

29

Las becas y ayudas al estudio como elemento determinante de la continuidad escolar en el nivel secundario post-obligatorio.

Un anàlisis de sensibilidad a partir de la aplicación del Propensity Score Matching

Mauro Mediavilla Bordalejo

Universitat de Barcelona

GIPE – IEB

mmediavilla@ub.edu

Las becas y ayudas al estudio como elemento determinante de la continuidad escolar en el nivel secundario post-obligatorio. Un análisis de sensibilidad a partir de la aplicación del Propensity Score Matching

Mauro Mediavilla Bordalejo

Universitat de Barcelona

GIPE – IEB

mmediavilla@ub.edu

El objetivo de la comunicación es comprobar la importancia de las becas y ayudas al estudio como un elemento determinante de la continuidad escolar de los alumnos del nivel secundario post-obligatorio en España. El diseño metodológico aplicado se basa en un análisis cuasi-experimental aplicando el Propensity Score Matching (PSM), a partir de la Encuesta de Condiciones de Vida de EUROSTAT (ECV), publicada en 2009, que brinda valiosa información acerca del individuo y su situación personal y familiar para el período 2004-2006. La aproximación empírica tiene dos partes claramente definidas. En primer lugar, la aplicación del PSM que permite confirmar la relevancia de las becas y ayudas al estudio como un elemento favorecedor de la continuidad escolar, para el nivel educativo analizado. En segundo lugar, la realización de un estudio para confirmar la robustez de las estimaciones, mediante la comprobación de la calidad de los emparejamientos realizados y la validez de los supuestos establecidos. Para intentar asegurar el cumplimiento de ambos requisitos se analiza la propia calidad de los emparejamientos, la correcta estimación de la varianza y la sensibilidad de los resultados obtenidos.

1 Introducción

La educación de las personas es un elemento indispensable para el desarrollo equilibrado y sostenible a largo plazo de las economías y de los sistemas políticos. Además, en los últimos tiempos dicho valor se ha incrementado debido a la creciente globalización de los mercados y el consecuente aumento en la competencia por los puestos de trabajo.

La literatura indica que la educación y el aprendizaje del individuo dependen de una multitud de factores personales y ambientales. Ello implica que las desigualdades existentes en la sociedad podrían llegar a ser reproducidas por el sistema educativo, generando una desigualdad manifiesta entre los ciudadanos en cuanto a las posibilidades reales de educarse y una clara pérdida de eficiencia en la medida que personas capacitadas se vean imposibilitadas de acceder niveles educativos superiores.

Esta situación fundamenta la intervención pública, que se aplica persiguiendo dos objetivos básicos: la igualdad de oportunidades y la eficiencia. En este sentido, las becas constituyen una herramienta que permite dirigir la acción pública hacia determinadas personas con el objeto de intentar disminuir los desequilibrios de partida.

En el ámbito español las becas y ayudas al estudio (becas, en adelante) intentan cubrir diversos costes asociados a la escolarización, siendo los mismos: los costes directos (matrícula), indirectos (material escolar, transporte, residencia) y los costes de oportunidad (salarios dejados de percibir al estar estudiando). En la actualidad su principal objetivo pasa por garantizar la igualdad de oportunidades de acceso a los diferentes niveles educativos.

En relación al nivel secundario post-obligatorio, dos han sido los principales elementos que han centrado el debate. En primer lugar, la gran cantidad de individuos que no continúan sus estudios más allá de la escolarización obligatoria y que se cuantifica mediante una alta tasa de abandono escolar prematuro¹ (tasa del 31,9 % en 2008, muy por encima del 10 % que plantea la Estrategia de Lisboa para 2010). En segundo lugar, la necesidad de impulsar estudios y análisis que evalúen la efectividad real de las políticas públicas y, entre ellas, las becas.

Son estos dos elementos anteriormente mencionados los que sirven como motivación para el presente trabajo, que pretende comprobar si las becas mejoran el nivel educativo y, en caso afirmativo, evaluar cual es su efecto “neto” en las posibilidades educativas de los individuos en el nivel secundario post-obligatorio. Finalmente, la literatura empírica, en los últimos años, sugiere la necesidad de realizar una comprobación múltiple de la robustez de los parámetros estimados centrada en la calidad de los emparejamientos, la correcta estimación de la varianza y la sensibilidad de los resultados.

El presente trabajo se inicia con una breve enunciación de los diferentes factores que afectan al rendimiento educativo (capítulo 2) y la metodología aplicada para la evaluación del impacto de las becas en el rendimiento educativo (capítulo 3). En el capítulo 4 se exponen una descripción de la base de datos y la metodología de imputación aplicada y; en el capítulo 5, se realiza el análisis de los principales resultados. En el capítulo 6, se plantean las diferentes aproximaciones empíricas destinadas a comprobar la robustez de las estimaciones y; por último, en el capítulo 7 se plantean las principales conclusiones.

2 Factores que afectan el rendimiento educativo

Existen en la literatura una serie de desarrollos teóricos y aplicaciones empíricas que se han elaborado con el objetivo de explicar las diferencias en los rendimientos académicos, mediante el

¹ Abandono escolar prematuro: porcentaje de jóvenes de 18 a 24 años que han completado como máximo la primera etapa de la educación secundaria y no continúan ningún estudio o formación adicional.

análisis de una serie de factores que previsiblemente impactarían en las posibilidades educativas reales de los alumnos.

En el caso del individuo se plantean diferentes elementos que influirían en los resultados educativos, como sus propias motivaciones, la calidad y el tiempo dedicado al estudio; su estado de salud; si es hijo biológico o adoptado; su condición de inmigrante; y el denominado “efecto calendario” (diferencia de aprendizaje en los individuos con menor edad derivada de una falta de madurez); el género y; por último, la escolaridad anterior (Chevalier, 2004). Otro elemento que se ha citado en las aproximaciones empíricas es la relación positiva entre la asistencia del alumnos al centro educativo y su rendimiento académico (Bedi y Marshall, 2002).

Dentro del entorno más cercano al individuo sobresalen el nivel educativo, la categoría ocupacional y los ingresos de los progenitores, ya que en conjunto conforman la estructura financiera y cultural con las que cuenta el individuo. La influencia de dichos factores no se limita a la ayuda que significa en el momento que la persona se encuentra inserta en el sistema educativo formal, sino que también influye en su futura introducción en el mercado de trabajo. Otros elementos referidos a los padres, también citados en la literatura, son la etnia, la edad y la condición de inmigrante de los mismos (Gang y Zimmerman, 2000; Chevalier, 2004).

