

Los programas de formación y su capacidad para reducir el tiempo de búsqueda de un empleo

José Manuel Cansino Muñoz – Repiso
Universidad de Sevilla
jmcansino@us.es

Antonio Sánchez Braza
Universidad de Sevilla
asb@us.es

RESUMEN

Después de más de 28 años de experiencia en España en la gestión de programas de formación para jóvenes desempleados, los efectos de este tipo de programas públicos permanecen sin evaluar mediante investigaciones basadas en métodos observacionales. Este artículo estima el efecto promedio del Programa de Escuelas Taller y Casas de Oficios sobre el periodo necesario para encontrar un empleo, el cual se considera como una variable respuesta adecuada para evaluar la eficacia del programa.

Los autores presentan dos tipos de estimadores: el estimador de emparejamiento o “matching” y el estimador basado en la ponderación sobre la probabilidad condicional de participación “propensity score”.

Para los participantes en el programa, el número de días hasta encontrar un empleo se reduce en 467 y 471, respectivamente, dependiendo del estimador utilizado. Se encuentra, por tanto, evidencia consistente entre los métodos empleados acerca de que el programa reduce el periodo de tiempo necesario para encontrar un empleo.

Palabras clave: evaluación de programas de formación, políticas públicas, “matching”, “propensity score”.

- 1. Introducción**
- 2. El programa de formación de las Escuelas Taller y Casas de Oficio**
- 3. La base de datos**
- 4. Características del modelo**
- 5. Estimación del efecto promedio: resultados empíricos**
- 6. Conclusiones**
- 7. Referencias utilizadas**

1. Introducción

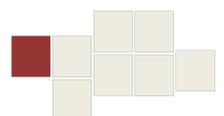
De acuerdo con Heckman, Clements y Smith (1997), generalmente se acepta que los programas sociales tienen un impacto diferente en función de las distintas características que presentan los individuos participantes. La asignación adecuada de los individuos entre un conjunto de programas alternativos puede convertirse en una cuestión crucial en las decisiones políticas.

Controlando las características diferenciadoras (covariables) de los individuos y obteniendo información del efecto promedio que un determinado programa provoca sobre una variable respuesta de interés, el decisor público puede determinar qué programa provocaría el mejor resultado sobre los participantes potenciales, precisamente teniendo en cuenta el efecto medio de los programas sobre subpoblaciones. Este efecto promedio sería estimado de manera previa a la asignación definitiva. El bienestar social puede aumentar si los decisores públicos siguen una regla de asignación que les permita determinar qué individuos deben asignarse a qué programas. Manski (2001) teorizó sobre esta cuestión para el caso de un conjunto finito de tratamientos o programas. Cansino y Román (2007) han explorado las posibilidades de esta cuestión para el caso del Tribunal de Cuentas español.

El objetivo de este artículo es estimar el efecto promedio de un tratamiento binario sobre una variable respuesta escalar a partir de dos tipos de estimadores típicamente utilizados en la literatura especializada, estimando la eficacia de este tratamiento a través de un análisis microeconómico. Dicho tratamiento es el Programa de Escuelas Taller y Casas de Oficios desarrollado en la provincia de Sevilla entre 1997 y 1999.

Los estimadores utilizados son los de emparejamiento y ponderación sobre la probabilidad condicionada de participación.

En primer lugar, se obtiene el estimador de emparejamiento sobre la probabilidad condicionada de participación, que es uno de los más utilizados en los trabajos empíricos realizados en este ámbito. En segundo lugar, se obtiene el estimador mediante la ponderación sobre dicha probabilidad condicionada, lo que permite la comparación de los resultados obtenidos por ambos métodos y verificar su consistencia.



Se ha seleccionado la provincia de Sevilla por ser la zona donde más ampliamente se ha desarrollado este programa de formación en el ámbito de Andalucía hasta el momento. La evaluación se lleva a cabo estimando el efecto promedio del programa sobre el tiempo que han necesitado los individuos de la muestra para encontrar un empleo. Los individuos son personas desempleadas con edades entre los 16 y los 25 años.

La dificultad para el acceso a los microdatos ha obstaculizado en España este tipo de evaluaciones, ampliamente desarrolladas en países como Francia, Alemania, Reino Unido y EE.UU¹. En relación con los participantes en las políticas formativas, no existen datos que puedan considerarse completos sobre el número de asistentes a cursos de formación no reglada en España. La corta, pero también heterogénea, duración de los cursos impide que estadísticas como la EPA capten los sucesos formativos con el rigor suficiente (Mato y Cueto, 2008).

Este artículo contribuye a la literatura en el sentido de que no abundan las investigaciones de tipo microeconómico sobre las políticas activas del mercado de trabajo. Aparte de los trabajos de Mato (2002), Arellano (2005), Mato y Cueto (2008) y las aportaciones de Cansino y Sánchez (2008a) y (2008b), no han proliferado evaluaciones similares en España basadas en métodos observacionales. Sobre la versión preliminar de Cansino y Sánchez (2008b), este artículo añade el uso del estimador de emparejamiento y contextualiza las conclusiones en investigaciones desarrolladas en otros países.

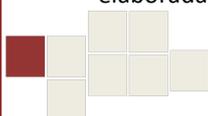
El artículo se estructura en seis epígrafes. Tras la introducción, en el epígrafe 2 se describe detalladamente el programa de formación evaluado. En el epígrafe 3 se expone la información contenida en la base de datos utilizada, mientras que en el epígrafe 4 se describen el modelo y las variables que en él intervienen. En el epígrafe 5 se procede a la aplicación de las técnicas de evaluación propuestas sobre las observaciones muestrales contenidas en la base de datos para la estimación del efecto promedio del programa de formación. El epígrafe 6 contiene las principales conclusiones.

2. El programa de formación de las Escuelas Taller y Casas de Oficio

El programa de formación de las Escuelas Taller y Casas de Oficios se creó en 1985, como una medida de Fomento de Empleo Juvenil² con la finalidad de fomentar el aprendizaje de un oficio o profesión. Este programa comenzó a desarrollarse con carácter experimental y con ámbito nacional bajo la dirección del Instituto Nacional de Empleo. Posteriormente, se llevó a cabo un proceso progresivo de transferencia de las competencias de su gestión a las Comunidades Autónomas, pasando a desarrollarse por los Servicios Públicos Regionales de Empleo de cada Comunidad Autónoma.

¹ Véanse Card y Sullivan (1988) y Manski y Garfinkel (1992) en EE.UU, Bonnall, Fougère y Sérandon (1997) en Francia, Andrews, Bradley y Upward (1999) y Blundell et al. (2004) en el Reino Unido, Bergemann, Fitzenberger y Speckesser (2005) en Alemania y Park et al. (1996) en Canadá, entre otras.

