

## ¿Quién se beneficia más de los estudios universitarios? Evidencia de selección negativa en los rendimientos económicos heterogéneos de la educación universitaria

### (Who Benefits Most from College? Evidence for Negative Selection in Heterogeneous Economic Returns to Higher Education)

Jennie E. Brand  
*University of California-Los Angeles*

Yu Xie  
*University of Michigan*

#### Resumen

En este artículo analizamos cómo varía el rendimiento económico de la educación universitaria en la población estadounidense. Siguiendo los principios de ventaja comparativa, generalmente los investigadores asumen que se da una selección positiva, esto es, que los individuos con mayor probabilidad de ir a la universidad son también quienes obtienen mayores réditos de dicha educación. Nuestro análisis sugiere que las personas con la *menor* probabilidad de obtener una educación universitaria son quienes más se benefician de ésta, controlando por factores económicos y no económicos que influyen en la asistencia a la universidad. Denominamos a esta teoría como *hipótesis de selección negativa*. Para decidir entre las dos hipótesis, estudiamos los efectos de haber finalizado la universidad sobre los ingresos por estrato de un *score* de propensión, utilizando un modelo lineal jerárquico innovador con datos provenientes de la Encuesta Nacional Longitudinal de Jóvenes de 1979 y el Estudio Longitudinal de Wisconsin. En ambas cohortes, tanto en mujeres como en hombres, en cada estadio observado a lo largo de la vida, encontramos evidencia que sugiere la selección negativa. Los resultados de los análisis auxiliares proporcionan mayor sustento a la hipótesis de selección negativa.

La expansión de la educación es uno de los rasgos más claros, perdurables y trascendentes de la sociedad moderna. Al considerar la importante expansión educativa que se produjo en los Estados Unidos durante el siglo XX, especialmente en el nivel posterior a los estudios secundarios, Fischer y Hout (2006:247) concluyen que “ha crecido la división entre los individuos menos y más educados y ha surgido como un determinante potente de las oportunidades vitales y los estilos de vida.” En 2007, el Bureau de Censos de los Estados Unidos comunicó que los graduados universitarios ganaban aproximadamente un promedio de \$55,000, en tanto que las personas que sólo habían finalizado sus estudios secundarios ganaban menos de \$30,000. Los científicos sociales se han interesado durante mucho tiempo en las cuestiones relacionadas con el acceso a la educación universitaria y su impacto (por ejemplo, Blau y Duncan 1967; Hout 1988; Hout y DiPrete 2006). Los investigadores han formulado las siguientes preguntas: (1) ¿Qué atributos familiares e individuales se asocian al acceso a la educación universitaria? Y (2) ¿Cuáles son los efectos causales de la educación universitaria en los resultados socioeconómicos subsiguientes? En el modelo de los agentes económicos racionales, usual en la bibliografía económica, las preguntas anteriores se entrelazan entre sí: los individuos deciden si emprender o no la educación universitaria sobre la base de un análisis de costo-beneficio. La gente elige la educación universitaria sólo si ésta aumenta sus expectativas de incremento de los ingresos durante la vida (Becker 1964; Card 1995, 2001; Heckman y Honoré 1990; Manski 1990; Mincer 1974; Willis y Rosen 1979). En otras palabras, si se deja de lado la información incorrecta, las restricciones a los préstamos para financiar la educación y la incertidumbre, los individuos eligen acceder a la educación universitaria de acuerdo con los rendimientos económicos esperados. La gente concurre a la universidad sólo si los rendimientos económicos superan los costos. Aunque este paradigma de maximización de las utilidades puede incluir, en principio, factores no económicos, los científicos partidarios de este enfoque rara vez consideran estos factores al investigar la educación universitaria.<sup>1</sup> En caso de que los factores económicos fueran los determinantes principales de la concurrencia a la universidad, los individuos con mayor probabilidad de concurrir a la universidad también deberían ser los que más se beneficiaran con la educación universitaria (Carneiro, Hansen y Heckman 2003; Carneiro, Heckman y Vytlačil 2001, 2007; Heckman, Urzua y Vytlačil 2006; Willis y Rosen 1979). Denominamos a este enunciado como *hipótesis de selección positiva*.

La bibliografía sociológica generalmente trata separadamente las dos preguntas planteadas anteriormente, porque considera que la educación universitaria es un estatus adquirido, sujeto a la influencia de numerosos factores (Boudon 1974; Bourdieu 1977; Bowles y Gintis 1976; Coleman 1988; DiMaggio

1982; Jencks et al. 1972; Lucas 2001; MacLeod 1989; Mare 1981; Morgan 2005; Sewell, Haller y Ohlendorf 1970). O sea que un tema clave para estos estudios es que el comportamiento que consiste en asistir a la universidad está determinado no sólo por una elección racional, sino también por normas y circunstancias culturales y sociales (Coleman 1988). Así, los mecanismos que influyen en la concurrencia a la universidad pueden diferir según el contexto social. Para algunos individuos que provienen de sectores sociales favorecidos, la universidad es un resultado culturalmente esperado. Para este grupo, la universidad está vinculada a la ganancia económica de un modo menos exclusivo e intencional que para la gente de los grupos menos favorecidos, para quienes la educación universitaria es una novedad que puede requerir una justificación económica (Beattie 2002; Boudon 1974; Smith y Powell 1990). Además, las perspectivas de ingresos de los trabajadores con menor nivel educativo son poco prometedoras, particularmente si provienen de contextos sociales desfavorecidos.

Por el contrario, la gente que proviene de contextos sociales favorecidos tiene una elevada probabilidad de asistir a la universidad y una perspectiva de tener ingresos relativamente altos. Cuando eliminamos la influencia de las covariables observadas que ayudan a predecir la educación universitaria, es posible que, debido a mecanismos de selección diferencial y a perspectivas de ganancias, los individuos que tienen *menos* probabilidades de concurrir a la universidad se benefician más con la educación universitaria. Denominamos a esta conjetura como *hipótesis de selección negativa*.

Para decidir entre las hipótesis de selección positiva y negativa, llevamos a cabo un estudio empírico que analizó datos de dos grandes estudios longitudinales de los Estados Unidos: la Encuesta Longitudinal Nacional de Jóvenes de 1979 y el Estudio Longitudinal de Wisconsin de 157 cohortes. El uso de estas dos fuentes de datos nos permitió controlar las fortalezas y debilidades relativas de las bases de datos así como la calidad de las variables de control disponibles y establecer la robustez de nuestros resultados. También nos permitió estudiar las posibles diferencias en el rendimiento entre las cohortes a lo largo de toda la vida. Como los individuos invierten en educación universitaria con la expectativa de obtener beneficios económicos en el curso de su vida (Mincer 1974), es importante considerar la variación en los rendimientos de la educación universitaria a lo largo de toda la vida.