Con respecto al hogar destacan diversos factores como la presencia de ambos padres en el hogar y el número de hermanos, elementos que afectarían el grado de seguimiento y control del proceso educativo del hijo (Baydar y Brooks-Gunn, 1991). Otros factores, algunos ya mencionados anteriormente, son su situación financiera y su posible fluctuación en el tiempo causado, en algunos casos, por el desempleo del cabeza de familia (Comi, 2003); la importancia de las redes sociales del hogar (Sullivan, 2001) y la persistencia de circunstancias adversas como la separación o divorcio (Price, 2008).

En referencia al impacto generado por la titularidad del centro escolar (que para el caso español, puede ser público, privado o privado con financiación pública -concertado-), Calero y Escardíbul (2007) estudian la eficacia comparada de los diferentes tipos de centro y, si bien los resultados generales muestran un mejor rendimiento de las escuelas privadas, tal efecto queda neutralizado al controlarlo por las variables del entorno.

En referencia al efecto generado por el compañero (*peer effects*), su existencia se fundamenta en que el individuo no aprende sólo del profesor en el aula, sino que también se nutre de conocimientos, competencias y diferentes estímulos originados en la interrelación constante con sus compañeros. Dentro de éstos últimos la literatura menciona la transmisión de actitudes, valores, modelos de comportamiento y de motivación (elementos que, en muchos casos, reflejan indirectamente las actitudes y valores de sus respectivos padres (Summers y Wolfe, 1977; Santín, 2006).

En cuanto a la calidad de los profesores (nivel educativo, conocimientos en general y experiencia) y su motivación, Rivkin et al. (2005) contrastan, con sus resultados, la relación positiva entre la capacidad del docente y los buenos resultados obtenidos por los alumnos. Otros elementos que influyen en el desarrollo educativo de los alumnos son la organización de la jornada escolar (duración que cada clase y cantidad de clases por día), el tiempo efectivo de enseñanza (por día y cantidad de días de clase por año lectivo) y la estructura jerárquica existente en el establecimiento educativo (Eren y Millimet, 2008).

Dentro del entorno más general del individuo se plantean las posibles diferencias entre habitar en una zona urbana o rural (Vandenberghe, 1999), así como el grado de incidencia de la pobreza o la delincuencia en el lugar de residencia, o la situación del mercado laboral durante los años de escolarización, en especial cuando el individuo debe decidir su continuidad en el sistema educativo no obligatorio (Chevalier, 2004).

En este marco se introduce la potencial acción del Estado a través de diferentes políticas públicas como, por ejemplo: las ayudas directas a los niños (en forma de becas u otras ayudas), transferencias a las familias, beneficios fiscales y ayudas relacionadas con la maternidad o dirigidas a las familias con diferente estructura a la considerada tradicional (Garib et al. 2007).

3 Metodología aplicada en la estimación del impacto de las becas

La evaluación de políticas públicas es un área de investigación aplicada cuyo objetivo fundamental es valorar la utilidad y el impacto de la intervención pública, teniendo en cuenta que en un entorno de escasez de recursos resulta fundamental evaluar, de forma rigurosa, la actuación que lleva a cabo el sector público, tanto desde la vertiente de la eficiencia en la asignación de los recursos de los que dispone, como de la efectividad de las políticas que desarrolla.

La primera pregunta que se busca contestar con la estimación cuantitativa del impacto de una intervención es cuál habría sido la situación de los beneficiarios si no hubieran participado en el programa. Se trata de una pregunta aplicable a cualquier tipo de intervención y su estudio permite analizar los efectos esperados y no esperados que la intervención genera en los beneficiarios. Esta idea se puede representar mediante la siguiente ecuación:

$$\alpha_i = Y_{1i} - Y_{0i},$$

donde Y_{1i} son las condiciones de vida que el individuo alcanzó cuando participó en el programa; Y_{0i} representa las condiciones de vida que el individuo hubiera alcanzado en ausencia del programa; y α_i el impacto del programa en la persona “i”. El impacto nunca puede ser observado directamente, siendo conocida esta cuestión como el “Problema Fundamental de Evaluación”, dado que sólo una de las dos situaciones potenciales (participar o no participar) es observada para cada individuo en un momento dado (Rubin, 1974).

Así, la evaluación de impacto contará sólo con observaciones de Y_{1i} para los individuos beneficiarios del programa, y el problema que se debe solucionar es la estimación de Y_{0i} , el cual es llamado comúnmente escenario contrafactual.

De acuerdo con Cook y Campbell (1979) el impacto se debe estimar mediante una comparación sistemática entre grupos de población que reciben los beneficios de la intervención (grupos de tratamiento) y grupos de no beneficiarios (grupos de control) con características muy similares en los aspectos más relevantes a los tratados antes de recibir el tratamiento. En este caso, el grupo de control simularía el escenario contrafactual Y_{0i} .

Los diseños de evaluación de impacto se clasifican en experimentales y cuasi-experimentales, siendo la principal diferencia la forma en que se asignan los participantes a los grupos de tratamiento y control (aleatoria en el diseño experimental y no aleatoria en el cuasi-experimental)².

En el caso del diseño cuasi-experimental y con el objeto de conocer y aislar el efecto neto de una política, se comparan dos grupos: el grupo experimental al cual se le ha aplicado el programa y el grupo de control, que estando en iguales condiciones para recibirlo, no lo recibe. Los efectos netos son el resultado de la diferencia de comportamiento entre ambos grupos sobre la variable objetivo considerada.

Se pueden utilizar distintas metodologías de comparación (*matching*) para construir el grupo de control. Se ha escogido el método de *Propensity Score Matching* (PSM, en adelante) puesto que permite escoger aquellos individuos que tienen la misma probabilidad o propensión para solicitar la ayuda a partir de una serie de características propias y del entorno de la persona, pero que también presentan un comportamiento similar de los factores no observables a priori (Caliendo y Kopeining, 2008). Mediante este método de emparejamiento se calcula el efecto medio del tratamiento en los tratados (*Average Treatment effect on the Treated* o ATT), que determina el valor medio del tratamiento para las personas que recibieron el tratamiento en comparación con los no tratados en el caso hipotético de que ellos también hubieran recibido el tratamiento.