² Para una justificación del programa nos remitimos a la memoria sobre la situación socioeconómica y laboral elaborada por el Consejo Económico y Social (1995, p. 296).



En la actualidad, este programa se rige por una Orden del Ministerio de Trabajo y Asuntos Sociales³, y en el ámbito regional andaluz, por una Orden de la Consejería de Empleo y Desarrollo Tecnológico⁴. La provisión del programa se realiza de forma descentralizada en diferentes centros formativos mediante convocatorias públicas.

En definitiva, el objetivo de este programa es incorporar al mercado de trabajo a jóvenes desempleados menores de veinticinco años, mediante su cualificación y profesionalización, alternando la formación teórica y la práctica profesional a través de su participación en la realización de obras y servicios de utilidad pública. Los participantes en este programa deberán haber cumplido 16 años y ser menores de 25 años, estar desempleados e inscritos en los servicios públicos de empleo y cumplir los requisitos establecidos para formalizar un contrato para la formación.

La implementación del programa se realiza en dos etapas, una primera de carácter formativo y una segunda etapa de formación en alternancia con el trabajo y la práctica profesional. En el caso de las Escuelas Taller, la duración de ambas etapas no será inferior a un año ni superior a dos, dividida en fases de seis meses, mientras que en el caso de las Casas de Oficios, la duración de cada una de estas etapas queda establecida en seis meses.

En la *Tabla 1* se recoge la evolución en toda España del número de proyectos dentro del Programa de Escuelas Taller y Casas de Oficios a lo largo del periodo 1995-2003. Asimismo, en la *Tabla 2* se recoge la evolución del número de participantes en estos proyectos de Escuelas Taller y Casas de Oficios a lo largo del periodo 1995-2003, alcanzándose el máximo en el año 1998 con 60450 alumnos.

Programas	1995	1996	1997	1998	1999	2000	2001	2002	2003
Escuelas Taller	912	949	941	1.043	1.021	1.106	973	971	907
Casas de Oficio	360	450	347	468	458	374	356	288	194
Total	1.272	1.399	1.288	1.511	1.479	1.480	1.329	1.259	1.101

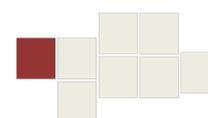
Fuente: INEM. Elaboración: J. M. Cansino y A. Sánchez, 2009.

Programas	1995	1996	1997	1998	1999	2000	2001	2002	2003
Alumnos de Escuelas Taller	37.419	39.883	40.143	44.636	42.909	44.868	36.888	36.404	32.350
Alumnos de Casas de Oficio	11.138	13.544	10.336	15.814	15.211	10.214	9.436	7.345	7.065
Total	48.557	53.427	50.479	60.450	58.120	55.082	46.324	43.749	39.415

Fuente: INEM. Elaboración: J. M. Cansino y A. Sánchez, 2009.

³ O. M. del Ministerio de Trabajo y Asuntos Sociales, de 14 de noviembre de 2001, por la que se regulan el programa de Escuelas Taller y Casas de Oficios y las Unidades de Promoción y Desarrollo y se establecen las bases reguladoras de la concesión de subvenciones públicas a dichos programas.

⁴ Orden de 8 de marzo de 2004, de la Consejería de Empleo y Desarrollo Tecnológico, por la que se regulan los Programas de Escuelas Taller, Casas de Oficios, Talleres de Empleo y Unidades de Promoción y Desarrollo y se establecen las bases reguladoras de la concesión de ayudas públicas a dichos programas.



Aunque la diversidad de agentes y de programas dificulta la obtención de series fiables de participantes en iniciativas de formación para desempleados, Mato y Cueto (2008, p. 62), a partir de datos de la OCDE, manejan la cifra de 500.000 personas desempleadas/año en formación durante los últimos años noventa. Por tanto, para el trienio 1997-1999, los alumnos de las Escuelas Taller y Casas de Oficios representaban el 11'26% de las personas desempleadas en formación.

3. La base de datos

La base de datos utilizada para la evaluación (*BASEVAFOR*) ha sido construida a partir de observaciones de individuos participantes en los programas de formación desarrollados por las Escuelas Taller y Casas de Oficios en la provincia de Sevilla a lo largo de la década de los años noventa.

De entre todos los participantes, se ha seleccionado a aquellos individuos que han participado en un programa de los concluidos a lo largo del año 1999, último año con datos desglosados disponibles cuando se inició la construcción de esta base de datos. El número total de participantes que forman este grupo es 1528 individuos. De entre estos individuos que constituyen la población objetivo, se ha seleccionado de forma aleatoria una muestra compuesta por 150 individuos que constituyen el grupo de participantes.

De forma paralela, se ha seleccionado también un grupo de control formado por 75 individuos, con características similares a los del grupo de participantes. El grupo de control ha sido elaborado para esta evaluación por la Delegación Provincial de Sevilla del Servicio Andaluz de Empleo (SAE). Las proporciones establecidas, el 10 por ciento de la población objetivo para los participantes y la relación 2 a 1 entre ambos grupos, se corresponde con las mismas proporciones mantenidas en otros estudios previos tomados como referencia. Las dimensiones de la base de datos son homologables a las evaluaciones similares resumidas en Deheija y Wahba (1999, p. 1056 -Tabla 1-)⁵.

Los datos han sido suministrados para esta investigación por el INEM y por el Servicio Andaluz de Empleo (SAE). Para los individuos de ambos grupos, la base de datos recoge información relacionada con su vida laboral, incluyendo datos referentes a los episodios de desempleo, a las solicitudes de demanda de empleo de los individuos, etc. También incluye información referente a las variables relevantes utilizadas para el desarrollo de la evaluación, entre las que se consideran edad, sexo y zona geográfica de residencia. Otros datos como el sector económico en el que los individuos encuentran empleo, el tipo de contrato o la duración del mismo, no han podido incorporarse a la *BASEVAFOR*, aun cuando su tratamiento enriquecería previsiblemente las conclusiones de la investigación. La información disponible tampoco permite hacer una evaluación de los efectos del programa a medio plazo, por lo que no resulta posible saber en qué medida los efectos de la formación se mantienen a lo largo del tiempo.

⁵ La relación 2/1 nos sitúa en la línea de estudios de evaluación previos como es el del Programa JTPA. Este programa asociado a la "Job Training Partnership Act -JTPA-" estadounidense fue objeto de muy influyentes estudios evaluatorios de entre los que destacan los de Heckman, Ichimura y Todd (1997, 1998) y el de Heckman, Ichimura, Smtih y Todd (1998).