Utilizamos un enfoque metodológico de tres pasos. Primero, consideramos un supuesto de ignorabilidad que implica que, después de haber controlado un conjunto rico de covariables observadas, no existen factores de confusión adicionales entre individuos que han completado la educación universitaria y los que no lo han hecho. Con el supuesto de ignorabilidad, resumimos en *scores* de propensión estimada las diferencias sistemáticas en las covariables entre concurrentes y no concurrentes a la universidad (Rosenbaum y Rubin 1983, 1984; Rubin 1997). En segundo lugar, estimamos los efectos de haber completado la educación universitaria sobre los ingresos por estratos de *scores* de propensión y estudiamos los patrones de estos efectos con un modelo lineal jerárquico (Xie y Wu 2005). Este paso innovador clave nos permitió encontrar un patrón positivo o negativo entre los efectos de una educación universitaria y la probabilidad de obtener una educación universitaria. En tercer lugar, reexaminamos el supuesto de ignorabilidad y llevamos a cabo análisis auxiliares que ayudaron a nuestra interpretación de los resultados. En un análisis de sensibilidad, omitimos varias covariables clave para estudiar las consecuencias de violar el supuesto de ignorabilidad. Todos los análisis se realizaron separadamente para hombres y mujeres.

## **CUESTIONES TEÓRICAS Y METODOLÓGICAS EN LA EVALUACIÓN DE LOS RENDIMIENTOS DE LA EDUCACIÓN UNIVERSITARIA**

### **DOS FUENTES DEL SESGO DE SELECCIÓN**

Se conoce en la bibliografía sobre inferencia causal, pero rara vez se admite en la investigación empírica sociológica, que existen dos tipos de sesgo de selección en los datos observacionales (Morgan y Winship 2007). El primer tipo se debe a la heterogeneidad de las condiciones preexistentes o atributos que se asocian tanto a la condición del tratamiento como al resultado. En el caso de los rendimientos económicos de la educación superior, atributos como la capacidad mental y los hábitos de trabajo pueden asociarse positivamente a la probabilidad de acceder a la educación superior y a mayores ingresos. El segundo tipo de sesgo de selección se debe a la heterogeneidad de los efectos del tratamiento, es decir, las diferencias sistemáticas en el efecto causal de la educación universitaria sobre los ingresos, entre los individuos que recibieron y los que no recibieron educación universitaria. Los rendimientos económicos de la educación superior deberían variar entre los miembros de una sociedad (Card 1999) porque es inverosímil suponer

que diferentes miembros de una población respondan de manera idéntica al tratamiento de la educación universitaria. Esta investigación se aparta del supuesto de la homogeneidad de la población y se centra en la variabilidad a nivel del grupo, al clasificar a los individuos según su probabilidad estimada de completar la educación universitaria.<sup>2</sup> Sobre la base de los atributos observados, nos preguntamos si los individuos que tienen más probabilidades de acceder a una educación universitaria obtienen rendimientos más altos o más bajos de este nivel educativo, comparados con los individuos que tuvieron menos probabilidades de acceder a estos estudios. Este enfoque nos permite examinar la asociación potencial entre dos fuentes de heterogeneidad de la población.

Para ilustrar la primera fuente del sesgo, comencemos con un modelo estándar en el que el efecto de una educación universitaria es homogéneo. Para la persona  $i$ , la siguiente función de regresión descompone los ingresos registrados observados en la suma de tres factores (más una constante) —el efecto de tratamiento de la educación superior, una combinación lineal de covariables y el residuo:

$$y_i = \alpha + \delta d_i + \beta'X_i + U_i, \quad (1)$$

donde  $y$  es el logaritmo natural de los ingresos,  $d$  es una variable *dummy* que representa el hecho de que el encuestado haya finalizado la universidad (1: sí; 0: cualquier otra respuesta),  $X$  es un vector de determinantes de ingresos que también podría influir en la probabilidad de finalizar los estudios universitarios y  $U$  es el residuo no explicado por el modelo de base. Los parámetros  $\beta$  son coeficientes de regresión que miden los cambios en el logaritmo de los ingresos asociados a cambios en los determinantes  $X$  de los ingresos que, en general, incluyen varias mediciones de estatus socioeconómico familiar, residencia geográfica, logros académicos y, en muchos estudios, algunas mediciones de capacidad mental. La transformación exponencial del coeficiente de regresión  $\delta$  representa el aumento multiplicativo en los ingresos asociado a la obtención de un grado universitario, *ceteris paribus*. En la ecuación 1, se asume que  $\delta$  es un parámetro constante desconocido, invariante entre todos los miembros de la población. Las extensiones de este modelo estándar pueden incorporar factores explicativos adicionales de los ingresos, como la selectividad institucional, la especialización académica y el rendimiento académico (Brewer, Eide y Ehrenberg 1999; Dale y Krueger 2002; Thomas 2003; Thomas y Zhang 2005; Zhang 2005).

Si la homogeneidad es cierta, la principal amenaza a la inferencia causal reside en que una regresión por el método de Mínimos Cuadrados Ordinarios (MCO) de  $y$  sobre  $d$ , incluso si se controla por el efecto  $X$ , está sujeta a la primera fuente del sesgo de selección, debido a una correlación no nula entre  $U$  y  $d$  (Griliches 1977). De acuerdo con el supuesto de homogeneidad, el saber convencional sostiene que los estimadores de MCO del rendimiento económico de la educación universitaria están sesgados positivamente (Griliches 1977; Hauser y Daymount 1977), porque factores como la capacidad no observable y la ética de trabajo deberían afectar positivamente tanto la educación como los ingresos. Sin embargo, no se ha establecido empíricamente la dirección real del sesgo. Por ejemplo, Ashenfelter y Krueger (1994) sostienen que los estimadores de MCO del efecto de la educación sobre los ingresos están sesgados negativamente porque estos estimadores con frecuencia están por debajo de los cálculos de la variable instrumental de los rendimientos de la educación universitaria. Sin embargo, si le restamos importancia al supuesto poco realista de homogeneidad, no es fácil responder si los estimadores de MCO están sesgados positiva o negativamente: el estimador de MCO es esencialmente un promedio ponderado de efectos heterogéneos, algunos de los cuales son necesariamente mayores y otros menores que el promedio poblacional (Angrist y Krueger 1999; Morgan y Winship 2007). Con esta conceptualización más realista de la heterogeneidad que subyace a los rendimientos de la educación, los individuos difieren no sólo en los atributos de su contexto de origen, sino también en los beneficios económicos que obtienen de la educación universitaria.