4 Selección muestral y metodología de imputación aplicada

4.1 Selección muestral

Para el análisis empírico se emplean los datos correspondientes a la *Encuesta de Condiciones de Vida (ECV)*³, elaborada por EUROSTAT con datos longitudinales para el período 2004-2006,

² Para una revisión de las diferentes técnicas aplicadas, véase Heckman et. al. (1999); Blundell y Costa Dias (2000, 2002).

³ Contrato n. EU-SILC/2006/19. En él se establece la obligación, en el momento de publicar los resultados, de comunicar lo siguiente: “EUROSTAT no es responsable de los resultados y las conclusiones, responsabilidad que corresponde al investigador”.

publicada en 2009. Los datos disponibles hacen referencia a los países de la Unión Europea y en el caso español, la muestra agrupada comprende 58.740 individuos.

Para el estudio de impacto de las becas y ayudas al estudio en el rendimiento educativo de los estudiantes, la variable dependiente hace referencia al nivel educativo que posee la persona a los 19 años. Si bien, la edad teórica para finalizar el nivel secundario post-obligatorio es a los 18 años, se ha optado por seleccionar un año más para evitar encontrar individuos con 18 años que aún no tengan este nivel educativo alcanzado sólo porque la encuesta se ha realizado antes de finalizar su curso lectivo. Como variables independientes, inicialmente, se consideran todas las variables, con información disponible, relacionadas al individuo, sus progenitores, su hogar y el entorno que lo rodea.

Variables individuales

- 1) Nivel educativo (19 años). Esta variable hace referencia al nivel educativo que posee el individuo a la edad de 19 años en 2006. Se estructura como una *dummy* con valor = 1 si la persona tiene un nivel educativo igual o superior al de secundaria post-obligatoria (ISCED-97) (Post_Oblig_con_19).
- 2) Beca. Variable *dummy* que indica si el individuo ha recibido una o más becas o ayudas al estudio durante el período 2004-2006 (Beca).
- 3) Género. *Dummy* género. Toma el valor 1 si el individuo es mujer (Mujer).
- 4) “Efecto calendario”. *Dummy* mes de nacimiento. Toma el valor 1 si el individuo nació en el último trimestre del año (Último_Tri).
- 5) Estado de salud . *Dummy* enfermedad crónica. Toma el valor 1 si el individuo padece una enfermedad o incapacidad crónica (Enf_Crónica).
- 6) Orden entre hermanos. Variable que hace referencia al orden que ocupa el individuo en relación con sus hermanos (Ejemplo: el hermano mayor tiene un número de orden igual a 1) (Posición).

Variables de los progenitores

- 1) Nivel educativo Padre. Máxima educación lograda por el padre (ISCED-97) (Educ_Padre).
- 2) Nivel educativo Madre. Máxima educación lograda por la madre (ISCED-97) (Educ_Madre).
- 3) Ocupación Padre. Variable que indica el tipo de ocupación del padre que se elabora a partir del ISCO-88⁴ (Ocup_Padre).
- 4) Ocupación Madre. Variable que indica el tipo de ocupación de la madre que se elabora a partir del ISCO-88 (Ocup_Madre).
- 5) Actividad Padre. *Dummy* activo. Toma el valor 1 si el padre se encuentra activo (ocupado o desempleado) (Activo_Padre).

⁴ ISCO-88: *Internacional Standard Classification of Occupations*.

- 6) Actividad Madre. *Dummy* activo. Toma el valor 1 si la madre se encuentra activa (ocupada o desempleada) (Activo_Madre).

Variables del hogar

- 1) Tamaño del hogar. Variable que hace referencia al número total de miembros del hogar (Tamaño_Hogar).
- 2) Número de hermanos. Variable que indica la cantidad de hermanos existentes en el hogar (Nro_Hermanos).
- 3) Nivel de ingresos (I). Quintil de ingresos disponibles equivalentes⁵ (Quintil).
- 4) Nivel de ingresos (II). *Dummy* dificultades económicas. Toma el valor 1 si el hogar declara tener problemas para asumir los gastos habituales del mes (Dificultad_Econ).
- 5) Régimen de la vivienda. *Dummy* propietario de la vivienda. Toma el valor 1 si los habitantes del hogar son propietarios de la misma (Vivienda_Prop).
- 6) Problemas estructurales. *Dummy* problemas estructurales. Toma el valor 1 si existen problemas estructurales en la vivienda (Prob_Estructural).
- 7) Dimensiones del hogar. *Dummy* más de cuatro ambientes en el hogar. Toma el valor 1 si el hogar tiene más de cuatro ambientes (Mas_4_Amb).
- 8) Capacidad de gasto del hogar. *Dummy* vacaciones. Toma el valor 1 si el hogar puede asumir los costes de una salida anual en concepto de vacaciones (Vacaciones).

Variables del entorno

- 1) Grado de urbanización. *Dummy* si el individuo vive en una zona de baja o media urbanización. (Baja_Media_Urb).

La base inicial se reduce al introducir los diferentes criterios empleados en las diferentes estimaciones realizadas. A partir del total de observaciones válidas para la variable dependiente (Post_Oblig_con_19), se genera una sub-base de datos con 783 observaciones que contiene valores perdidos para algunas de las variables independientes, básicamente localizadas en aquellas que hacen referencia a los progenitores. En el apartado siguiente se explicita la técnica de imputación aplicada, con el objeto de realizar las futuras estimaciones con el mayor grado de ajuste posible y aprovechando toda la información disponible en la base de datos.

4.2 Metodología de imputación de los valores no observados

Habitualmente, la metodología seguida por la literatura empírica indica que las estimaciones se realizan sólo a partir de las observaciones con información válida para todas las variables intervinientes, por lo que se eliminan todas aquellas que presenten falta de información en alguna de las variables empleadas en la estimación. A consecuencia de la pérdida de información, se

⁵ El ingreso equivalente se calcula teniendo en cuenta el ingreso disponible anual del hogar, el factor de no respuesta y el tamaño equivalente del hogar, el cual pondera de manera diferencial a los adultos y a los menores del hogar.

genera un incremento de la varianza y una mayor probabilidad de realizar estimaciones a partir de una muestra poco representativa de la población analizada.