4. Características del modelo

4.1 Definición de las variables D e Y

El desarrollo de la evaluación de políticas públicas se ha beneficiado del uso de la inferencia causal estadística⁶. Uno de los resultados más sobresalientes ha sido el Modelo de Resultados Potenciales -POM-, en el que las variables relevantes son comparadas para el caso de individuos participantes y no participantes en determinadas iniciativas públicas⁷. Un desarrollo prolífico del POM, aplicado a la evaluación de programas públicos de formación, se debe a Roy (1951) y Rubin⁸ (1974, 1978). Este artículo toma como referencia el Modelo Causal de Roy-Rubin⁹ -RRM-.

De acuerdo con el modelo RRM, se define D como la variable binaria que indica la participación de los individuos de la muestra en el programa evaluado, que toma los valores 1 ó 0 dependiendo de si el individuo considerado ha participado o no. De esta forma, $D_i = 1$ indicará que el individuo i es un individuo participante y $D_i = 0$ indicará que i es un individuo no participante (individuo de control).

Por otra parte, el escalar Y es la variable respuesta a partir de la cual se evaluarán los efectos promedios del programa. Se define Y_i como la capacidad del individuo i para encontrar un empleo, y muestra cuánto tiempo necesita emplear hasta encontrarlo¹⁰. La construcción de esta variable se ajusta a la siguiente expresión:

$$Y_i = 1 - \frac{\text{Nº de días transcurridos hasta que el individuo "i" encuentra un empleo}}{\text{Tiempo total del periodo de observación considerado}} \quad (1)$$

La elección de la variable respuesta se justifica en que los individuos de la muestra, tanto los participantes como los de control, están inicialmente desempleados e incluidos en el censo oficial de demandantes de empleo. Muchos de ellos no han trabajado antes o tienen poca experiencia laboral debido a su juventud. Por esa razón, resulta relevante

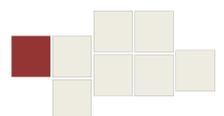
⁶ Una referencia amplia del enfoque teórico de la causalidad y su consideración en los experimentos aleatorios puede encontrarse en Cox (1992). Otros autores como Dawid (1979, 2000), Holland (1986), Heckman (1990) y Pearl (2000) también discuten qué debe entenderse por causalidad en este ámbito. Finalmente, para el caso concreto de los programas de formación, resultan trabajos pioneros el ya citado de Rubin (1974) y el de Heckman y Hotz (1989).

⁷ Cameron y Trivedi (2005, pp. 31 y ss.) exponen las ventajas del Modelo de Resultados Potenciales frente a otros modelos alternativos.

⁸ Como referencias de partida, Rubin tomó los trabajos de Neyman (1923, 1935) y Fisher (1928, 1935).

⁹ Para estimar la causalidad en el MRR se ha de recurrir al supuesto clásico de independencia de los resultados potenciales individuales de la participación de otros individuos, conocida como "Stable Unit Treatment Value Assumption" (SUTVA). En consecuencia, no se tienen en cuenta los efectos de equilibrio general -Cameron y Trivedi (2005, p. 872)-.

¹⁰ Para una investigación posterior en la que se disponga de datos más desagregados está prevista la descomposición de esta variable respuesta Y en dos variables. La primera nos permitiría conocer el efecto promedio del programa sobre la probabilidad individual de encontrar un empleo. La segunda podría definirse como el tiempo necesario para encontrar un empleo condicionado al subconjunto de desempleados (participantes y no participantes) que han encontrado un trabajo.



para la evaluación del programa considerar una variable respuesta que nos permita medir la capacidad de estas personas para encontrar empleo¹¹.

El periodo de observación considerado ha sido de tres años¹² (1095 días). Se empieza a computar este periodo desde el momento en el que los participantes finalizan el programa de formación (generalmente al final de 1999), o bien el 1 de enero de 2000 para los individuos de control¹³.

El valor de Y varía entre 0 y 1. Si Y es igual a 0, significa que el individuo i no ha encontrado empleo durante el periodo considerado. Este es el peor escenario para la eficacia del programa. Un valor de Y próximo a 1 indicará que el individuo i ha encontrado un empleo en un corto periodo de tiempo.

La *Tabla 3* resume los principales datos de estadística descriptiva de la variable Y .

Media	0,578147	Curtosis	-1,392020
Mediana	0,773516	Coefficiente de asimetría	-0,533554
Moda	0	Mínimo	0
Desviación típica	0,380657	Máximo	1

Elaboración: J. M. Cansino y A. Sánchez, 2009.

4.2 El efecto promedio del programa

En los desarrollos del POM y del RRM, los valores individualizados de las variables relevantes pueden extraerse bien de experimentos aleatorios o bien de datos observacionales. Ambos entornos condicionarán notablemente la evaluación y darán lugar a distintos desarrollos metodológicos. Sin embargo, en las ciencias sociales, los experimentos aleatorios se enfrentan a importantes problemas relacionados con su elevado coste, posibles limitaciones de índole moral, el desgaste de la muestra (o efecto abandono) y complicaciones derivadas del conocido como efecto Hawthorne -Burtless (1995) y Cameron y Trivedi (2005)-.

Estas dificultades pueden salvarse mediante el uso de datos observacionales¹⁴. En estos casos, Rosenbaum (1999, p. 266) afirma que el investigador puede diseñar un grupo de tratamiento y un grupo de control a partir de los individuos con el objetivo de reproducir un escenario lo más parecido posible a un experimento aleatorio¹⁵. No

¹¹ En definitiva, el Programa de Escuelas Taller y Casas de Oficios tiene como objetivo "constituirse como una medida de inserción en el mercado de trabajo de jóvenes desempleados menores de 25 años".

¹² Se ha considerado un periodo de tiempo relativamente elevado, tres años (1095 días), por los especiales problemas de este colectivo para encontrar empleo.

¹³ Fecha fijada como punto de partida para el grupo de control ya que coincide con el punto de partida también para gran parte de los individuos del grupo de participantes.

¹⁴ Los estudios pioneros en este ámbito se desarrollaron en Medicina. Como introducción a los planteamientos presentados por los métodos observacionales remitimos a dos de los ejemplos a los que recurre Rosenbaum (1995, pp. 2 y ss.) en su exposición sobre estos métodos: Cochran (1968) y Cameron y Pauling (1976). También resultan de interés los trabajos de Billewicz (1965) y Moertel et al. (1985), relacionados con los dos anteriores.