Para estudiar de manera sistemática los efectos de tratamiento heterogéneo de la educación superior sobre los ingresos, hemos adoptado un enfoque simple que utiliza covariables ricas y apela a la ignorabilidad, al menos provisionalmente. Este enfoque nos permite encontrar patrones empíricos de la heterogeneidad del efecto del tratamiento como función de las covariables observadas. Para estudiar los efectos heterogéneos del tratamiento por covariables observadas, es habitual examinar la interacción entre la educación y los factores específicos que influyen en los salarios y en la probabilidad de acceder a la educación universitaria, como la raza o el género (Barrow y Rouse 2005; Perna 2005; Welch 1973) o la educación u ocupación de los padres (Altonji y Dunn 1996; Hauser 1973; Olneck 1979). No obstante, cuando se comparan los rendimientos de la educación universitaria entre los individuos que la completaron y los que no lo hicieron, la interacción más significativa se observa entre la educación universitaria y la propensión

a finalizar la universidad (Heckman et al. 2006).<sup>3</sup> Agrupamos efectos heterogéneos de la universidad con los efectos medios de *scores* de propensión por nivel del grupo y observamos directamente las tendencias en los efectos (Xie y Wu 2005).<sup>4</sup>

El supuesto de ignorabilidad (también denominado de “selección en base a observables” o *unconfoundedness*) establece que los resultados potenciales no se correlacionan con el estatus de tratamiento, condicional a las covariables observables. El supuesto no puede verificarse nunca y, por consiguiente, no debería tomarse como cierto en la práctica para los datos observacionales. Su plausibilidad depende de la disponibilidad de las covariables observadas que difieren entre graduados universitarios y graduados no universitarios, y que también influye en los ingresos. Es razonable suponer que modelos que no controlan la capacidad cognitiva, por ejemplo, no satisfacen la ignorabilidad. Aún así, la medición de los factores de confusión significativos vuelve tentativamente más plausible la ignorabilidad, aunque no necesariamente verdadera. Si bien no consideramos que el supuesto de ignorabilidad sea cierto, los análisis realizados según este supuesto son lo mejor que los datos pueden darnos sin recurrir a supuestos adicionales no verificables.<sup>5</sup> Utilizando nuestra estrategia, nos centramos en las diferencias entre grupos en su propensión a completar la educación universitaria y adjudicamos entre dos patrones potenciales en los efectos heterogéneos observados de haber finalizado la educación universitaria sobre los ingresos: selección positiva (los individuos que tienen más probabilidades de beneficiarse con la universidad tienen mayores probabilidades de finalizar la universidad) versus selección negativa (los individuos que tienen más probabilidades de beneficiarse con la educación universitaria tienen *menos* probabilidades de finalizar la universidad).

#### SELECCIÓN POSITIVA VERSUS SELECCIÓN NEGATIVA

En economía, la teoría del capital humano es una explicación sobre la adquisición de la educación que ha ejercido una influencia importante (Becker 1964; Mincer 1974). La idea central de esta teoría es que una gradación de los ingresos por nivel de educación refleja los rendimientos de la inversión racional de los individuos en la educación. Si  $\lambda$  representa el valor presente del rendimiento económico de la concurrencia a la universidad durante la vida y  $c$  el costo de la universidad, concurrir a la universidad produce una ganancia neta si  $\lambda > c$ , con el beneficio así definido como  $\pi = \lambda - c$ . La asociación entre los rendimientos de la universidad y la decisión de concurrir a la universidad es nuclear en la bibliografía más reciente, que vincula la variación en los rendimientos de la educación al comportamiento heterogéneo de la educación universitaria. Sobre la base de los principios de la autoselección y de las ventajas comparativas, la tesis es que los individuos más “dignos de ir a la universidad”, en el sentido de que obtienen los mayores rendimientos de la universidad, son los que más probabilidad tienen de poder acceder a la universidad (Averett y Burton 1996; Carneiro, Hansen y Heckman 2003; Carneiro, Heckman y Vytlačil 2001, 2007; Roy 1951; Willis y Rosen 1979). Estos individuos también están en una mejor posición para cubrir los costos económicos de una educación universitaria, especialmente en instituciones más costosas (Zhang 2005). De acuerdo con estos estudios, la selección positiva debería producirse porque los individuos que están en condiciones de beneficiarse más con la educación universitaria tienen mayores probabilidades de acceder a la universidad.

En economía, la tesis de la selección positiva es aceptada de manera amplia, aunque no universal. En nuestra opinión, se trata más de un argumento teórico que de una propuesta que puede someterse realmente a pruebas empíricas. En economía, la investigación empírica sobre la elección se basa fuertemente en el marco de preferencia revelada (por ej., Manski y Wise 1983; Train 2003). Este marco, aplicado al problema que investigamos aquí, en esencia establece que un investigador puede inferir que  $\lambda > c$ , al menos como expectativa, si se observa que una persona completa la educación universitaria y  $\lambda \leq c$  en el caso contrario. Willis y Rosen (1979) usan esta estrategia en su estudio clásico que aplica el modelo de Roy (1951) a la cuestión de la educación universitaria: la diferencia en la utilidad esperada entre la educación universitaria y la secundaria determina la probabilidad de concurrir a la universidad. Más recientemente, Carneiro, Heckman y Vytlačil (2007) también presentaron evidencia que sugiere la autoselección positiva (es decir, los individuos con los mayores rendimientos esperados tienen mayores probabilidades de concurrir a la universidad).

Los sociólogos también reconocen la heterogeneidad de los rendimientos de la educación superior. Raftery y Hout (1993: 57), por ejemplo, sostienen que “parece probable que el beneficio percibido de la educación varíe entre individuos” en función de los atributos individuales. Como los economistas, los sociólogos infieren que la elección de concurrir a la universidad puede ser el resultado de un análisis costo-beneficio (Boudon 1974; Breen y Goldthorpe 1997; Raftery y Hout 1993); sin embargo, enfatizan

que los costos y los beneficios no son meramente económicos. Por ejemplo, en términos de costos, los sociólogos han considerado la heterogeneidad tanto como carga financiera como de presión familiar derivada del hecho de desviarse de las normas culturales de clase (Boudon 1974; Raftery y Hout 1993).

