

PROPENSITY SCORE-MATCHING METHODS FOR NONEXPERIMENTAL CAUSAL STUDIES

Análisis del paper: Rajeev H. Dehejia and Sadek Wahba

José Palma González

Universidad de Santiago de Chile

Magister en Ciencias Económicas

Tópicos de microeconometría

Índice

- [Introducción](#)
- [Métodos de emparejamiento](#)
- [Algoritmos de propensity score matching](#)
- [Data](#)
- [Resultados de emparejamiento](#)
- [Testeo](#)
- [Conclusiones](#)

Introducción

- Se discuten métodos de emparejamiento del propensity score y los problemas de dimensionalidad.
- 2 Enfoques:
 - Información demasiado costosa de conseguir.
 - El proceso de búsqueda para un mejor subgrupo del grupo de control para revelar la extensión del overlap.

Métodos de emparejamiento

- El rol de la randomización
 - El problema básico al identificar el efecto causal se ve ya sea bajo tratamiento o bajo control, pero nunca ambos. El efecto básico del tratamiento de interés bajo ajustes no experimentales es el efecto esperado del tratamiento para la población tratada.
- Emparejamiento exacto en covariables
 - Cuando tenemos una falta de unidades de control para experimentar, asumimos que la información puede ser obtenida para un grupo unidades potenciales de comparación, que no son necesariamente sacadas de la misma población que los tratados donde observamos el set de covariables X .

Métodos de emparejamiento

- Propensity Score y reducción de dimensionalidad
 - Rubin propone usar el propensity score para reducir la dimensionalidad del problema del emparejamiento.

Algoritmos de propensity score matching

- Asumimos que la propensity score es conocida, pero en la realidad no lo es.
- Emparejar con propensity score es esencialmente un esquema de peso, lo que determina el peso en las unidades de comparación al estimar los efectos del tratamiento.
- Este estimador sigue la propuesta 3. Las expectativas son reemplazadas por medias muestrales y condicionamos $p(X_i)$ al emparejar cada unidad del tratamiento i a un grupo de unidades de comparación j , con un propensity score similar. El objetivo se convierte en emparejar tratados con unidades de comparación cuyos propensity score son lo suficientemente cercanos para condicionar $p(x)$ a la propuesta 3 para ser válida.

Algoritmos de propensity score matching

- Ocurren 3 problemas al emparejar:
 - Si emparejar o no con un reemplazo.
 - Cuántas unidades de control emparejar para cada unidad de tratados.
 - Qué método de emparejamiento usar.

Data

- El programa “National Supported Work”
 - Era un programa federal que proveía experiencia laboral a aquellos que habían enfrentado problemas económicos y sociales antes de entrar al programa. Los candidatos para el experimento fueron seleccionados por un criterio elegible, y después se asignó (o no) al tratamiento.
- Distribución de las muestras de tratados y de control
 - Las muestras de la PSID y de la CPS son 8 o 9 años mas viejas que las del grupo NSW, además contienen su composición étnica y si tienen educación secundaria completa, mientras que el NSW contaba si eran dropouts de secundaria, e incluso, las ganancias pre tratamiento son mucho mas altas para los de control que de tratamiento.