¹⁵ Conocidos trabajos basados en estos métodos cuasi-experimentales son, entre otros, los de Kiefer (1979) y Bassi (1984). Ambos trabajos son comentados por Lalonde (1986) en su estudio comparativo entre los métodos experimentales y no experimentales.

obstante, los modelos que incluyen sucesos contrafactuales (como que un individuo participe y no participe, al mismo tiempo, en un programa de formación) no permiten estimar efectos causales individuales. Holland (1986) lo denominó como el “problema fundamental de la identificación causal”, que obliga a optar por la estimación de un efecto promedio.

Particularmente, el efecto promedio sobre los participantes en el programa¹⁶, *ATET* (“*Average Treatment Effect on the Treated*”), se obtiene como valor medio esperado de la diferencia entre los resultados potenciales Y_1 e Y_0 , pero sólo con respecto a los individuos que han participado en el programa (esto es, condicionada al valor $D=1$):

$$ATET = E [Y_1 - Y_0 | D = 1] \quad (2)$$

Los distintos métodos de evaluación tratan de construir el suceso contrafactual a través de una redefinición del problema, que implica pasar del nivel individual de la evaluación al nivel poblacional (Heckman et al., 1999).

4.3 La selección sobre variables observables

Dado que la validez del efecto promedio estimado puede verse lesionada si los individuos participantes y de control difieren en características distintas a las derivadas de la propia participación en el programa, habrá que controlar estas características debido a los efectos que pueden inducir sobre los valores de la variable respuesta. En la medida en que estas características ajenas al programa puedan ser observadas y que los individuos (participantes y de control) difieran sólo en ellas¹⁷ se podrán controlar tales diferencias. Sobre esta base se fundamentan los métodos cuasi-experimentales de selección sobre variables observables¹⁸.

La selección sobre variables observables nos permite aislar el efecto de una covariable (o de un vector de covariables) manteniendo la independencia entre la variable D indicativa de la participación y la variable respuesta Y , condición que puede ser expresada como:

$$(Y_1, Y_0) \perp D | X \quad (3)$$

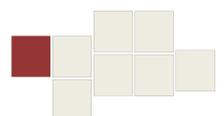
La selección sobre variables observables permite sostener la condición de independencia propia de los experimentos aleatorios, facilitando la comparación entre los individuos participantes y los individuos de control.

Siguiendo a Heckman y Hotz (1989, p. 865), la selección sobre variables observables procede cuando la dependencia entre D e Y es debida a la covariable X , que influye en el

¹⁶ Pese a que en este artículo sólo se utiliza el ATET, en Imbens (2004) se resumen una serie de efectos medios que han sido considerados por la literatura.

¹⁷ Cuando los participantes y no participantes difieren en características personales inobservables, el efecto promedio puede estimarse utilizando el estimador de diferencias en diferencias. Para la investigación desarrollada en este artículo, véase Cansino y Sánchez (2008a).

¹⁸ El marco de la selección sobre variables observables se define en la línea de Barnow, Cain y Goldberger (1980), Heckman y Hotz (1989) y Dehejia y Wahba (1999).



proceso de selección de los individuos, de forma que, controlando X se soluciona el sesgo en la selección, eliminándose la dependencia entre D e Y .

En la selección sobre variables observables, garantizada la condición de independencia, se considera, siguiendo a Dehejia y Wahba (1999, p. 1057), que:

$$E [Y_1 - Y_0 | X] = E [Y | X, D = 1] - E [Y | X, D = 0] \quad (4)$$

De esta forma, es posible calcular el efecto promedio del programa de formación solo para los participantes ($ATET$) como:

$$ATET = E [Y_1 - Y_0 | D = 1] = \int (E[Y | X, D = 1] - E[Y | X, D = 0]) dP(X | D = 1) \quad (5)$$

Así, el $ATET$ será igual a la diferencia de los valores promedios observados en la variable respuesta de los individuos participantes y de control, para cada uno de los posibles valores que registra X cuando $D = 1$.

4.4 Definición del vector de covariables $X^3 = (X_1, X_2, X_3)$

Definida la variable D , X será una variable predeterminada¹⁹ con respecto a D si, para todos y cada uno de los individuos observados, su valor se mantiene constante para cualquier valor de D . Es decir, $X_{0i} = X_{1i}$, siendo X_{0i} el valor de X previo al suceso D ($D=0$) y X_{1i} el valor de X tras producirse D ($D=1$). Esta variable predeterminada X también es denominada covariable o variable contaminante, ya que “contaminará” la variable resultado Y con sus efectos²⁰, que se añadirán a los producidos por D .

El hecho de que X sea predeterminada con respecto a D no implica que esta independencia sea bidireccional, ya que es posible que, como característica de la población considerada, sí pueda presentarse la relación de dependencia en dirección inversa, ocurriendo que el valor de D se vea afectado por el de X .

A partir de la información de la muestra incluida en la base de datos, se han considerado tres variables predeterminadas que conforman el vector de covariables $X^3 = (X_1, X_2, X_3)$. La base de datos sólo nos permite incluir en el modelo información completa sobre estas tres covariables, que quedan definidas de la siguiente forma:

- X_1 : sexo. Muestra si el individuo considerado es hombre o mujer. Al ser una variable cualitativa se incluirá en el modelo mediante el uso de variables ficticias. Para prevenir multicolinealidad perfecta, se incluyen tantas variables instrumentales como categorías menos una, de forma que se incluye la variable ficticia X_{11} , que podrá tomar los valores 0 (si el individuo es mujer) y 1 (si el individuo es hombre).
- X_2 : edad. Recoge la edad del individuo al comienzo del periodo de observación. Teniendo en cuenta que la edad de los participantes es uno de los requisitos para acceder al programa y que el intervalo permitido está entre los 16 y los 24 años, y

¹⁹ Aunque hablamos de una covariable, cuanto se sostiene puede extrapolarse al caso de que X sea un vector de n covariables, definido como $X^n = (X_1, X_2, \dots, X_n)$.

²⁰ Para profundizar en este aspecto, véanse las consideraciones realizadas sobre las variables contaminantes o covariables en Rubin (1978).

considerando que el programa puede extenderse entre 1 y 2 años, la covariable X_2 tomará valores entre los 17 y los 26 años de edad.

- X_3 : zona. Variable cualitativa que muestra la zona de realización del curso o de residencia para el caso de los individuos de control. Para definir esta covariable se ha dividido la provincia de Sevilla en cuatro zonas, siguiendo un criterio de zonificación meramente operativo: zona 1 (Sevilla ciudad), zona 2 (zona este y noreste de la provincia), zona 3 (zona sur y suroeste) y zona 4 (zona oeste y noroeste). La inclusión de la covariable zona en el modelo se hará a partir de la construcción de tres variables ficticias binarias, X_{31} (=1 si el individuo pertenece a la zona 1), X_{32} (=1 si pertenece a la zona 2) y X_{33} (=1 si pertenece a la zona 3).