En contraste con el modelo estrictamente económico de costo-beneficio de la concurrencia a la universidad, muchas investigaciones indican que múltiples actores y factores influyen en ella. A partir del modelo de Blau-Duncan, los sociólogos han reconocido el significado de numerosos factores de contexto familiar para el logro educativo, como la educación y la ocupación de los padres, la estructura familiar (McLanahan y Sandefur 1994) y la cantidad de hijos (Blake 1981).<sup>6</sup> Además, el “modelo de Wisconsin” de logro del estatus especifica los procesos concretos por los cuales el contexto familiar afecta el logro educativo: estatus socioeconómico familiar y capacidad medida afectan las aspiraciones ocupacionales y educativas, de la misma manera que lo hacen el estímulo de los padres y de otras personas importantes para el individuo (Hauser, Tsai y Sewell 1983; Sewell et al. 1970; Sewell, Haller y Portes 1969; Sewell y Hauser 1975).<sup>7</sup> También Coleman (1988) ayuda a comprender cómo los factores del contexto familiar influyen en los logros de los niños a través del concepto de capital social, esto es, las relaciones sociales que consisten en expectativas, canales de información y normas sociales. Apoyo, expectativas, información y normas difieren según el contexto familiar, lo que genera mecanismos diferenciales de acceso a la universidad (Morgan 2005).

Además, los sociólogos han desarrollado una perspectiva neomarxista del conflicto, que ayuda a explicar las diferencias en los logros educativos a través del contexto social. Por ejemplo, los estudiosos del capital cultural enfatizan la importancia del contexto familiar para los logros educativos, poniendo el acento en el contexto cultural general, el conocimiento, la disposición y las habilidades que los niños adquieren de sus padres. Los sociólogos también sostienen que las escuelas sistemáticamente premian el capital cultural de las clases favorecidas y devalúan el de las clases más bajas (Bourdieu 1977; DiMaggio 1982; Lareau 2003). Los teóricos de la reproducción social analizaron este tema en detalle: sostienen que las escuelas primarias y secundarias entrenan a los estudiantes aventajados para que se posicionen en el punto más alto del orden socioeconómico (por ejemplo, continuando los estudios después de la escuela secundaria), en tanto que condicionan a los pobres para que acepten su estatus inferior en la estructura de clases (Bowles y Gintis 1976; MacLeod 1989). En resumen, estos estudios sugieren que los individuos que provienen de contextos sociales altos tienen mayores probabilidades de concurrir a la universidad aun en ausencia de un análisis racional de costo-beneficio, en tanto que los individuos que provienen de contextos sociales bajos tienen que superar gran cantidad de obstáculos para concurrir a la universidad.

Los estudios realizados sobre estratificación social han aportado fundamentos teóricos y empíricos convincentes para postular una variación en los efectos de la educación sobre los ingresos de acuerdo con el contexto social. Dichos estudios muestran que la relación directa entre el origen social y los resultados finales (ambos medidos por el estatus ocupacional) es mucho más débil para los graduados universitarios que para los trabajadores sin títulos universitarios (Hout 1984, 1988). La figura 1 describe este patrón empírico. Si modificamos el punto de vista y consideramos los rendimientos de la educación (Goldthorpe y Jackson 2008) (por ejemplo, la diferencia en los resultados entre trabajadores con y sin educación universitaria) como una función del origen social, este patrón de interacción produce una diferencia menor por educación universitaria para los individuos de origen social alto ( $\delta_2$ ) que para los individuos de origen social bajo ( $\delta_1$ ). En otras palabras, los individuos que provienen de contextos sociales relativamente desfavorecidos, o los que tienen las menores probabilidades de completar la educación universitaria, se benefician *más* con la finalización de los estudios universitarios. Este patrón es el resultado de las perspectivas especialmente malas del mercado laboral para los trabajadores con niveles educativos bajos, combinadas con bajos niveles de otras formas de capital humano, social o cultural.<sup>8</sup> Esta tradición colectiva teórica y empírica conduce a nuestra hipótesis de selección negativa.

No somos los primeros sociólogos en argumentar sobre la posibilidad de patrones de selección negativa (Brand y Halaby 2006; Bryk, Lee y Holland 1993; DiPrete y Engelhardt 2004; Hoffer, Greeley y Coleman 1985; Morgan 2001; Tsai y Xie 2008). Los estudios muestran, por ejemplo, que el ambiente de la escuela secundaria tiene un efecto más fuerte sobre quienes tienen posibilidades marginales de concurrir a la universidad que sobre los estudiantes más favorecidos (Bryk et al. 1993; Hoffer et al. 1985). La literatura económica también proporciona evidencia empírica directa a favor de la selección negativa en la educación superior. Un estudio económico informa que una persona elegida al azar podría esperar recibir un 9 % de incremento salarial a causa de la educación universitaria, en tanto que los que realmente concurren a las universidades reciben aproximadamente un 4 % de aumento (Heckman, Tobias, y Vytlačil 2001). Además, los estudios que incluyen la consideración de leyes de escolaridad obligatoria,

diferencias en la accesibilidad de las escuelas o aspectos similares como variables instrumentales encuentran rendimientos económicos mayores que los estimadores de MCO. Esto sugiere mayores rendimientos de la educación para los individuos cuyas posibilidades de continuar su escolaridad son marginales.<sup>9</sup>

### MODELO DE COMPORTAMIENTO

Especificamos de la siguiente manera el modelo de comportamiento para la educación universitaria:  $d_i^*$  representa la posibilidad potencial de que la persona  $i$  complete sus estudios universitarios y  $d_i$  el resultado observado (1: sí; 0 cualquier otra respuesta). Constituye una práctica habitual relacionar a ambos mediante un modelo de medición del umbral:

$$d_i = 1 \text{ si } d_i^* > 0; \quad (2)$$

$$d_i = 0 \text{ para el complemento.}$$

Además, especificamos que el logro de una educación universitaria está determinado por un promedio ponderado de un componente económico  $\pi_i$ , un componente no económico  $\eta_i$  y un residuo  $\varepsilon_i$ :

$$d_i^* = w_i\pi_i + (1 - w_i)\eta_i + \varepsilon_i, \quad (3)$$

donde se supone que  $\varepsilon_i$  es independiente de  $\pi_i$ ,  $\eta_i$ , y  $w_i$ , con  $0 \leq w_i \leq 1$ . Una noción clave de la bibliografía sociológica es que el peso relativo  $w_i$ , que se otorga al componente económico, puede disminuir con el determinante no económico  $\eta_i$  (es decir, una correlación negativa entre ambos en la población). Además, suponemos que  $\pi_i$  es una función lineal de las covariables observadas ( $\lambda_1'X$ ) más un componente no observado  $\mu_i$ , y que  $\eta_i$  es una función lineal de las covariables observadas ( $\lambda_2'X$ ). Podemos reescribir la ecuación 3 de la siguiente manera:

$$d_i^* = w_i\lambda_1'X_i + (1 - w_i)\lambda_2'X_i + w_i\mu_i + \varepsilon_i. \quad (4)$$

