

Laboratorio de Cómputo

Pareamiento por Puntaje de Propensión

Emmanuel Chávez

Centro de Investigación y Docencia Económicas
Microeconometría
7° Semestre - Licenciatura en Economía

Otoño 2023

Efectos de programas de capacitación

Dehejia, R. H., & Wahba, S. (1999). Causal Effects in Nonexperimental Studies: Reevaluating the Evaluation of Training Programs, *Journal of the American Statistical Association*.

- NSW Demonstration: programa de capacitación laboral aplicado en EUA entre 1975 y 1977.
- El programa se aplicó de manera aleatoria. Esto permite obtener un estimador insesgado de su efecto sobre el ingreso.
- Objetivo de Dehejia, R., & Wahba, S. (1999): comprobar qué tan confiable es el método de PSM, tomando como referencia al estimador insesgado del experimento aleatorio.

- La base `nsw_dw_new.dta` tiene observaciones del experimento controlado aleatorio del programa de capacitación NSW Demonstration.
- La base `psid_controls_new` tiene individuos que recibieron el programa de capacitación NSW Demonstration, y una muestra de individuos (tomada de una encuesta) que no recibieron el programa de capacitación.

A entregar

- Con la base `nsw_dw_new.dta`, hacer una tabla que contenga estimadores regresiones de MCO que muestren la relación entre haber recibido el programa de capacitación y el ingreso de 1978. Hacer tres especificaciones:
 - 1) sin controles; 2) controles I: edad, educación y estado civil; 3) controles II: las variables anteriores, más ser de raza negra, ser hispano, no tener diploma, el ingreso de 1974 y el ingreso de 1975.
 - Para las tres regresiones, incluir estimador del programa de capacitación, el error estándar, el nivel de significancia (con asteriscos), los intervalos de confianza al 95%, y el número de observaciones. No es necesario incluir el coeficiente de los controles.
- Con la base `psid_controls_new`, hacer una tabla que contenga:
 - La diferencia de medias incondicional del ingreso de 1978 entre los tratados y los no tratados. Incluir el error estándar de la media y el nivel de significancia (con asteriscos).
 - El estimador del efecto de haber recibido el curso de capacitación bajo el método de propensity score matching:
 - Construir el score de dos formas: 1) con la edad, educación y estado civil; 2) con las variables anteriores más ser de raza negra, ser hispano, no tener diploma, el ingreso de 1974 y el ingreso de 1975. Incluir coeficiente de estimador de PSM, su error estándar y el nivel de significancia (con asteriscos).
- Al comparar los estimadores de las regresiones MCO y de los PSM, ¿qué se puede concluir sobre el estimador de PSM?

Comandos a tener en cuenta

- *regress*: hace una regresión de MCO.
- *ttest*: hace tests de medias.
- *psmatch2*: estima con el método de propensity score matching.

Resultados: Regresiones de MCO

Variable dependiente: ingreso en 1978			
	Sin controles	Con controles I	Con controles II
Participar en programa de capacitación	1,794*** (633) 551 - 3,038	1,655*** (633) 411 - 2,900	1,676*** (639) 421 - 2,932
Observaciones	445	445	445

Nota: Errores estándar en paréntesis. Intervalos de confianza debajo de los errores estándar. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Controles I: edad, educación y estado civil. Controles II: Controles I más raza negra, ser hispano, no tener diploma, el ingreso de 1974 y el ingreso de 1975.

Resultados: Propensity Score Matching

	Media Incondicional	Estimador PSM controles I	Estimador PSM controles II
$Y_1 - Y_0$	-15205*** (1155)	-3350** (1459)	1546 (1746)

Nota: Errores estándar en paréntesis. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Controles I: edad, educación y estado civil. Controles II: Controles I más raza negra, ser hispano, no tener diploma, el ingreso de 1974 y el ingreso de 1975.

Respuesta a la pregunta

- Al comparar los estimadores de las regresiones MCO y de los PSM, ¿qué se puede concluir sobre el estimador de PSM?
- **Respuesta:** El estimador de MCO con los datos del experimento aleatorio es muy robusto ante diferentes especificaciones de la regresión, ya que no cambia significativamente al añadir o quitar variables de control. Esto es de esperarse, ya que las regresiones se hacen con datos de un experimento aleatorio. Y si el RCT tiene validez interna, el estimador por MCO es insesgado.

Por otro lado, los estimadores PSM con datos que no se obtienen de un experimento aleatorio, sino de una encuesta son altamente sensibles a la especificación utilizada. Según las variables de control que se utilicen para construir el propensity score, el estimador cambia enormemente (incluso de signo). La alta sensibilidad del PSM al conjunto de controles X_i indica que este método no siempre es confiable, ya que es crucial encontrar el conjunto exacto de variables X_i que garantice la independencia condicional. Si no tenemos un experimento aleatorio contra el cual comparar el PSM, nos quedará la duda de si nuestro conjunto X_i elegido es el óptimo.