5. Estimación del efecto promedio: resultados empíricos

5.1 Definición de la probabilidad condicional de participación

Rosenbaum y Rubin²¹ (1983, p. 42) definen la probabilidad condicional de participación (“*propensity score*”) como la probabilidad para un individuo de participar en un programa condicionada sobre los valores de un vector X de variables predeterminadas o covariables. Esta probabilidad será una función de X , normalmente desconocida, y que deberá ser estimada a partir de los datos muestrales.

Siguiendo a Hahn (1998, p. 316), el cálculo de la probabilidad condicional de participación, dadas ciertas características observables, juega un papel crucial en el control del sesgo con objeto de obtener un estimador de los efectos del programa.

Mediante la probabilidad condicional de participación se procede como si de una covariable unidimensional se tratase, ganando la evaluación en operatividad al evitar manejar el extenso número de covariables que puede incluir el vector X .

Denotando dicha probabilidad como $\varepsilon(X)$, puede expresarse como:

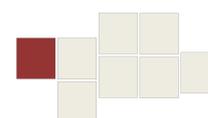
$$\varepsilon(X) = P(D = 1 | X) \quad (6)$$

Para garantizar la condición de independencia propia de los experimentos aleatorios, Rosenbaum y Rubin (1983) formulan la proposición de independencia de la probabilidad condicional de participación como:

$$(Y_1, Y_0) \perp D | \varepsilon(X) \quad (7)$$

A partir de esta proposición es posible plantear que, para todas las observaciones que presenten una misma probabilidad condicional, la distribución del vector X de variables predeterminadas será la misma, de forma que es posible comparar los resultados observados para los individuos participantes y los de control con un mismo valor de la probabilidad condicional. Esto permitirá conocer los efectos del programa

²¹ Véase también el trabajo posterior de Rosenbaum y Rubin (1984).



aislando los efectos contaminantes de todas las covariables consideradas conjuntamente.

5.2 Estimación de la probabilidad condicional de participación

Para determinar la probabilidad condicional de participación²², se expresa la probabilidad de que el individuo participe en el programa condicionada sobre el vector de covariables X como:

$$\varepsilon(X) = P(D = 1 | X) = F(\beta X) \quad (8)$$

donde β es el vector de parámetros asociado a las covariables.

El valor de esta probabilidad quedará condicionado al valor de la función de distribución en el punto βX_j , siendo X_j cada uno de los posibles valores que puede adoptar el vector de covariables X , con $j = 1, \dots, k$. Según la hipótesis sobre la forma de la función de distribución F , se podrán especificar diferentes modelos de elección de respuesta binaria. De entre las posibles opciones, se han seleccionado los dos modelos más comúnmente utilizados: el modelo Probit y el modelo Logit.

No existe un criterio de elección definido para optar por uno u otro modelo, de manera que la elección se realiza por razones meramente operativas, estimando ambos modelos y eligiendo el que presente unos mejores resultados en los términos que más adelante se especificarán. Llevando a cabo las regresiones sobre el vector de covariables $X^3 = (X_1, X_2, X_3)$, los resultados obtenidos aparecen en las *Tablas 4 y 5*.

Tabla 4: Estimación de los parámetros de la probabilidad condicional según el Modelo Probit
Variable Dependiente: D (Probabilidad de que $D = 1$)

Variable	Coficiente	Valor	Error estándar	t -estadístico	Probabilidad
Efecto fijo	μ	-2'366701	0'848762	-2'788414	0'0053
X_{11}	β_{11}	0'580077	0'185036	3'134942	0'0017
X_2	β_2	0'125535	0'040936	3'066652	0'0022
X_{31}	β_{31}	-0'254362	0'285963	-0'889491	0'3737
X_{32}	β_{32}	0'002826	0'254795	0'011091	0'9912
X_{33}	β_{33}	0'058707	0'257340	0'228131	0'8195

Elaboración: J. M. Cansino y A. Sánchez, 2009.

²² El método de la probabilidad condicional de participación queda delimitado según las aportaciones de Hahn (1998), Hirano, Imbens y Ridder (2003) y Abadie e Imbens (2002, 2005 y 2006).

Tabla 5: Estimación de los parámetros de la probabilidad condicional según el Modelo Logit Variable Dependiente: D (Probabilidad de que $D = 1$)					
Variable	Coefficiente	Valor	Error estándar	t -estadístico	Probabilidad
Efecto fijo	μ	-3'801614	1'384561	-2'745719	0'0060
X_{11}	β_{11}	0'959225	0'311150	3'082833	0'0021
X_2	β_2	0'201706	0'067164	3'003174	0'0027
X_{31}	β_{31}	-0'428509	0'477954	-0'896549	0'3700
X_{32}	β_{32}	0'007757	0'422600	0'018354	0'9854
X_{33}	β_{33}	0'090384	0'427989	0'211182	0'8327

Elaboración: J. M. Cansino y A. Sánchez, 2009.

En ambos casos, para corregir los posibles problemas de heterocedasticidad, los valores del t-estadístico aparecen ajustados por el método de White. De acuerdo con los resultados, las variables X_{11} y X_2 aparecen como significativas a la hora de determinar la probabilidad de participación en ambos modelos, mientras que las variables ficticias construidas para incluir la covariable zona, X_{31} , X_{32} y X_{33} , aparecen como no significativas. Sin embargo, se opta por mantenerlas ya que contribuyen a mejorar la significatividad conjunta de todos los parámetros estimados y a obtener una mejor bondad de los ajustes. El modelo más eficiente será aquel que presente menores valores de los criterios de información de Akaike, Schwarz y Hannan-Quinn y un mayor valor de la función de verosimilitud. En la *Tabla 6* se recogen, de forma comparativa, los valores obtenidos de estos criterios y de la función de verosimilitud para cada modelo.

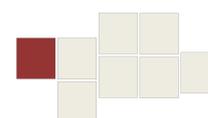
Tabla 6: Comparación de los resultados según los Modelos Probit y Logit		
	Modelo Probit	Modelo Logit
Log. func. verosimilitud	-134'5663	-134'6875
Criterio de infor. Akaike	1'249478	1'250555
Criterio Schwarz	1'340574	1'341651
Criterio Hannan-Quinn	1'286245	1'287322

Elaboración: J. M. Cansino y A. Sánchez, 2009.

El modelo Probit es el que presenta mejores resultados, por lo que es el que resulta seleccionado. La *Tabla 7* recoge los principales datos de estadística descriptiva de la probabilidad condicional de participación estimada mediante el Modelo Probit.