La probabilidad de completar la educación universitaria es alta cuando  $d_i^*$  es grande. Si se formula en detalle el modelo de las ecuaciones 1 a 4, resulta más sencillo apreciar la diferencia clave entre los factores económicos y no económicos que influyen en el logro de una educación universitaria. En el modelo tradicional de comportamiento universitario de tipo Roy,  $w_i = 1$  y  $\mu_i$  impulsa la decisión acerca de la educación universitaria, condicional a  $X$  (Willis y Rosen 1979). En la mayoría de los estudios sociológicos, las características familiares, personales e institucionales dominan (es decir,  $w_i$  es mucho menor que 1), de manera que las covariables observadas  $X$  principalmente determinan la regla de decisión mientras el componente de autoselección cumple un papel secundario o a veces es ignorado (es decir,  $w_i\mu_i = 0$ ).

No puede calcularse la ecuación 4 porque no está identificada. Como estrategia de investigación, apelamos al supuesto de ignorabilidad y de este modo desestimamos el componente no observado de autoselección ( $\mu_i$ ), como un primer paso en el análisis de los datos. Simplificamos aún más la ecuación en un modelo de *score* de propensión de forma reducida, mal especificado pero estimable.

$$d_i^* = \lambda'X_i + v_i. \quad (5)$$

¿Cómo afecta la mala especificación de la ecuación 5 nuestra capacidad para realizar inferencias acerca de los efectos causales específicos de los *scores* de propensión de la educación universitaria sobre los ingresos? La bibliografía sociológica sugiere que dado que  $w$  debería correlacionar negativamente con el *score* de propensión observado, el alcance de la mala especificación causada por la omisión de  $\mu$  declina con el *score* de propensión observado, es decir, en los niños de familias de estatus alto, la decisión de ir a la universidad está menos dictada por la elección racional y la autoselección que en los niños de familias de estatus bajo. Cuando, sobre la base de características observadas, una persona sin expectativas de ingresar a la universidad lo hace, es porque hay factores poderosos involucrados, uno de los cuales puede ser el incentivo económico.

### MODELOS ESTADÍSTICOS

Para precisar las ideas, adoptamos el enfoque de resultados potenciales para la inferencia causal. Éste se utiliza desde hace tiempo en los diseños experimentales (Neyman 1923) y en la teoría económica (Roy 1951) y se ha extendido y formalizado en los estudios observacionales de estadística (por ejemplo, Holland 1986; Rosenbaum y Rubin 1983, 1984; Rubin 1974), economía (por ejemplo, Heckman 2005; Manski 1995) y sociología (por ejemplo, Morgan y Winship 2007; Sobel 2000; Winship y Morgan 1999). El enfoque explicita los temas relevantes para la identificación y estimación de los efectos causales. Se establece que  $y$  son los ingresos registrados y nuevamente  $d$  es una variable con valor 1 para un individuo que completa la universidad y 0 para la situación complementaria. Nos preguntamos cuáles serían los ingresos del individuo  $i$  si recibiera el tratamiento (es decir, completar su educación universitaria), comparado con no haberlo recibido (es decir, no haber completado la educación universitaria). Como sólo se observa realmente uno de los dos valores de ingresos,  $y_i^1$  o  $y_i^0$ , la inferencia causal es imposible a nivel de los individuos. Siempre se requiere un análisis estadístico a nivel de grupo, sobre la base del supuesto de homogeneidad (Holland 1986).

Para inferir causalidad con datos observacionales, es necesario introducir supuestos no verificables. En esta investigación, primero introducimos el supuesto de ignorabilidad:

$$E(y^0 | \mathbf{X}, d = 1) = E(y^0 | \mathbf{X}, d = 0) \quad (6a)$$

y

$$E(y^1 | \mathbf{X}, d = 0) = E(y^1 | \mathbf{X}, d = 1). \quad (6b)$$

La ecuación 6a supone que los ingresos promedio de los trabajadores con educación universitaria, si no la hubieran completado, deberían ser iguales a los ingresos promedio de los trabajadores sin educación universitaria, de modo condicional a las covariables observadas. Asimismo, la ecuación 6b supone que los ingresos promedio de los trabajadores sin educación universitaria, si hubieran completado la universidad, serían iguales a los ingresos promedio de los trabajadores con educación universitaria, de modo condicional a las covariables observadas.

## MODELOS PARA LOS EFECTOS HETEROGÉNEOS DEL TRATAMIENTO

Cuando los efectos de tratamiento son heterogéneos, puede haber dos tipos de sesgos de selección, tal como lo planteamos anteriormente: sesgo de heterogeneidad del pretratamiento y sesgo de heterogeneidad del efecto del tratamiento. Ambos tipos de sesgo pueden amenazar la validez de la inferencia causal en caso de datos observacionales. Estimadores como los de efectos fijos y de diferencia en diferencias intentan eliminar el sesgo de la heterogeneidad del pretratamiento, pero no el sesgo de heterogeneidad del efecto del tratamiento (Angrist y Krueger 1999).

Si permitimos que el coeficiente de tratamiento de la ecuación 1 sea heterogéneo, podemos, al menos teóricamente, formular en detalle los dos tipos de componentes heterogéneos, la ecuación 1 se convierte en:

$$y_i = \alpha_i + \delta_i d_i + \beta' \mathbf{X}_i + U_i. \quad (7)$$

En esta especificación,  $\alpha_i$  representa la heterogeneidad del pretratamiento, en tanto que  $\delta_i$  representa la heterogeneidad del efecto del tratamiento.<sup>10</sup> Si hay sesgo de heterogeneidad del pretratamiento, la correlación  $\rho(\alpha, d) \neq 0$ . Si hay sesgo de heterogeneidad de efecto del tratamiento, la correlación  $\rho(\delta, d) \neq 0$  (Heckman et al. 2006; Winship y Morgan 1999). El modelo de heterogeneidad a nivel individual no es identificable, ya que  $\alpha_i$  y  $\delta_i$  no pueden separarse de  $U_i$  sin restricciones ulteriores; apelamos el supuesto de ignorabilidad. En la práctica, el condicionamiento a  $\mathbf{X}$ , que es generalmente multidimensional, resulta difícil a causa de la “maldición de la dimensionalidad”. No es frecuente encontrar unidades tratadas y no tratadas con valores idénticos a  $\mathbf{X}$ , si  $\mathbf{X}$  tiene una dimensión elevada. Sin embargo, Rosenbaum y Rubin (1983, 1984) muestran que, dado el supuesto de ignorabilidad, es suficiente condicionar al *score* de propensión como una función de  $\mathbf{X}$ . El *score* de propensión se define como la probabilidad de asignación al grupo de tratamiento (carrera universitaria completa) dadas las covariables  $\mathbf{X}$ :