Data

FIGURE 1.—HISTOGRAM OF ESTIMATED PROPENSITY SCORE,
NSW AND CPS

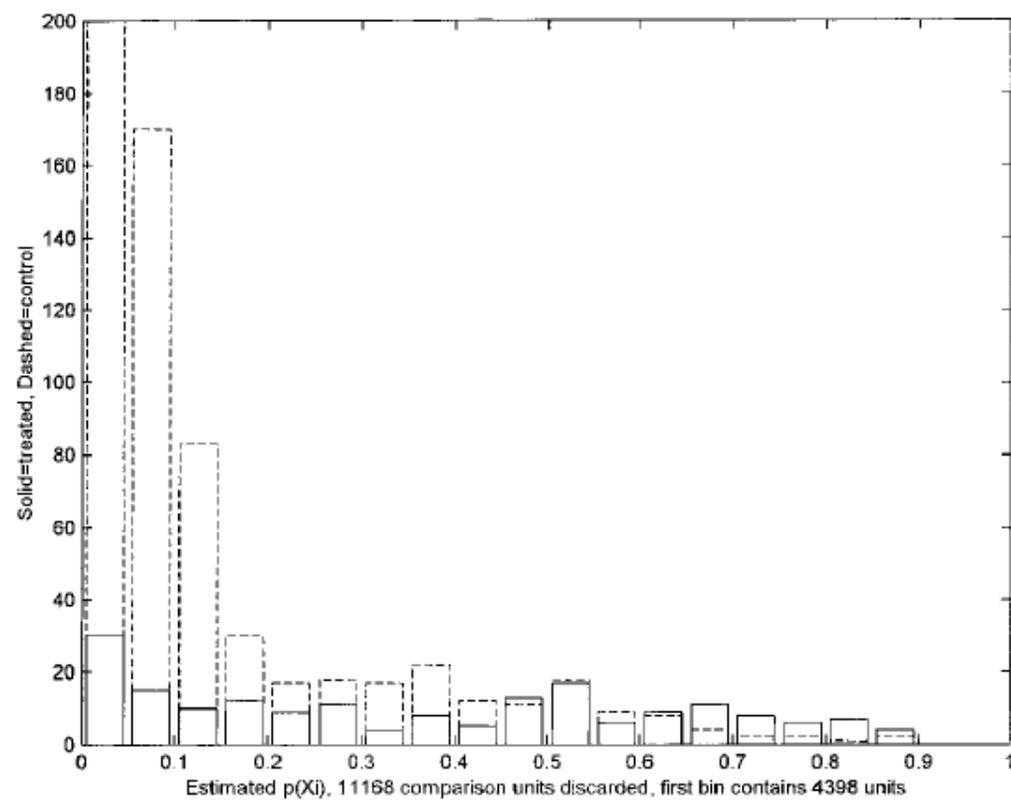
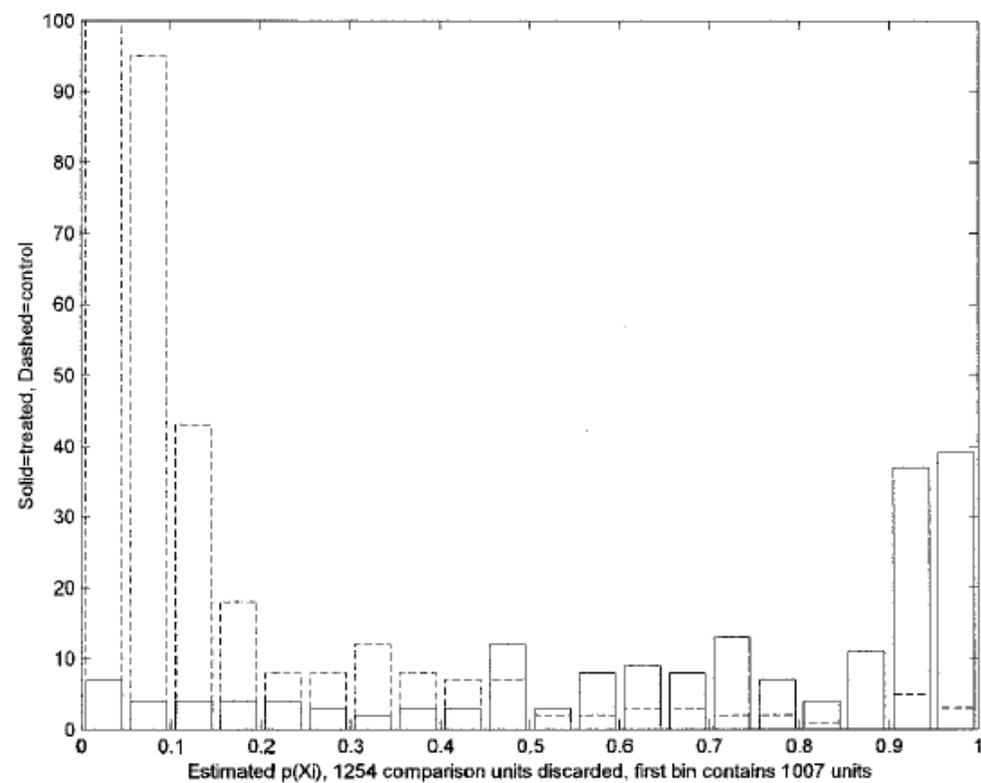


