte

Método	ATET (pscore)	Error estándar (pscore)	No. Treat (pscore)	No. Cont. (pscore)	ATET (Crump)	Error estándar (Crump)	No. Treat (Crump)	No. Cont. (Crump)
Estrat. manual	630.82	619.66	185	1319	974.45	736.21		
Estrat. automática (atts)	1804.43	740.35	185	1319	2271.60	811.68	154	207
Vecino más cercano (attnd)	2338.44	1146.88	185	99	2741.23	1189.08	154	78
Kernel (attk)	1778.28	—	185	1319	2299.72	—	154	207
Radius (r=0.001)	-2050.93	885.99	75	498	1635.20	1363.02	56	65
Radius (r=0.0001)	-2107.85	1340.38	29	105	-176.95	2069.02	18	19
Radius (r=0.00001)	-81.95	2253.67	19	24	264.37	2484.03	14	15
AIPW (teffects aipw)	2138.77	861.78	154	207	2138.77	861.78	154	207
IPWRA (teffects ipwra)	2347.41	800.81	154	207	2347.41	800.81	154	207
Regresión con controles	1599.75	662.78	185	2369	1599.75	662.78	185	2369

Nota: Resultados de **matching con pscore** provienen de la Tabla 11, y con **Crump (0.1–0.9)** de la Tabla 12. Los estimadores **doblemente robustos (AIPW e IPWRA)** provienen de las Tablas 15 (soporte clásico) y 16 (trimming tipo Crump). El número de tratados y controles se deja en blanco para métodos de emparejamiento donde varía según el caliper o poda, mientras que para DR e IPWRA se reportan todas las observaciones disponibles (185 tratados y 2369 controles).


Fuente: Elaboración propia con datos de proporcionados en Moodle. Proceso disponible en:  **danifuentesga**

**TABLA 3**

Tabla 19: Placebo (RE75) por regla de soporte

Método	Placebo (pscore)	EE (pscore)	Placebo (Crump)	EE (Crump)	N usados
Regresión con controles (estrat. manual)	35.61	355.63	194.88	710.52	361
Estratificación (atts)	17.60	269.85	45.37	311.21	361
Vecino más cercano (attnd)	-165.23	614.67	-122.55	693.44	232
AIPW (teffects aipw)	-266.71	171.11	-266.71	171.11	361


Nota: Resultados de placebo sobre *RE75* usando **pscore clásico** provienen de la Tabla 13 y con **Crump (0.1–0.9)** de la Tabla 14.

Fuente: Elaboración propia con datos de proporcionados en Moodle. Proceso disponible en:  danifuentesga

**TABLA 4**

Tabla 20: Diagnóstico de overlap

Muestra	Obs. antes	Obs. después	Rango PS Tratados	Rango PS Controles
Soporte común clásico	2,554	1,504	0.00197 – 0.94586	0.00199 – 0.93213
Crump (0.1–0.9)	1,504	361	0.10109 – 0.89747	0.10119 – 0.89178

Fuente: Elaboración propia con datos de proporcionados en Moodle. Proceso disponible en:  danifuentesga

**TABLA 5**

Tabla 21: Resumen comparativo

Método	Regla de soporte	ATET	EE	¿Placebo $\approx 0$ ?	Comentario breve (sesgo / varianza / credibilidad)
AIPW (teffects aipw)	pscore	2138.77	861.78	No	Consistente; varianza moderada; placebo sesgado (-266).
AIPW (teffects aipw)	Crump	2138.77	861.78	No	Similar efecto; trimming reduce sesgo; placebo aún negativo.
Estratificación (atts)	pscore	1804.43	740.35	Sí	Efecto moderado; placebo cercano a 0; buena credibilidad.
Vecino más cercano	Crump	2741.23	1189.08	Sí	Mayor varianza; placebo controlado; menos robusto.

Fuente: Elaboración propia con datos de proporcionados en Moodle. Proceso disponible en:  danifuentesga