$$P = p(d_i = 1 | \mathbf{X}) \quad (8)$$

En este estudio evaluamos la heterogeneidad en los efectos del tratamiento descomponiendo  $\delta$  en la ecuación 7 en una función no paramétrica del *score* de propensión y usamos un modelo lineal jerárquico para revelar un patrón de los rendimientos.<sup>11</sup> Sobre la base de los atributos familiares y personales observados, podemos dividir un grupo en subpoblaciones con similares *scores* precalculados de propensión a completar la educación universitaria. Luego, evaluamos si la heterogeneidad de la población en la propensión a completar la educación universitaria se asocia a la heterogeneidad de los rendimientos de la educación universitaria. Específicamente, nos preguntamos si el efecto estimado de la educación universitaria se asocia positiva o negativamente a la propensión estimada de completar los estudios universitarios.

Nuestra estrategia analítica se desarrolla en tres pasos: (1) Estimamos la regresión logística binaria que estima la probabilidad de completar la educación universitaria y derivamos *scores* de propensión para cada individuo de la muestra. Agrupamos a los sujetos en estratos de *scores* de propensión estimada para equilibrar las distribuciones de covariables entre graduados universitarios y graduados no universitarios ( $p < .001$ ). (2) En el nivel 1, estimamos los efectos del tratamiento que sirven específicamente para equilibrar los estratos de *score* de propensión utilizando el método de regresión de mínimos cuadrados ordinarios. (3) En el nivel 2, estudiamos los resultados heterogéneos por estratos de *score* de propensión y resumimos la tendencia en la variación de los efectos utilizando un modelo lineal jerárquico (Xie y Wu 2005). Nuestro enfoque es similar al de correspondencia por *score* de propensión, dado que las diferencias observadas en los entrevistados se caracterizan por *scores* de propensión. Los dos métodos difieren en la construcción de las comparaciones. En un análisis clásico de correspondencia por *score* de propensión, la comparación por estatus de tratamiento se realiza según una base individual y se promedia para una población o subpoblación. En nuestro enfoque, la comparación por estatus de tratamiento se construye para un grupo relativamente homogéneo, sobre la base de *scores* de propensión y se estudia en diferentes grupos con *scores* de propensión similares mediante un modelo lineal jerárquico.<sup>12</sup>

## DATOS, MEDIDAS Y ESTADÍSTICA DEScriptIVA

### DESCRIPCIÓN DE LOS DATOS

Para estudiar los efectos heterogéneos del tratamiento de la educación sobre los ingresos, utilizamos dos grandes paneles de conjuntos de datos que contienen información amplia sobre el contexto social, las capacidades y las experiencias educativas de los entrevistados: la Encuesta Longitudinal Nacional de Jóvenes de 1979 (NLSY)<sup>13</sup> y el Estudio Longitudinal de Wisconsin (WLS).<sup>14</sup> Ambas muestras están basadas en cohortes. La ventaja de las encuestas longitudinales de cohorte única es que permiten controlar la potencial confusión entre efectos de pertenecer a una misma cohorte y los efectos de la experiencia. La NLSY es una muestra representativa a nivel nacional de 12.686 entrevistados que tenían entre 14 y 22 años cuando se los encuestó por primera vez en 1979. Esos individuos fueron entrevistados anualmente hasta 1994 y después, cada dos años. Restringimos nuestra muestra a entrevistados que tenían entre 14 y 17 años en la encuesta de base en 1979 (N=5.581), no habían finalizado todavía la escuela secundaria cuando se les pidió que completaran el Examen de Aptitud Vocacional para las Fuerzas Armadas (ASVAB) (N= 3.885), habían completado al menos el 12º grado en 1990 (N= 3.034) y no había datos faltantes en el conjunto de covariables utilizadas en nuestro análisis (N= 2.474). Establecimos estas restricciones a la muestra para estudiar una cohorte con poca variación etaria, para asegurarnos de que todas las medidas que utilizábamos fueran previas a la educación universitaria y para comparar graduados universitarios con entrevistados que habían completado al menos la educación secundaria. Evaluamos los efectos de haber completado la educación universitaria sobre los ingresos para entrevistados entre 29 y 32 años (en 1994), entre 33 y 36 años (en 1998) y entre 37 y 40 años (en 2002), es decir, desde los años iniciales hasta los intermedios de la carrera laboral.

El WLS es un estudio de panel regional, basado en una muestra aleatoria de 10.317 hombres y mujeres graduados de las escuelas secundarias de Wisconsin en 1957. La investigación muestra que para los procesos de logros socioeconómicos, los patrones hallados en el WLS son los mismos que los encontrados en las muestras probabilísticas nacionales (Sheridan 2001). Restringimos nuestra muestra a los entrevistados que no tenían ningún dato faltante en el conjunto de covariables utilizadas en nuestro análisis (N=7.905).<sup>15</sup> La replicación del análisis a través de estas fuentes de datos para dos cohortes diferentes nos permite verificar la robustez de los resultados centrales. Las dos fuentes de datos son también complementarias en sus fortalezas y debilidades relativas. En tanto que la NLSY ofrece una representación a nivel nacional, el WLS consta de una muestra mucho más grande de entrevistados

relativamente homogéneos con muchas covariables medidas adecuadamente antes de la educación universitaria, entre las cuales se incluye una medida particularmente confiable de capacidad cognitiva.