FIGURE 2.—HISTOGRAM OF ESTIMATED PROPENSITY SCORE,
NSW AND PSID



Resultados del emparejamiento

FIGURE 3.—PROPENSITY SCORE FOR TREATED AND MATCHED COMPARISON UNITS, RANDOM WITHOUT REPLACEMENT

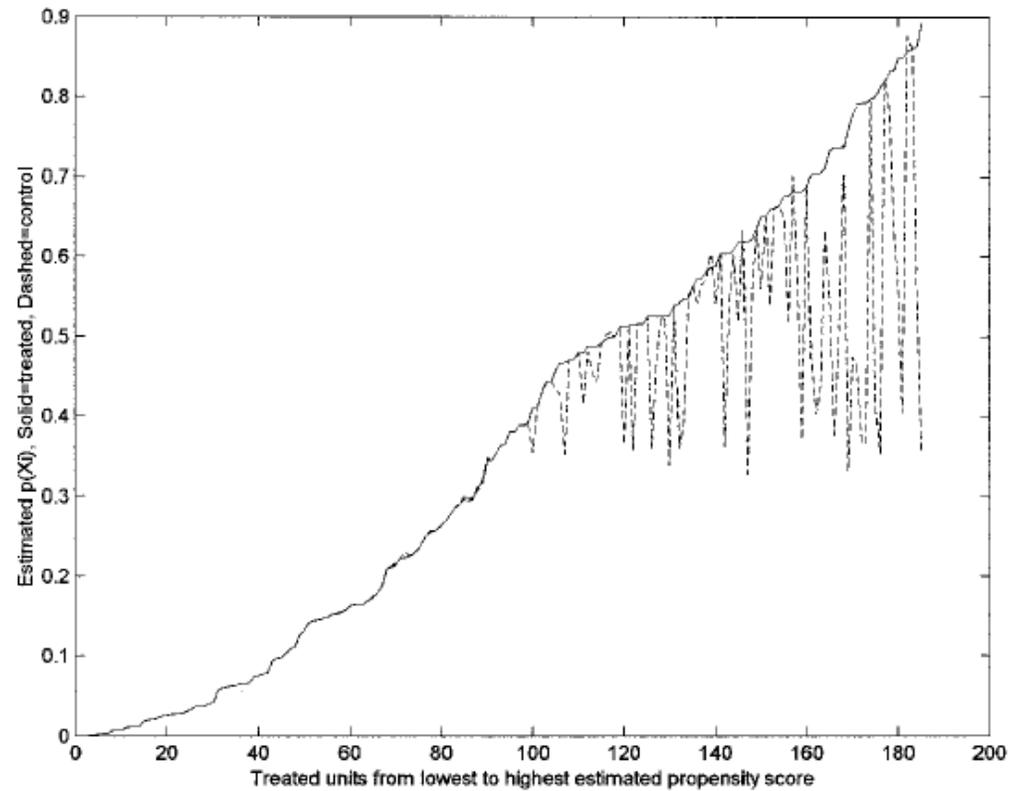
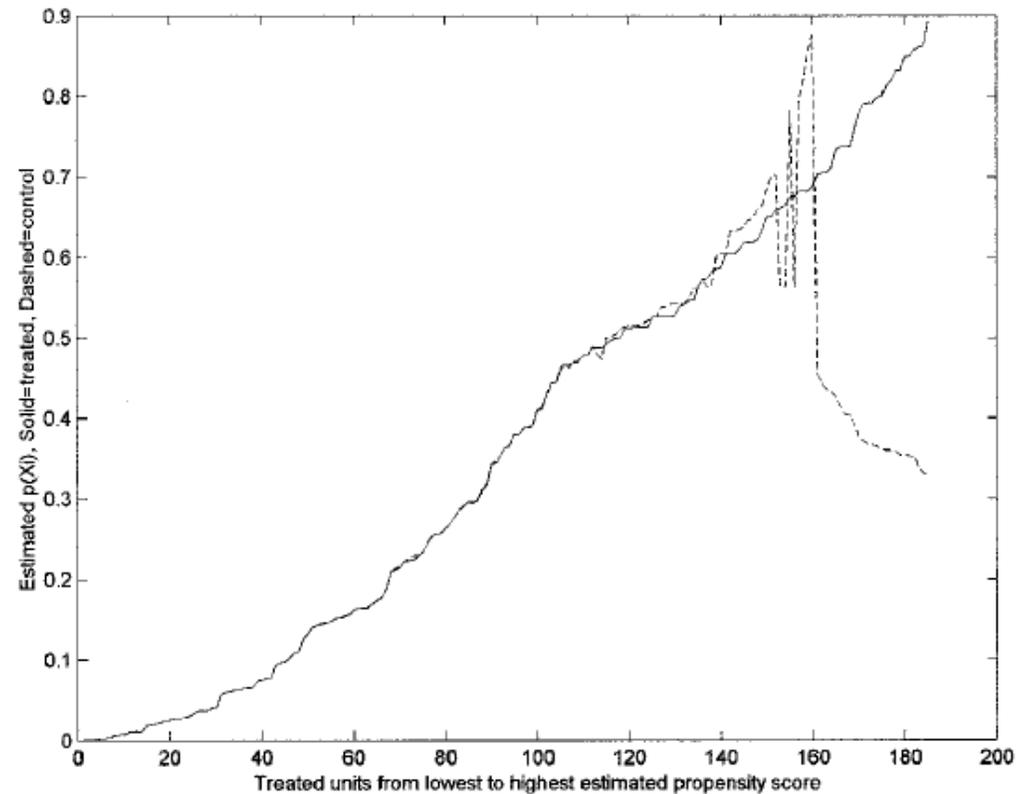