En los últimos años, la literatura empírica se ha decantado por sugerir que la manera más adecuada es la sustitución de los valores faltantes mediante la generación de valores plausibles, siguiendo alguna de las técnicas de imputación desarrolladas⁶. En este trabajo se aplica la técnica de imputación múltiple, cuyo desarrollo y aplicación se ha ido extendiendo con el paso de los años, favorecida, principalmente, por dos factores. En primer lugar, a causa del desarrollo de los diferentes programas econométricos que han permitido su generalización entre la comunidad académica (Van Buuren et. al., 1999; Royston, 2004, 2005; Reiter y Raghunathan, 2007) y; en segundo lugar, a partir de la publicación de diferentes estudios que han demostrado las ventajas de la imputación múltiple frente a los procedimientos tradicionales (Acock, 2005; Ambler y Omar, 2007).

Esta técnica, de características estocásticas, permite hacer un uso eficiente de los datos, obtener estimadores no sesgados y reflejar la incertidumbre que la no-respuesta parcial introduce en la estimación de los parámetros (Rubin, 1996). Su aplicación se basa en sustituir los datos faltantes por $m > 1$ valores plausible simulados⁷. La aplicabilidad de este método se ha visto potenciada con la incorporación en su esquema general de los métodos de Monte Carlo basados en cadenas de Markov, conocidos como algoritmos MICE (*Multiple imputation by chained equations*). Asimismo, a la imputación múltiple se la considera una metodología flexible que permite trabajar con datos multivariados y con distribuciones monótonas o arbitrarias de los valores perdidos. Su aplicabilidad requiere que el patrón de distribución de los valores perdidos sea aleatoria (MCAR -*Missing Completely At Random*- o MAR -*Missing At Random*).

El proceso de imputación múltiple consta de tres etapas. En la primera, cada valor perdido se reemplaza por un conjunto de $m > 1$ valores generados por simulación, con los que se crean m matrices de datos completas. Para generar estos valores plausibles el investigador debe decidir que método de estimación emplea para cada variable a imputar, teniendo en cuenta sus características propias. En la segunda etapa, el investigador debe aplicar a cada matriz simulada el análisis deseado que se hubiese aplicado a la base original en caso de no haber contenido observaciones faltantes. Por último, se combinan los resultados obtenidos en cada matriz para obtener una estimación global, mediante el cálculo de la media, para conocer el parámetro estimado.

Mediante el análisis de la base de datos se llega a la conclusión que existe una distribución arbitraria de los datos faltantes, en cuyo caso y siguiendo la literatura, se aplica una estimación múltiple a partir del algoritmo MICE, sistematizado para el programa STATA mediante el comando *ice*. La imputación aquí aplicada genera valores plausibles a partir de una serie de

⁶ Para una revisión de la literatura, véase Acock (2005).

⁷ Para una explicación teórica detallada, véase Schafer (1999).

modelos univariantes en los cuales una variable única es imputada en base de un grupo de variables.

En la tabla 1 se expone un análisis descriptivo de las variables existentes en la base de datos y, en los casos que la variable posea valores faltantes, con sus valores originales e imputados. Si bien no se puede exigir que el análisis descriptivo para las variables con datos faltantes sea el mismo para la base imputada que para la base original, sí que se espera encontrar valores medios no muy disímiles.

Los resultados indican que no se generan diferencias destacables en la media al incorporar los valores imputados y además, los desvíos estándar continúan en el mismo nivel. Por último, cabe destacar que el proceso de imputación no disminuye el rango de variabilidad existente en las observaciones originales, elemento que la literatura valora como positivo de esta técnica. Finalmente, queda establecida una base de datos completa, con datos para 783 individuos.

5 Análisis de los resultados

Para intentar conocer el efecto “neto” de las becas y ayudas al estudio en el rendimiento educativo de los alumnos en el nivel secundario post-obligatorio se aplica el *Propensity Score Matching* (PSM) que, según Rosenbaum y Rubin (1983), tiene la capacidad de eliminar los diferentes sesgos asociados a las diferencias existentes antes de aplicar el tratamiento entre el grupo tratado y de control.

5.1 Evaluación del efecto “neto” de las becas

El método de emparejamiento para estimar los efectos causales bajo selección de observables mediante el PSM se constituye en dos etapas. En una primera etapa se estima el *Propensity Score* que indica la probabilidad de participar en el programa, mediante un modelo *logit* u otro modelo de elección binaria. Finalmente, se realiza el análisis de emparejamiento que permitirá conocer el efecto “neto” de la política.

5.1.1 Primer paso: estimación de la probabilidad de obtener una beca

La evaluación del impacto de las becas se hará mediante una técnica de emparejamiento cuyo proceso se inicia con un análisis de máxima verosimilitud (*logit*) para estimar las probabilidades de cada individuo de obtener una beca en el nivel secundario post-obligatorio, a partir de todas las variables individuales y del entorno antes descritas (resultados no citados).

Tabla 1. Análisis descriptivo. Valores con y sin imputar.

	Media		Desvío Estandar		Mínimo	Máximo
	Original	Imputada (*)	Original	Imputada (*)		
Post_oblig_con_19	0,45	-	0,49	-	0	1
Beca	0,19	-	0,39	-	0	1
Mujer	0,48	-	0,49	-	0	1
Último_Tri	0,26	0,26	0,44	0,44	0	1
Enf_Crónica	0,07	-	0,26	-	0	1
Posición	1,57	1,57	0,69	0,70	1	3
Educ_Padre	2,45	2,47	1,49	1,51	1	5
Educ_Madre	2,29	2,28	1,37	1,38	1	5
Ocup_Padre	3,74	3,69	1,02	1,08	1	5
Ocup_Madre	3,50	3,28	1,45	1,54	1	5
Activo_Padre	0,92	0,89	0,27	0,32	0	1
Activo_Madre	0,51	0,52	0,50	0,50	0	1
Tamaño_Hogar	4,17	-	0,99	-	1	6
Nro_Hermanos	1,33	1,32	0,87	0,88	0	3
Quintil	2,63	-	1,37	-	1	5
Dificultad_Econ	0,63	-	0,48	-	0	1
Vivienda_Prop	0,83	-	0,37	-	0	1
Prob_Estructural	0,17	-	0,38	-	0	1
Mas_4_Amb	0,71	-	0,45	-	0	1
Vacaciones	0,54	-	0,49	-	0	1
Baja_Media_Urb	0,54	-	0,49	-	0	1

(*) El valor imputado surge como el promedio de los valores obtenidos en las 20 bases de datos completas generadas por el proceso de imputación.

Fuente: Elaboración propia a partir de microdatos de EUSILC LONGITUDINAL UDB 2006 – version.2 – de Marzo 2009.