## **MEDICIÓN DE LAS VARIABLES**

La tabla 1 presenta la lista de las variables previas a la educación universitaria que utilizamos para construir los estratos del *score* de propensión para nuestras dos fuentes de datos. La mayoría de estas medidas tienen un lugar destacado en los estudios de sociología de la educación y de logros ocupacionales y su medición es sencilla. No obstante, existen pequeñas diferencias entre las fuentes de datos en la medición de estas variables. Los ingresos de los padres se miden como ingresos familiares totales netos en dólares de 1979 en la NLSY; el WLS utiliza los ingresos de los padres en dólares de 1957. En la NLSY, la categoría “Residencia/proximidad a la Universidad” indica si un entrevistado vivía en una SMSA (Área Estadística Metropolitana Estándar) en 1979; en el WLS, si el entrevistado vivía en una escuela secundaria vivía dentro de las 15 millas de una universidad. En la NLSY, “Programa de preparación para la universidad” (“*College-prep*”) indica si un estudiante se había inscrito en un programa de preparación para la universidad; en el WLS, si había completado los requisitos para la UW-Madison. La medición de la capacidad mental también difiere en ambas fuentes de datos. En 1980, al 94% de los entrevistados de la NLSY se les tomó la prueba ASVAB, una batería de 10 pruebas de inteligencia que mide el conocimiento y las habilidades en campos como matemáticas y lenguaje. Primero residualizamos cada una de las pruebas ASVAB según la edad en el momento de realización de la prueba, separando por raza y género, con residuos estandarizados para tener una media de cero y varianza uno. Luego combinamos los ítems (con cargas iguales cuya suma fuera uno) en una escala compuesta (Cronbach’s  $\alpha = .92$ ) (Cawley et al. 1997). En el WLS, utilizamos la Prueba de Henmon-Nelson de 1957 de índices de Capacidad Mental. Usamos la paga por hora como variable de resultado final en forma logarítmica.<sup>16</sup> En la NLSY, nuestro resultado consiste en las pagas por hora y los salarios registrados para los entrevistados que se encontraban en la etapa final de la década de los 20 años y en la etapa inicial de la década de los 40 años (en 1994, 1998 y 2002). En el WLS, nuestro resultado consiste en los ingresos anuales registrados a los 35 años (en 1975) y la paga por hora registrada a los 53 años (en 1993).<sup>17</sup> Sumamos una pequeña constante positiva (\$.50) antes de calcular logaritmos. Se eliminaron los trabajadores desempleados.<sup>18</sup>

## **ESTADÍSTICAS DESCRIPTIVAS**

La mayor probabilidad de completar los estudios universitarios es uno de los mecanismos causales más importantes que permiten comprender la ventaja asociada al origen socioeconómico alto; éste es un descubrimiento clave del estudio clásico de Blau y Duncan (1967). Como muestra la tabla 1, es más verosímil que los graduados universitarios, comparados con los individuos que no lo son, provengan de familias con ingresos elevados, padres con alto nivel educativo, estructura familiar intacta y pocos hermanos. Asimismo, el éxito académico en la escuela secundaria, una buena capacidad cognitiva y el estímulo de los padres y maestros para que concurran a la universidad, así como el hecho de que los amigos planifiquen concurrir a la universidad son predictores de la educación universitaria. Estas estadísticas sugieren que muchos factores no económicos son relevantes para que los jóvenes obtengan logros educativos. Finalmente, por razones multifacéticas (Kao y Thompson 2003), la posibilidad de completar los estudios universitarios varía según la raza y el origen hispano: los blancos y asiáticos tienen más posibilidades de completarlos que los negros y los hispanos.

## **PRINCIPALES ANÁLISIS Y RESULTADOS**

### **RENDIMIENTOS ECONÓMICOS DE LA EDUCACIÓN UNIVERSITARIA SEGÚN EL SUPUESTO DE HOMOGENEIDAD**

La tabla 2 proporciona los efectos estimados de la finalización de la educación universitaria sobre los ingresos por sexo, a través del análisis de regresión conforme al supuesto efecto homogéneo, dado el conjunto completo de covariables descriptas anteriormente.<sup>19</sup> En el caso de los hombres con empleo de la NLSY, la finalización de los estudios universitarios tiene un efecto positivo muy significativo sobre la paga por hora registrada, que se incrementa regularmente a lo largo del tiempo, desde un 20% de ventaja para los hombres de entre fines de la década de los 20 y comienzos de la de los 30 años, hasta un 51% de ventaja para aquellos que estaban entre fines de los 30 y comienzos de los 40 años. Estos resultados son consistentes con el modelo del capital humano. Dada la conocida tendencia al crecimiento de los rendimientos de la educación universitaria a largo del tiempo, no es sorprendente que el resultado de la

finalización de la educación universitaria sea menor en las primeras cohortes del WLS. Aun así, los resultados para los hombres del WLS indican rendimientos asociados a un título universitario significativos y crecientes a lo largo de la vida.

En cuanto a las mujeres con empleo de la NLSY, los resultados revelan un efecto amplio y significativo de la finalización de la educación universitaria en la franja etaria que va de fines de la década de los 20 años a comienzos de la de los 30 años, un efecto menor a mediados de los 30 años, comparado con las que están a comienzos de los 30 y, luego, un efecto comparativamente mayor en la franja etaria comprendida entre fines de los 30 y comienzos de los 40.<sup>20</sup> Las diferencias en los patrones a lo largo de la vida entre hombres y mujeres pueden reflejar las influencias de los tradicionales papeles de los géneros en la familia y el correspondiente vínculo intermitente de las mujeres (en relación con los hombres) con la fuerza de trabajo, especialmente durante los años reproductivos (Becker 1991; Bianchi 1995; Mincer y Polachek 1974). De manera que el patrón de los efectos a lo largo de la vida de las mujeres puede reflejar la decisión de algunas mujeres de quedar fuera de la fuerza de trabajo o de tener menores rendimientos de la educación universitaria durante los años reproductivos. En el WLS, el efecto del título universitario también declina durante la vida en el caso de las mujeres, pero el efecto cuando están en la mitad de la tercera década es mayor para las encuestadas del WLS que para las de la NLSY. Creemos que la selección de la muestra puede explicar estos resultados un tanto peculiares. La participación de las mujeres en la fuerza laboral era mucho menor en la cohorte del WLS (el 57% de las mujeres del WLS tenía empleo a los 35 años, en comparación con el 76% de las mujeres de la NLSY en la mitad de la tercera década), así que esta cohorte fue más selectiva en relación con los ingresos que la de la NLSY.