FIGURE 4.—PROPENSITY SCORE FOR TREATED AND MATCHED COMPARISON UNITS, LOWEST TO HIGHEST



Resultados del emparejamiento

FIGURE 5.—PROPENSITY SCORE FOR TREATED AND MATCHED COMPARISON UNITS, HIGHEST TO LOWEST

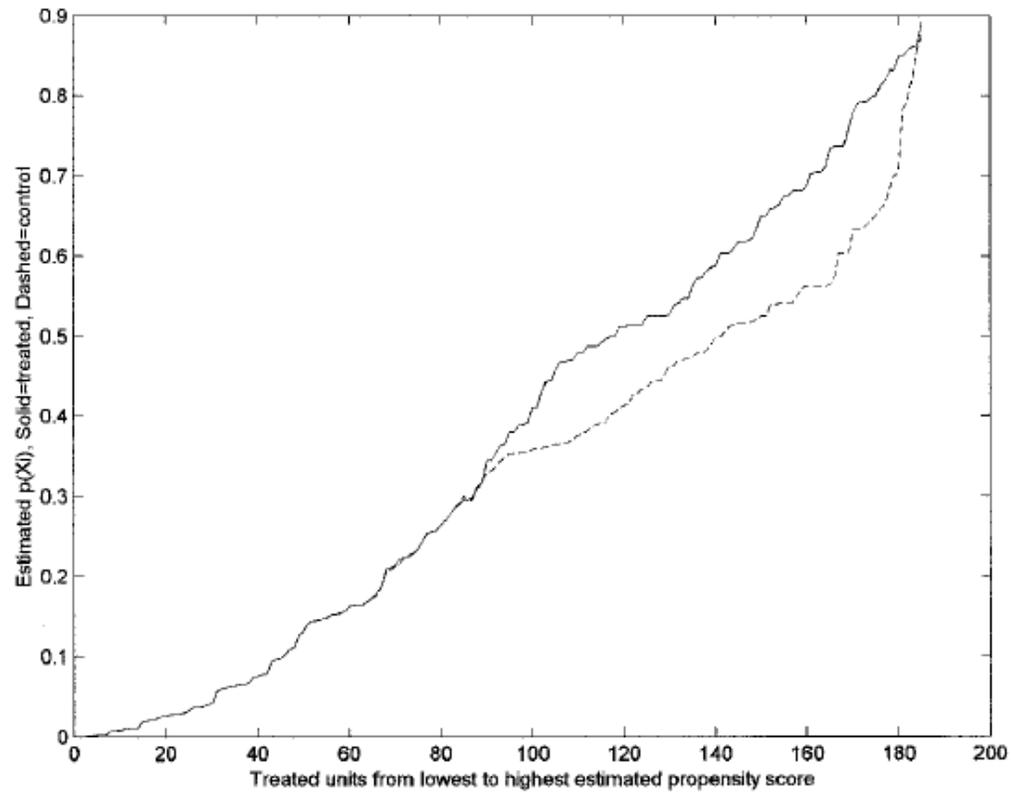
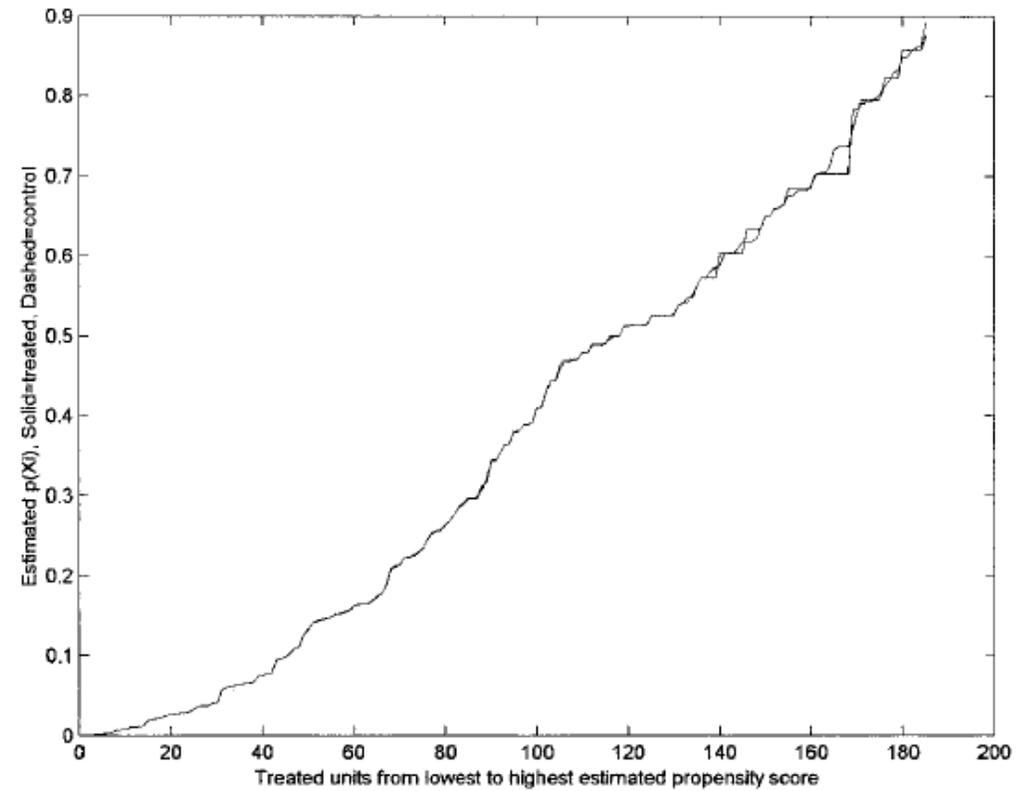