Tabla 7: Estadística descriptiva de la probabilidad condicional obtenida según el Modelo Probit			
Media	0'667158	Curtosis	-0'795769
Mediana	0'681896	Coefficiente de asimetría	-0'145460
Moda	0'682903	Mínimo	0'358889
Desviación típica	0'129550	Máximo	0'920798

Elaboración: J. M. Cansino y A. Sánchez, 2009.



5.3 Los estimadores de emparejamiento

Una vez asignado a cada individuo el valor de su probabilidad condicional de participación en el programa, es posible obtener el estimador del *ATET* mediante la obtención de los estimadores de emparejamiento (o “*matching*”)²³, cuyo objetivo es restablecer las condiciones de un experimento aleatorio.

Con los estimadores de emparejamiento, se busca la similitud entre los individuos participantes y de control asignando a cada individuo del grupo de participantes o del grupo de control, un individuo²⁴ del otro grupo que presente las mismas, o similares, características respecto al vector de covariables X . Como argumentan Dehejia y Wahba (2002, p. 1), este método permite hacer pares de individuos participantes y de control que sean similares en relación a sus características observables. La hipótesis básica es que el sesgo de selección²⁵ se elimina si se condiciona sobre las covariables observables (Heckman et al., 1998a). Cuando todas las diferencias relevantes entre los dos miembros del par quedan recogidas por las covariables observables, este método podrá proporcionar un estimador insesgado de los efectos del programa.

Así, para un individuo i participante (con un valor estimado de su probabilidad condicional de participación $\hat{\varepsilon}(X_i)$), se busca un individuo m de control (con un valor estimado de su probabilidad condicional de participación $\hat{\varepsilon}(X_m)$) tal que $\hat{\varepsilon}(X_i) = \hat{\varepsilon}(X_m)$, o al menos, $\hat{\varepsilon}(X_i) \approx \hat{\varepsilon}(X_m)$. A este individuo m lo denotaremos como $m(i)$.

Para llevar a cabo este proceso de emparejamiento, se construye una función de distancia entre el valor de la probabilidad condicional de participación de un individuo y los valores de los individuos del otro grupo. El individuo del otro grupo que haga mínima esta distancia será el asignado.

Realizado este procedimiento para cada uno de los individuos participantes, podemos obtener el estimador de emparejamiento del *ATET*, que suele ser el más utilizado cuando se aplica el método de emparejamiento, a partir de la siguiente expresión:

$$\hat{\alpha}_{ATET \rightarrow EMPAREJAMIENTO} = \frac{1}{n_1} \sum_{i=1}^{n_1} (Y_i - Y_{m(i)}) \quad (9)$$

²³ El cálculo de los estimadores de emparejamiento sigue las aportaciones de Hahn (1998) e Hirano, Imbens y Ridder (2003). Remitimos también a Rubin (1973), Heckman, Ichimura y Todd (1997, 1998), Smith (1997), Angrist (1998), Dehejia y Wahba (2002), Smith y Todd (2005) y Abadie e Imbens (2002, 2005 y 2006).

²⁴ Existen otros estimadores de emparejamiento más complejos que asignarán a cada individuo no un solo individuo similar, sino un conjunto de ellos, tomando para la comparación el valor medio de sus resultados obtenidos. Una exposición detallada puede seguirse en Cameron y Trivedi (2005, p. 871 y ss.).

²⁵ El sesgo de selección se deriva de las diferencias entre los individuos del grupo de tratamiento y del grupo de control y puede clasificarse en tres tipos: el derivado de comparar individuos diferentes en características observables, el derivado de una diferente distribución de las características observables en ambos grupos y el derivado de características no observables (Mato y Cueto, 2008).

donde Y_i es el valor que toma la variable respuesta Y para el individuo participante i , $Y_{m(i)}$ es el valor de la variable respuesta Y para el individuo de control $m(i)$ que le ha sido asignado como par y n_1 es el número de individuos participantes.

Considerando la información de la *BASEVAFOR*, el estimador del *ATET* ($\hat{\alpha}_{ATET}$) toma un valor positivo, evidenciando un efecto causal favorable del programa sobre los participantes, indicando que éstos ven incrementada su capacidad de inserción laboral, en términos medios, en 0'426581.

5.4 Ponderación sobre la probabilidad condicional de participación

En segundo lugar, es posible obtener también el estimador del *ATET* mediante la ponderación de los valores observados de la variables respuesta Y a partir de los valores obtenidos de la probabilidad condicional de participación para cada individuo, lo que permite la comparación de los resultados con los obtenidos en el apartado anterior. Esta ponderación se hará de acuerdo con la expresión²⁶:

$$\hat{\alpha}_{ATET \rightarrow PONDERACION} = \frac{1}{n_1} \sum_{i=1}^n Y_i \cdot \frac{D_i - \hat{\varepsilon}(X_i)}{1 - \hat{\varepsilon}(X_i)} \quad (10)$$

donde $\hat{\varepsilon}(X_i)$ es el valor estimado de la probabilidad condicional de participación para el individuo i sobre el vector X de variables predeterminadas.

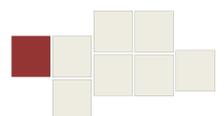
Aplicando esta expresión, el estimador obtenido del *ATET* ($\hat{\alpha}_{ATET}$) toma un valor positivo, evidenciando un efecto causal favorable del programa sobre los participantes, indicando que éstos ven incrementada su capacidad de inserción laboral, en términos medios, en 0'430409.

6. Conclusiones

Los resultados de este trabajo permiten concluir que el Programa de Escuelas Taller y Casas de Oficios evaluado tiene un impacto positivo significativo en la inserción laboral de los individuos participantes. Se han desarrollado dos métodos para la obtención de estimadores del efecto promedio del programa a partir de datos observacionales. La aplicación de estos métodos sobre la muestra a partir de la base de datos construida para esta investigación, nos permite comparar los resultados obtenidos mediante ambos métodos y contextualizarlos con trabajos similares realizados en otros países.

El uso del estimador de emparejamiento permite concluir que el efecto promedio de este Programa se traduce en la reducción en 467 días del periodo de tiempo necesario para encontrar un empleo de los individuos por parte de los individuos participantes. Por su parte, el efecto promedio del Programa estimado mediante la ponderación de las observaciones por la inversa de la probabilidad condicional de participación permite

²⁶ La obtención del estimador basado en la ponderación sobre la probabilidad condicional de participación se realiza de acuerdo con Hirano, Imbens y Ridder (2003).



concluir que el tiempo necesario para encontrar un empleo de los individuos participantes se reduce en 471 días.