5.1.2 Segundo paso: análisis de emparejamiento

La estimación de la probabilidad de cada persona de obtener una beca (PS) no es suficiente para conocer el real impacto de la misma. Para ello es necesario conocer el efecto diferencial de haber recibido una beca en personas con igual probabilidad de obtenerla (ATT) y, por lo tanto, similares al individuo tratado en el resto de aspectos conocidos. Debido a la dificultad de encontrar dos personas con idéntica probabilidad se han desarrollado diversos métodos de emparejamiento que intentan estimar el ATT, siendo cuatro más conocidos los que se aplicarán en la presente evaluación: el método del Vecino más Próximo, de Estratificación, Kernel y Radius (Becker e Ichino, 2002).

Los resultados obtenidos (véase tabla 2) muestran la efectividad de las becas y ayudas al estudio otorgadas debido a su efecto positivo sobre el nivel educativo logrado por los individuos de 19 años. Las diferentes técnicas de emparejamiento calculadas indican la existencia de un efecto medio del tratamiento que es positivo y significativamente diferente de cero, con lo que se puede afirmar que existe un efecto “neto” positivo de las becas en las personas beneficiarias. Asimismo, es preciso mencionar tres elementos centrales que contribuyen a confirmar la validez de las estimaciones. En primer lugar, que el parámetro estimado no presenta variaciones importantes entre las diferentes técnicas de emparejamiento siendo, en todos los casos, significativo al 1%. En segundo lugar, la existencia de errores estándar relativamente bajos y por último, el empleo, en el emparejamiento, de un alto porcentaje de las observaciones totales (no citados en el cuadro).

Tabla 2. Impacto “neto” de recibir una beca en el rendimiento educativo

	Estadístico t			Individuos	
	ATT	Analítico	<i>Bootstrapping</i>	Tratados	Control
Vecino M.P.(1)	0,234 ^a	5,063	(*)	150	528
Error estándar		0,046			
Vecino M.P.(2)	0,235 ^a	5,085	(*)	150	528
Error estándar		0,046			
Estratificación	0,257 ^a	5,838 ^a	6,002	150	621
Error estándar		0,044	0,043		
Kernel (**)	0,243 ^a		5,580	150	621
Error estándar			0,044		
Radius	0,222 ^a	4,630	4,529	150	550
Error estándar		0,048	0,049		

^a Significativa al 1%.

Se expresan los valores medios de los resultados obtenidos para 20 bases de datos imputadas. Todas las estimaciones se han realizado con reemplazo y soporte común. (1) Vecino más próximo con selección aleatoria. (2) Vecino más próximo con ponderativos idénticos. (*) No se calculan siguiendo a Abadie e Imbens (2006). (**) En el caso de la aproximación Kernel mediante la función de densidad Epanechnikov kernel el resultado es de 0,251.

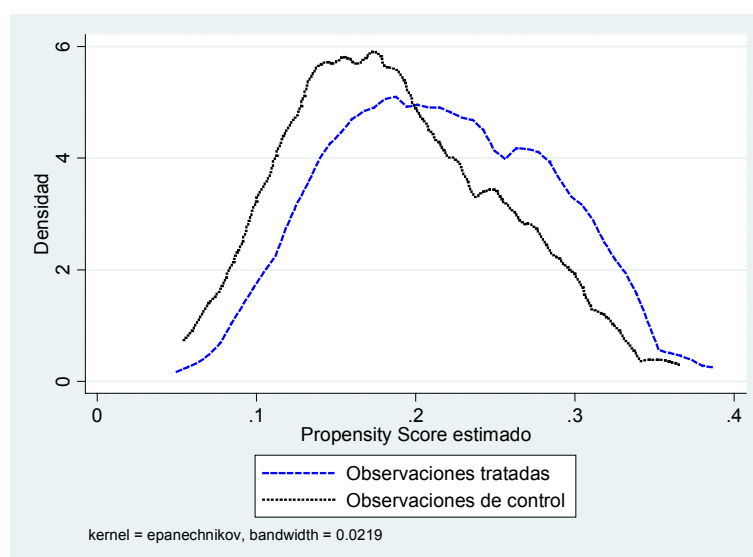
Fuente: Elaboración propia a partir de microdatos de EUSILC LONGITUDINAL UDB 2006 – version.2 – de Marzo 2009.

5.2 Comprobación de la conveniencia de la imposición del soporte común

Para comprobar que la aplicación del soporte común es una pre-condición válida para realizar las estimaciones, se realiza un análisis visual a través de un gráfico que muestra las funciones de densidad Kernel para el caso de las observaciones tratadas y de control. (véase gráfico

1). Se observa que para casi la totalidad de los *Propensity Scores* estimados se encuentra una observación tratada y de control. Esto sugiere que el supuesto de soporte común permite mejorar la calidad de los emparejamientos sin perder observaciones relevantes, que se reafirma conociendo que el 98,47% de la base original queda incluida en el soporte común y sólo 12 de las 783 observaciones originales se excluyen del análisis de comparación.

Gráfico 1. Densidad de Kernel para las observaciones tratadas y de control



Fuente: Elaboración propia en base a EUSILC LONGITUDINAL UDB 2006 – version.2 – de Marzo 2009.

6 Análisis de la robustez de la estimación

La robustez de las estimaciones dependen, básicamente, de la calidad de los emparejamientos realizados y la validez de los supuestos establecidos, entre los que destaca el referido a la independencia condicional (CIA -*Conditional Independence Assumption*-). Para intentar asegurar la existencia de ambos requisitos, en la presente estimación se realiza la comprobación centrada en tres elementos: la calidad de los emparejamientos, la correcta estimación de la varianza y la sensibilidad de los resultados.

6.1 Calidad de los emparejamientos

Una de las opciones que plantea la literatura, con el objeto de conocer si la aplicación de la metodología cuasi-experimental ha permitido generar dos grupos de observación similares, es la realización de un análisis de diferencia de medias para ambos grupos. En caso de obtener un resultado satisfactorio, se podría asegurar que los parámetros estimados de impacto realmente estarían detectando la diferencia existente en la variable resultado sólo como consecuencia del tratamiento realizado.