FIGURE 6.—PROPENSITY SCORE FOR TREATED AND MATCHED COMPARISON UNITS, NEAREST MATCH



Testeo

- Supuesto de la prueba del emparejamiento
 - La estructura especial de la información que usamos nos permite probar el supuesto bajo el propensity score matching, porque tenemos un grupo experimental de control y 2 grupos no experimentales de comparación. Podemos probar el supuesto que, al condicionar el propensity score, las ganancias en un estado no tratado son independientes de la asignación del tratamiento.
 - Los estimadores sesgados serian las ganancias para el grupo de control experimental menos las ganancias para el grupo de comparación no experimental condicionado al propensity score estimado.
 - Tanto para el CPS y el PSID, vemos un rango de estimadores sesgados que son grandes para valores bajos del propensity score estimado. Ellos representan al grupo que eran menos probables de estar en el grupo de tratamiento.
 - Los estimadores sesgados no son estadísticamente significativos.

Testeo

- Probar la sensibilidad especificando propensity score
 - Una limitación potencial de los métodos del propensity score es tener que estimar el propensity score.
 - En el caso que investigador que no tenga el beneficio “experimental benchmark estimate” preferirían estimadores que especifiquen completamente y de manera exitosa todas las covariables observadas.

Conclusiones

- La calidad del estimador que surge del resultado de comparar esta limitado generalmente por la calidad del grupo de comparación usado.
- La conclusión del análisis es que es extremadamente valioso revisar si es comparable el tratamiento y unidades de comparación en términos de las características pre tratamiento.
- El método del propensity score realmente remarca el hecho que la mayoría de las unidades de comparación son muy diferentes de las unidades tratadas.

Alguna Reflexiones

- Fortalezas

- Existe 2 grupos de comparación para la NSW, lo que puede ayudar a la investigación.
- Posee una gran cantidad de formas de estimar los resultados.
- Los métodos pueden estimar de manera acertada los efectos del tratamiento en un contexto no experimental

- Posibles criticas

- Un problema importante es si el supuesto de variables observables son validas o si el proceso depende de las variables que no son observables.
- No todos los estimadores son realmente útiles.
- En la practica no se pueden hacer ciertas pruebas.

Referencias

- “Causal Effects in Non-Experimental Studies: Re-Evaluating the Evaluation of Training Programs,” *Journal of the American Statistical Association* 94:448 (December 1999), 1053–1062.
- LaLonde, Robert, “Evaluating the Econometric Evaluations of Training Programs,” *American Economic Review* 76:4 (September 1986), 604–620.
- Rosenbaum, P., and D. Rubin, “The Central Role of the Propensity Score in Observational Studies for Causal Effects,” *Biometrika* 70:1 (April 1983), 41–55.
- Entre otros

PROPENSITY SCORE-MATCHING METHODS FOR NONEXPERIMENTAL CAUSAL STUDIES

Análisis del paper: Rajeev H. Dehejia and Sadek Wahba

José Palma González

Universidad de Santiago de Chile

Magister en Ciencias Económicas

Tópicos de microeconometría