Dado que el Programa se diseñó para promover el empleo entre los jóvenes desempleados, este resultado respalda la eficacia del mismo. Como se puede comprobar, los dos estimadores obtenidos indican resultados similares: una significativa reducción del tiempo necesario para encontrar empleo por parte de los individuos participantes.

No obstante, las limitaciones que atañen a la *BASEVAFOR* aconsejan valorar prudentemente estos resultados, que resultan más benévolos que los encontrados por Arellano (2005) y Mato y Cueto (2008) en sus investigaciones. Posteriores investigaciones permitirían robustecer estas conclusiones en el caso de que las Administraciones Públicas implicadas permitiesen ampliar el soporte empírico con datos sobre características individuales adicionales cuyo efecto no ha podido analizarse en este caso, ya que los datos individuales, que permiten la definición de las variables predeterminadas o covariables, proporcionadas hasta el momento son limitados.

Los resultados obtenidos son coherentes con los aportados por otros trabajos, como son los de Card y Sullivan (1988) en EE.UU y de Blundell et al. (2004) en el Reino Unido. En ambos artículos se concluye, a partir de distintos estimadores, la existencia de efectos positivos notorios de los programas de formación evaluados sobre la probabilidad o capacidad de encontrar un empleo de los individuos participantes. En la misma dirección se sitúa el trabajo de Park et al. (1996) para Canadá, que estima un efecto positivo de los programas sometidos a evaluación, aunque en este caso se estima el efecto promedio sobre el nivel de ingresos de los individuos, tomando éste como variable respuesta. Cansino y Román (2007) han expuesto la dificultad de emplear esta variable para la evaluación de programas de formación en España.

7. Referencias utilizadas

ABADIE, A. y G. W. IMBENS (2002): "Simple and Bias-Corrected Matching Estimators for Average Treatment Effects", *NBER Working Paper* nº 283, octubre 2002.

----- (2005): "On The Failure of the Bootstrap for Matching Estimators", *Working Paper*, Octubre 2005.

----- (2006): "Large Sample Properties of Matching Estimators for Average Treatment Effects", *Econometrica*, 74 (1): 235-267.

ANGRIST, J. D. (1998), "Estimating the Labor Market Impact of Voluntary Military Service Using Security Data on Military Applicants", *Econometrica*, 66 (2): 249-288.

ANDREWS, M., S. BRADLEY y R. UPWARD (1999): "Estimating Youth Training Wage Differentials During and After Training", *Oxford Economic Papers*, vol 51(3): 517-544.

ARELLANO, A. (2005): "Do Training Programmes Get the Unemployed Back to Work?: A Look at the Spanish Experience", *Working Paper* 05-25, Economic Series 05, Abril, Universidad Calors III de Madrid.

BARNOW, B., G. CAIN y A. GOLDBERGER (1980): "Selection on Observables", en Stromsdorfer, E. W. and Farkas, G. (eds.), *Evaluation Studies Review Annual*, vol. 5: 43-59, Beverly Hills, California, Ed.: Sage Publications.

BASSI, L. J. (1984): "Estimating the Effect of Training Programs with Non-Random Selection", *Review of Economics and Statistics*, 66 (1): 36-43.

BERGEMANN, A., B. FITZENBERGER y S. SPECKESSER, S. (2005): Evaluating the Dynamic Employment Effects of Training Programs in East Germany Using Conditional Difference-in-

Differences, *Documento de trabajo nº 1848* del “Institute for the Study of Labor (IZA)”, Noviembre, 2005.

BILLEWICZ, W. Z. (1965): “The Efficiency of Matched Samples: an Empirical Investigation”, *Biometrics*, 21 (3): 623-644.

BLUNDELL, R., M. COSTA, C. MEGHIR y J. VAN REENEN (2004): “Evaluating the Employment Impact of a Mandatory Job Search Program”, *Journal of the European Economic Association*, 2 (4): 569-606.

BONNALL, L., D. FOUGÈRE y A. SÉRANDON, A. (1997): “Evaluating the Impact of French Employment Policies on Individual Labour Market Histories”, *Review of Economic Studies*, 64 (4): 683-713.

BURTLESS, G. (1995): “The Case for Randomized Field Trials in Economic and Policy Research,” *Journal of Economic Perspectives*, 9 (2): 63-84.

CAMERON, E. y L. PAULING (1976): “Supplemental Ascorbate in the Supportive Treatment of Cancer: Prolongation of Survival Times in Terminal Human Cancer”, *Proceedings of the National Academy of Sciences of the United States of America*, 73 (10): 3685-3689.

CAMERON, A. C. y P. K. TRIVEDI (2005): *Microeconomic. Methods and Applications*, New York, Ed.: Cambridge University Press.

CANSINO, J.M., M. A. CARDENETE y R. ROMAN: (2007): “Regional evaluation of a tax on retailer sales of some fuels through social accounting matrix”, *Applied Economics Letters*, 14 (12): 877-880.

CANSINO, J.M. y R. ROMÁN (2007): “Evaluación de políticas públicas sobre poblaciones heterogéneas ¿pueden los órganos de control externo contribuir a su avance?”. *Revista Española de Control Externo*, Vol. IX, nº 25, pp. 107-129.

CANSINO, J.M. y A. SÁNCHEZ (2008a): “Evaluación del programa de Escuelas Taller y Casas de Oficios a partir de su efecto sobre el tiempo de búsqueda del primer empleo. El caso de Sevilla”, *Estudios de Economía Aplicada*, próxima aparición. Versión disponible en www.revista-eea.net/aceptados.php

----- (2008b): “Average effect of training programs on the time needed to find a job. The case of the training programs schools in the south of Spain (Seville, 1997-1999)”. Documentos de trabajo de Funcas. Versión disponible en www.funcas.es.

CARD, D. y D. SULLIVAN (1988): “Measuring the Effects of Subsidized Training Programs on Movements In and Out of Employment”, *Econometrica*, 56 (3): 497-530.

COCHRAN, W. G. (1968): “The Effectiveness of Adjustment by Subclassification in Removing Bias in Observational Studies”, *Biometrics*, 24 (2): 295-313.

CONSEJO ECONÓMICO Y SOCIAL (1995): *España... economía, trabajo y sociedad: memoria sobre la situación socioeconómica y laboral*, Madrid, Ed.: Consejo Económico y Social.

COX, D. R., (1992): “Causality: Some Statistical Aspects,” *Journal of the Royal Statistical Society, Series A (Statistics in Society)*, 155 (2): 291-301.

DAWID, A. P. (1979): “Conditional Independence in Statistical Theory”, *Journal of the Royal Statistical Society, Series B (Statistics Methodological)*, 41 (1): 1-31.