Tabla 3. Test de diferencia de medias. Observaciones tratadas y de control

Variables	Media de las observaciones		Test de diferencia de medias	
	Tratadas	Control	Estadístico t	Resultado
<i>Individuales</i>				
Mujer	0,53	0,47	-1,22	No Sig.
Último_Tri	0,25	0,26	0,38	No Sig.
Enf_Crónica	0,09	0,07	-0,96	No Sig.
Posición	1,46	1,59	2,10	Sig.
<i>Padre/Madre</i>				
Educ_Padre	2,33	2,48	1,09	No Sig.
Educ_Madre	2,24	2,30	0,50	No Sig.
Ocup_Padre (*)	Valor sig. 3,29	3,08	-2,51	Sig.
	Valor no sig. 3,30	3,15	-1,29	No Sig.
Ocup_Madre	2,99	2,74	-1,75	Sig.
Activo_Padre	0,92	0,91	-0,41	No Sig.
Activo_Madre	0,49	0,51	0,56	No Sig.
<i>Hogar</i>				
Nro. Hermanos	1,18	1,35	2,11	Sig.
Quintil	2,45	2,67	1,80	Sig.
Vivienda_Prop	0,89	0,83	-1,97	Sig.
Problemas_Estruct	0,19	0,17	-0,34	No Sig.
Vacaciones	0,53	0,54	0,37	No Sig.
Más_de_4_Amb	0,77	0,70	-1,72	Sig.
<i>Entorno</i>				
Baja_Media_Urb	0,65	0,52	-2,92	Sig.

La estimación tiene en cuenta los valores de las 20 bases de datos imputadas. (*) Debido a que existen diferencias de significatividad según la base imputada utilizada se muestran los valores extremos de cada opción. En el caso de las variables dummies se emplea el test de diferencia de proporciones.

Fuente: Elaboración propia en base a EUSILC LONGITUDINAL UDB 2006 – versión 2 – de Marzo 2009.

Los resultados obtenidos indican que, de las 17 variables incorporadas en el análisis, siete muestran diferencias significativas, nueve presentan diferencias no significativas y una, la referida a la ocupación del padre, no ofrece un resultado concluyente (véase tabla 3). Seguidamente, se analiza con detalle aquellas variables con diferencias significativas en sus valores medios.

Casi la totalidad de variables presentan diferencias que, aunque estadísticamente significativas, no son relevantes en sus valores absolutos. En esta clasificación se ubicarían las variables referidas a la posición del individuo, la ocupación de la madre, el número de hermanos, el quintil de ingresos y la tenencia en propiedad de la vivienda. En segundo lugar, la variable referida a la ubicación del hogar (Baja_Media_Urb) se revela como la única con una diferencia significativa

y relevante en sus medias. Aún así, un análisis global a partir de la totalidad de variables tenidas en cuenta, permite concluir que el análisis de comparación se realiza a partir de dos grupos similares y, por tanto, que el efecto en el rendimiento educativo encontrado proviene realmente de la posibilidad de haber gozado de algún programa de becas y/o ayudas al estudio durante el período analizado (2004-2006).

6.2 *Estimación de la varianza*

El cálculo de la significatividad de los parámetros estimados y el cómputo de los errores estándar no resulta una tarea sencilla, debido a que la estimación de la varianza debe incluir la varianza generada en la estimación del PS, la imputación del soporte común y, en el caso de realizar una estimación sin reemplazo, también el orden en que las observaciones tratadas son emparejadas.

Siguiendo la literatura empírica se ha aplicado las diferentes maneras de estimar la varianza, sea por la vía analítica o mediante el *bootstrapping*, sugerido por Lechner (2002). Las diferentes aproximaciones con sus similares resultados, reflejados en los coeficientes para el test t (véase tabla 2), indican un alto grado de certeza en los parámetros determinados.

6.3 *Sensibilidad de los resultados*

El método de emparejamiento, que sirve como herramienta para evaluar el impacto de una política, tiene como base teórica el supuesto de la independencia condicional (CIA), que se basa en suponer que, dadas las características observables, la participación en la política o tratamiento es independiente de los valores potenciales de la variable objetivo. En este marco, comprobar la sensibilidad de los resultados estimados con respecto a posibles desviaciones que pudieran generarse por la acción de componentes no observables se ha convertido en un elemento relevante en la literatura empírica de los últimos años.

Debido a que directamente no es posible realizar esta comprobación, Rosebaum (2002) desarrolla una metodología teórica alternativa que se centra en responder si la inferencia sobre el efecto de una política puede ser alterada por factores no observados. Claramente, si el resultado se revela muy sensible, el investigador debería valorar la validez del supuesto testeado y considerar otras aproximaciones metodológicas. Empíricamente, esta metodología ha sido plasmada en dos aproximaciones diferentes que se desarrollan a continuación.

6.3.1 *Primera aproximación empírica*

Esta primera aproximación la proponen Becker y Caliendo (2007). Los autores parten del supuesto de ausencia de sesgo de selección que se asume en toda la estimación del *Propensity Score Matching*. Tal supuesto significa, en su desarrollo teórico, tener un *odds ratio* ($e^{\gamma(U_i - U_j)}$) igual a 1. Este resultado se obtendría a partir de dos vías: que las variables no observables no

tengan influencia en la posibilidad de ser tratado ($\gamma = 0$) o bien, a causa de la ausencia de diferencias entre las variables no observables para los dos grupos comparados ($U_i - U_j$). El análisis de sensibilidad, en este caso, se basa en evaluar como los valores estimados pueden verse afectados por una alteración artificial que simule la existencia de ambas fuentes de sesgo de selección.

En el caso aquí presentado (variable resultado binaria) y siguiendo el desarrollo de Aakvik (2001), Becker y Caliendo emplean para realizar el análisis el test estadístico no paramétrico de Mantel y Haenszel (MH). La metodología propone estimar un modelo inicial donde no existe sesgo de selección, que servirá como parámetro de comparación. Luego, se replica la prueba para diferentes niveles de sesgo con el objetivo de analizar la sensibilidad de los resultados a los mismos.

Los valores obtenidos, siguiendo esta metodología, se observan en la tabla 4. La interpretación de los mismos es la siguiente. Si comparamos dos individuos con las mismas características observables, un *odds ratio* igual a 1,25 significa que se simula un sesgo de selección en cuanto a la participación en el tratamiento de un 25% por encima de la situación sin sesgo (que es la aplicada para la estimación real del parámetro de impacto). En este caso se observa que se debe llegar a simular un sesgo que multiplica por 2,25 la situación inicial para encontrar un primer caso donde el parámetro de impacto obtenido (ATT) se revele sensible a este sesgo. Esto indica que sería necesario un sesgo que más que duplique las posibilidades de participación en el programa evaluado para generar un resultado no robusto. Tal extremo no es esperable cuando la estimación ha tenido en cuenta una serie de variables individuales y del entorno del individuo y, por lo tanto, esta prueba sugiere una alta estabilidad de los resultados ante el sesgo de selección.

Tabla 4. Análisis de sensibilidad del impacto estimado según diferentes *odds ratios*

<i>Odd ratios</i>	Chi-cuadrado	Sensibilidad
1	3,39 – 3,39	(No sesgo)
1,25	2,90 – 3,95	No sensible
1,50	2,49 – 4,39	No sensible
1,75	2,14 – 4,77	No sensible
2	1,84 – 5,10	No sensible
2,25	1,58 – 5,40	Sensible al 5%

La estimación tiene en cuenta los valores de las 20 bases de datos imputadas. Se citan los resultados obtenidos en la base estimada donde la sensibilidad al sesgo resultó significativa con el menor *odds ratio*.

Fuente: Elaboración propia en base a EUSILC LONGITUDINAL UDB 2006 – versión 2 – de Marzo 2009.

6.3.2 Segunda aproximación empírica

La segunda aproximación empírica la sugieren Ichino et. al. (2006) y Nannicini (2007). Se instrumenta mediante el cálculo del efecto medio del tratamiento en los tratados (ATT) bajo diferentes escenarios posibles de desvío del supuesto de independencia condicional (CIA), con el objetivo de comprobar la robustez de las estimaciones.

Se supone que la CIA no se satisface con los valores observables pero que podría serlo a partir de la incorporación de una variable binaria adicional. Esta nueva variable puede ser simulada a partir de la fijación *ad-hoc* de los parámetros que caracterizan su distribución o bien, utilizar los parámetros de una variable observable presente en la base de datos. En ambos casos, el objetivo es simular la capacidad de generar sesgo que podría tener una supuesta variable no observable. La comparación entre el valor del impacto neto calculado originariamente y el que surge a partir de la incorporación de esta nueva variable en la estimación, permite conocer la estabilidad del resultado estimado. En este caso se ha simulado la distribución de una variable no observable a partir de la distribución de las diferentes variables observables en la base de datos con características binarias.

La interpretación de los resultados de las variables no observables simuladas proviene del análisis de las diferencias encontradas entre el impacto original y los impactos que incorporan la supuesta variable no observada (véase tabla 5). Ichino et. al. (2006) plantean una diferencia mayor al 5% entre los parámetros obtenidos como valor umbral para evaluar el grado de sensibilidad de los resultados. En este caso, observamos que de todas las pruebas realizadas sólo en dos casos una variable simulada no observada provoca un diferencial superior al 5% en el ATT estimado. Son los casos de las variables no observables que simulan la distribución de dos variables existentes: la referida a la tenencia en propiedad de la vivienda (6,8%) y aquella que indica el grado de urbanización donde habita el individuo analizado (10,25%). En el primer caso, la variable no observable simulada debería ser lo suficientemente relevante para más que triplicar el rendimiento educativo y casi duplicar las posibilidades de ser tratado. En el segundo caso, la supuesta variable no incorporada al modelo original debería reducir el rendimiento educativo a la mitad e multiplicar por 1,8 las posibilidades de ser tratado.

En conclusión, se observa que la estimación realizada muestra una destacada estabilidad a la incorporación de variables no observable y, por tanto, se confirma la robustez del parámetro estimado.

Tabla 5. Análisis de sensibilidad del impacto estimado a partir de la simulación de variables no observables

Variable no observada simulada	Efecto resultado (neutral = 1)	Efecto selección (neutral = 1)	ATT	Error Estándar
Sin variables no observadas (*)	-	-	0,234	0,046
Incorporación de una variable no observada simulada a partir de una variable observable				
Mujer	0,9	1,3	0,235	0,050
Último_Tri	1,2	0,9	0,236	0,050
Enf_Crónica	0,6	1,4	0,237	0,048
Activo_Padre	3,0	1,8	0,227	0,049
Activo_Madre	1,3	0,8	0,239	0,051
Vivienda_Prop	3,3	1,9	0,218	0,049
Problemas_Estruc	0,4	1,1	0,235	0,050
Vacaciones	4,4	1,0	0,244	0,053
Más_de_4_Amb	1,4	1,5	0,229	0,050
Baja_Media_Urb	0,5	1,8	0,258	0,051

(*) Estimación original. Caso Vecino más Próximo con selección aleatoria. El efecto resultado y el efecto selección muestran el grado de desviación que cada variable no observada incorporaría al modelo original. Debido a que los mismos se expresan en odds ratio un valor 1 equivaldría a decir que la variable no observada no incorpora ningún grado de desviación. En la medida que los valores difieren de 1, el sesgo potencial que incorpora cada variable es mayor en relación al resultado obtenido y a la posibilidad de ser tratado.

7 Conclusiones

El objetivo de la comunicación era comprobar la importancia de las becas y ayudas al estudio como un elemento determinante de la continuidad escolar de los alumnos del nivel secundario post-obligatorio en España.

La motivación de la misma proviene de la gran cantidad de individuos que no continúan sus estudios más allá de la escolarización obligatoria y que se cuantifica mediante una alta tasa de abandono escolar prematuro y; en segundo lugar, de la necesidad de impulsar estudios y análisis que evalúen la efectividad real de las políticas públicas.

La aplicación del PSM permite comprobar la efectividad de las becas y ayudas al estudio otorgadas debido a su efecto positivo sobre el nivel educativo logrado por los individuos de 19 años. Las diferentes técnicas de emparejamiento calculadas indican la existencia de un efecto medio del tratamiento que es positivo y significativamente diferente de cero, con lo que se puede afirmar que existe un efecto “neto” positivo de las becas en las personas beneficiarias.

La robustez de las estimaciones queda asegurada a partir de la comprobación de la calidad de los emparejamientos y la validez de los supuestos establecidos, mediante el análisis de la propia calidad de los emparejamientos, la correcta estimación de la varianza y la baja sensibilidad de los resultados a la existencia de sesgos.