----- (2000): “Causal Inference Without Counterfactuals”, *Journal of the American Statistical Association*, 95 (2): 407-448.

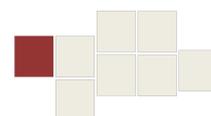
DEHEJIA, R. H. y S. WAHBA (1999): “Causal Effects in Non-Experimental Studies: Reevaluating the Evaluation of Training Programs”, *Journal of the American Statistical Association*, 94 (448): 1053-1062.

----- (2002): “Propensity Score Matching Methods for Non-experimental Causal Studies”, *Review of Economics and Statistics*, 84 (1): 151-161.

FISHER, R. A. (1928): *Statistical Methods for Research Workers*, 2ª edición, London, Ed.: Oliver and Boyd.

----- (1935): *The design of experiments*, Edinburgh, Ed.: Oliver and Boyd.

HAHN, J. (1998): “On the Role of the Propensity Score in Efficient Semiparametric Estimation of Average Treatment Effects”, *Econometrica*, 66 (2): 315-331.



HECKMAN, J., N. CLEMENTS y J. SMITH (1997): "Making the Most out of Programme Evaluations and Social Experiments: Accounting for Heterogeneity in Programme Impacts" *Review of Economic Studies*, 64 (4): 487-535.

HECKMAN, J. J. y V. J. HOTZ (1989): "Choosing Among Alternative Nonexperimental Methods for Estimating the Impact of Social Programs: The Case of Manpower Training", *Journal of the American Statistical Association*, 84 (408): 862-874.

HECKMAN, J. J., H. ICHIMURA, J. SMITH y P. E. TODD (1998a): "Characterizing Selection Bias Using Experimental Data", *Econometrica*, 66 (5): 1017-1098.

----- (1998b): "Matching As an Econometric Evaluation Estimator", *Review of Economics Studies*, 65 (2): 261-294.

HECKMAN, J. J., H. ICHIMURA y P. E. TODD (1997): "Matching As an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme", *Review of Economics Studies*, 64 (4): 605-654.

HECKMAN, J. J., R. J. LALONDE y J. A. SMITH (1999): "The Economics and Econometrics of Active Labour Market Programms", en Ashenfelter, O. y Card, D. (eds.): *Handbook of Labor Economics*, vol. III, Elsevier Science B.V.

HIRANO, K., G. IMBENS y G. RIDDER, (2003): "Efficient Estimation of Average Treatment Effects Using the Estimated Propensity Score", *Econometrica*, 71(4): 1161-1189.

HOLLAND, P. W. (1986): "Statistics and Causal Inference" (con comentarios), *Journal of the American Statistical Association*, 81 (396): 945-970.

IMBENS, G. W. (2004): "Nonparametric Estimation of Average Treatment Effects Under Exogeneity: A Review", *Review of Economics and Statistics*, 86 (1): 4-29.

KIEFER, N. (1979): *Economic Benefits from Four Manpower Training Programs*, Garland Series of Outstanding Dissertations in Economics, New York, Ed.: Garland Press.

LALONDE, R. J. (1986): "Evaluating the Econometric Evaluation of Training Programs with Experimental Data", *American Economic Review*, 76 (4): 604-620.

MANSKI, C. F. (2001): "Designing programs for heterogeneous populations: The value of covariate information", *American Economic Review*, 91 (2): 103-106.

MANSKI, C. e I. GARFINKEL (1992): *Evaluating Welfare and Training Programs*, Cambridge, Massachusetts, Ed.: Harvard University Press.

MATO, F. J. (2002): *La formación para el empleo: una evaluación cuasi-experimental*. Madrid, Consejería de Trabajo y Promoción de Empleo del Principado de Asturias y Civitas Eds.

MATO, F. J y B. CUETO. (2008): "Efectos de las Políticas de Formación a Desempleados", *Revista de Economía Aplicada*, 46 (16): 61-83.

MOERTEL, C., T. FLEMING, E. CREAGAN, J. RUBIN, M. O'CONNELL y M. AMES (1985): "High-dose Vitamin C Versus Placebo in the Treatment of Patients with Advanced Cancer Who Have Had no Prior Chemotherapy: a Randomized Double-blind Comparison", *New England Journal of Medicine*, 312: 137-141.

NEYMAN, J. (1923): "On the Application of Probability Theory to Agricultural Experiments. Essay on Principles.", re-edited en *Statistical Science*, 1990, 5 (4): 465-472.

----- (1935): "Statistical Problems in Agricultural Experimentation", *Supplement to the Journal of the Royal Statistical Society*, 2: 107-180.

PARK, N., B. POWER, W. C. RIDDELL y G. WONG (1996): "An assessment of the Impact of Government-Sponsored Training", *Canadian Journal of Economics*, 29 (Special Issue: Part I): S93-S98.

PEARL, J. (2000): *Causality: Models, Reasoning and Inference*, Cambridge, Ed.: Cambridge University Press.

ROSENBAUM, P. R. (1995): *Observational Studies*, Springer Series in Statistics, Nueva York, Ed.: Springer-Verlag.

----- (1996): "Observational Studies and Nonrandomized Experiments", en S. Ghosh and Rao, C. R. (eds.), *Handbook of Statistics*, vol. 13, capítulo 6: 1277-1366, New York, Ed.: Elsevier.

----- (1999): "Choice As an Alternative to Control in Observational Studies" (con comentarios), *Statistical Science*, 14 (3): 259-304.

ROSENBAUM, P. R. y D. B. RUBIN (1983): “The Central Role of the Propensity Score in Observational Studies for Causal Effects”, *Biometrika*, 70 (1): 41-55.

----- (1984): “Reducing Bias in Observational Studies Using Subclassification on the Propensity Score”, *Journal of the American Statistical Association*, 79 (387): 516-524.

ROY, A. (1951): “Some Thoughts on the Distribution of Earnings”, *Oxford Economic Papers*, 3 (2): 135-146.

RUBIN, D. B. (1973): “Matching to remove bias in observational studies”, *Biometrics*, 29 (1): 159-183.

----- (1974): “Estimating Causal Effects of Treatments in Randomized and Non-randomized Studies”, *Journal of Educational Psychology*, 66 (5): 688-701.

----- (1978): “Bayesian Inference for Causal Effects the Role of Randomization,” *Annals of Statistics*, 6 (1): 34-58.

SMITH, H. L. (1997): “Matching with multiple controls to estimate treatment effects in observational studies”. *Sociological Methodology*, 27: 325-353.

SMITH, J. y P. TODD (2005): “Does Matching Overcome Lalonde’s Critique of Nonexperimental Estimators”, *Journal of Econometrics*, 125 (1-2): 305-353.