8 Referencias

- Aakvik, Arild. 2001. "Bounding a matching estimator: the case of a Norwegian training program." *Oxford Bulletin of Economics and Statistics*, 63:1, pp. 115-43.
- Abadie, Alberto y Guido Imbens. 2006. "On The Failure of the Bootstrap for Matching Estimators." NBER Technical Working Paper Series No. 325.
- Acock, Alan. 2005. "working with Missings Values." *Journal of Marriage and Family*, 67, pp. 1012-28.
- Ambler, Gareth y Rumana Omar. 2007. "A comparison of imputation techniques for handling missing predictor values in a risk model with a binary outcome." *Statistical Methods in Medical Research*, 16, pp. 227-98.
- Baydar, Nazli y Jeanne Brooks-Gunn. 1991. "Effects of maternal employment and child-care arrangement on preschoolers' cognitive and behavioral outcomes: Evidence from the childrens of the National Longitudinal Survey of Youth." *Developmental Psychology*, 27, pp. 932-45.
- Becker, Sasha y Marco Caliendo. 2007. "Sensitivity Analysis for Average Treatment Effects." *Stata Journal*, 7:1, pp. 71-83.
- Becker, Sasha y Andrea Ichino. 2002. "Estimation of average treatment effects based on propensity scores." *The Stata Journal*, 2:4, pp. 358-77.
- Bedi, Arjun y Jeffery Marshall. 2002. "Primary School attendance in Honduras." *Journal of Development Economics*, 69, pp. 129-53.
- Blundell, Richard y Monica Costa Dias. 2000. "Evaluation Methods for Non-Experimental Data." *Fiscal Studies*, 21:4, pp. 427-68.
- Blundell, Richard y Monica Costa Dias. 2002. "Alternative approaches to evaluation in empirical microeconomics." *Portuguese Economic Journal*, 1, pp. 91-115.
- Calero, Jorge y J.Oriol Escardíbul. 2007. "Evaluación de servicios educativos: el rendimiento en los centro públicos y privados medido en PISA-2003." *Hacienda Pública Española* 183:4, pp. 33-66.
- Caliendo, Marco y Sabine Kopeinig. 2008. "Some Practical Guidance for the Implementation of Propensity Score Matching " *Journal of Economic Surveys*, 22:1, pp. 31-72.
- Chevalier, Arnaud. 2004. "Parental Education and Child's Education: A Natural Experiment." *Discussion Paper Series. IZA:1153*.

- Comi, Simona. 2003. "Intergenerational Mobility in Europe: Evidence from ECHP." Working Papers del Dipartimento de Economia Politica e Aziendale. Università degli Studi di Milano:3, pp. 29.
- Cook, Thomas y Donald Campbell. 1979. Quasi-experimentation: design and analysis issues for field settings. Chicago: Rand McNally College Pub. Co.
- Eren, Ozkan y Daniel Millimet. 2008. "Time to learn? The organizational structure of schools and student achievement," in *The Economics of Education and Training*. Dustmann Christian, Bernd Fitzenberger and Stephen Machin eds. Meppel, The Netherlands: Physica-Verlag Heidelberg.
- Gang, Ira y Klaus Zimmerman. 2000. "Is Child like Parent? Educational Attainment and Ethnic Origin." *The Journal of Human Resources*, 35:3, pp. 550-69.
- Garib, Geetha, Teresa Martín Garcia, and Jaap Dronkers. 2007. "Are the effects of different family forms on children's educational performance related to the demographic characteristics and family policies of modern societies?," in *Changing families and their lifestyles*. Hester Moerbeek, Anke Niehof and Johan van Ophem eds. Netherlands: Wageningen Academic Publishers, pp. 27-50.
- Heckman, James, Robert Lalonde, y Jeffrey Smith. 1999. "The Economics and Econometrics of Active Labor Market Programs," in *Handbook of Labor Economics*. A. Ashenfelter and D. Card eds: Elsevier Science, pp. 1865-2097.
- Ichino, Andrea, Fabrizia Mealli, y Tommaso Nannincini. 2006. "From Temporary Help Jobs to Permanent Employment: What Can We Learn from Matching Estimators and their Sensitivity?" CERP, Discussion Paper N°5736.
- Lechner, Michael. 2002. "Program Heterogeneity and Propensity Score Matching: An Application to the Evaluation of Active Labor Market Policies." *The Review of Economics and Statistics*, 84:2, pp. 205-20.
- Nannincini, Tommaso. 2007. "A Simulation-Based Sensitivity Analysis for Matching Estimators." *Stata Journal*, 7:3, pp. 334-50.
- Price, Joseph. 2008. "Parent-Child Quality Time. Does Birth Order Matter?" *Journal of Human Resources*, 43:1, pp. 240-65.
- Reiter, Jerome y Trivellore Raghunathan. 2007. "The Multiple Adaptations of Multiple Imputation." *Journal of the American Statistical Association*, 102:480, pp. 1462-71.
- Rivkin, Steven, Eric Hanushek, y John Kain. 2005. "Teachers, Schools, and Academic Achievement." *Econometrica*, 73:2, pp. 417-58.
- Rosenbaum, Paul. 2002. *Observational Studies*. New York: Springer.
- Rosenbaum, Paul y Donald Rubin. 1983. "The Central Role of the Propensity Score in Observational Studies for Causal Effects." *Biometrika*, 70:1, pp. 41-55.

- Royston, Patrick. 2004. "Multiple imputation of missing values." *The Stata Journal*, 4:3, pp. 227-41.
- Royston, Patrick. 2005. "Multiple imputation of missings values: update." *The Stata Journal*, 5:2, pp. 1-14.
- Rubin, Donald. 1974. "Estimating causal effects of treatment in randomized and nonrandomized studies." *Journal of Educational Psychology*, 66, pp. 685-701.
- Rubin, Donald. 1996. "Multiple Imputation After 18+ Years." *Journal of the American Statistical Association*, 91:434, pp. 473-89.
- Santin, Daniel. 2006. "La medición de la eficiencia de las escuelas: una revisión crítica." *Hacienda Pública Española* 177:2, pp. 57-82.
- Schafer, Joseph. 1999. "Multiple imputation: a primer." *Statistical Methods in Medical Research*, 8, pp. 3-15.
- Sullivan, Alice. 2001. "Cultural Capital and Educational Attainment." *Sociology*, 35:4, pp. 893-912.
- Summers, Anita y Barbara Wolfe. 1977. "Do schools Make a Difference?" *The American Economic Review*, 67:4, pp. 639-52.
- Van Buuren, S., H. Boshuizen, and D. Knook. 1999. "Multiple imputation of missing blood pressure covariates in survival analysis." *Statistics in Medicine*, 18, pp. 681-94.
- Vandenberghe, V. 1999. "The need to go beyond human capital theory and production-function analysis." *Educational Studies*, 25:2, pp. 129-43.