#### GENERACIÓN DE LOS ESTRATOS DE SCORE DE PROPENSIÓN

Nuestro siguiente objetivo consistió en estudiar los efectos heterogéneos de la finalización de la educación universitaria por estratos de *scores* de propensión. Estimamos regresiones logísticas binarias que predicen la probabilidad de completar la educación universitaria mediante las covariables descritas en la tabla 1 para cada fuente de datos, separadamente por sexo, y derivamos *scores* de propensión estimados para cada individuo (Becker e Ichino 2002). La tabla A1, que figura en el Apéndice, da cuenta de los resultados de las regresiones logísticas. Luego generamos estratos de *score* de propensión balanceados; el balance se satisface cuando dentro de cada intervalo del *score* de propensión, el *score* de propensión promedio y las medias de cada covariable no difieren significativamente entre graduados y no graduados universitarios. Restringimos el algoritmo de balance a la zona de apoyo común, es decir, las regiones de *scores* de propensión en las que se observaron tanto unidades tratadas como unidades control.<sup>21</sup> Para demostrar que se ha logrado el balance en cada estrato, en la tabla 3 presentamos las medias de las covariables por estrato de *score* de propensión para los hombres de la NLSY.<sup>22</sup> La tabla 3 también aclara las características de un individuo típico dentro de cada estrato. Por ejemplo, una persona típica del estrato 1 tiene padres que abandonaron el colegio secundario, tres hermanos, baja capacidad cognitiva, amigos que no han planificado concurrir a la universidad y ha iniciado una carrera laboral no académica. En cambio, una persona típica del estrato 5 tiene padres con educación universitaria parcial, un hermano/a, elevada capacidad cognitiva, amigos que han planificado concurrir a la universidad y ha iniciado una carrera académica. La tabla 4 muestra la cantidad de casos en cada estrato, separadamente por logro educativo universitario, género y fuente de datos. Tal como se esperaba, las distribuciones de frecuencia para los individuos con y sin educación universitaria van en direcciones opuestas. En el caso de los individuos con educación universitaria, el recuento de frecuencia aumenta con el *score* de propensión, en tanto que para los individuos sin educación universitaria, el recuento disminuye con el *score* de propensión. Aún así, encontramos superposición dentro de cada estrato: en cada estrato de *score* de propensión existen individuos con  $d = 1$  y otros individuos con  $d = 0$ .<sup>23</sup>

#### RENDIMIENTOS HETEROGÉNEOS DE LA EDUCACIÓN UNIVERSITARIA

Las figuras 2 a 5 presentan los resultados principales de nuestro estudio. En primer término, estimamos los efectos específicos del tratamiento de los estratos de *score* de propensión y luego detectamos el patrón de los efectos por *score* de propensión con un modelo lineal jerárquico. Los puntos en las figuras 2 a 5 representan las estimaciones de efectos específicos por estratos de la finalización de los estudios universitarios sobre los ingresos registrados. Los trazados lineales y las pendientes de nivel 2 que aparecen en las figuras se basan en los modelos lineales jerárquicos (es decir, modelos de mínimos cuadrados de nivel 2 ponderados por varianza, estimados por los efectos de la educación universitaria de nivel 1 específicos por estratos de *score* de propensión regresionados sobre el rango de propensión del

estrato). En el apéndice, en la tabla A2 (para la NLSY) y en la tabla A3 (para el WLS) se presentan todos los puntos estimados y los valores de  $t$  asociados correspondientes a las figuras 2 a 5.

La figura 2 describe los resultados de los efectos de la educación universitaria sobre los ingresos de los hombres de la NLSY de 29 a 32 años (en 1994), de 33 a 36 años (en 1998) y de 37 a 40 años (en 2002). Las pendientes lineales descendentes ilustran la tendencia descendente de los efectos con rango de estrato de propensión y en todos los períodos observados. Por ejemplo, en el caso de los hombres de un rango etario entre fines de los 20 y comienzos de los 30 años, un cambio de unidad en el rango del estrato se asocia a una reducción del 5% en el efecto del tratamiento: el efecto estimado de la finalización de la educación universitaria sobre los ingresos en el estrato 1 es de aproximadamente el 30%, en tanto que el efecto estimado en el estrato 5 es de aproximadamente el 10%. Esto significa, por ejemplo, que un individuo cuyos padres abandonaron el colegio secundario y que, a su vez, tiene una capacidad cognitiva baja, se beneficia más por haber completado la educación universitaria, en una magnitud estimada del 20%, que un individuo cuyos padres asistieron a la universidad y que tiene elevada capacidad cognitiva. También encontramos evidencia que sugiere una tendencia descendente en los efectos de la educación universitaria sobre los ingresos en los hombres que se encuentran a mediados y fines de la tercera década y comienzos de la cuarta década. Por ejemplo, el 5% de reducción en el efecto del tratamiento por rango de estrato entre los que se encuentran entre fines de la tercera década y comienzos de la cuarta década da como resultado una diferencia estimada del 20% entre los estratos más bajos y los más altos o entre los concurrentes a la universidad menos y más favorecidos. Así, las pendientes de nivel 2 sustentan la hipótesis de selección negativa en cada estadio de la vida observado. Como se esperaba, la finalización de la educación universitaria se asocia a rendimientos económicos crecientes durante el transcurso de la vida y esto es cierto en todos los estratos de *scores* de propensión.

Los resultados para las mujeres de la NLSY, que se muestran en la figura 3, son similares a los de los hombres, lo que parece indicar selección negativa en cada etapa de la vida observada. Por ejemplo, en las mujeres que se encuentran entre fines de la tercera y comienzos de la cuarta década, un cambio de unidad en el rango de estrato se asocia a un 4% de reducción en el efecto del tratamiento. El efecto estimado por la finalización de la educación universitaria sobre los ingresos es de aproximadamente el 40% para las mujeres del estrato 1, que provienen de contextos socioeconómicos desfavorecidos, versus aproximadamente el 25% para las mujeres del estrato 5, que provienen de contextos socioeconómicos favorecidos. Sin embargo, a diferencia de lo que sucede con los hombres, entre las mujeres observamos rendimientos oscilantes de la educación universitaria durante la vida, como discutimos en los resultados según el supuesto de homogeneidad. Una vez más, las diferencias en los rendimientos económicos de los ingresos durante la vida entre hombres y mujeres posiblemente reflejen el vínculo intermitente de las mujeres con la fuerza laboral durante los años reproductivos. Estos procesos familiares podrían afectar los estratos de *scores* de propensión de modo diferente en las mujeres.

Las figuras 4 y 5 presentan los resultados para los hombres y mujeres del WLS, respectivamente.<sup>24</sup> La figura 4 describe los ingresos de los hombres del WLS a los 35 años (en 1975) y a los 53 años (en 1993). Las pendientes de nivel 2 en el WLS indican una reducción de menos del 1% por rango de estrato a los 35 años (la pendiente de nivel 2 más plana observada) y una reducción del 2% por rango de estrato a los 53 años. En el WLS hay nueve estratos; en el NLSY, cinco. Un 2% de reducción por rango de estrato a los 53 años significa un 18% de disminución en los rendimientos de la educación universitaria sobre los ingresos para el estrato más alto, comparado con el estrato más bajo. Aunque generalmente observamos rendimientos menores de la educación universitaria en la cohorte más joven, los resultados del WLS son consistentes con los de la NLSY y brindan sustento para la selección negativa. En cada período de la vida observado, las pendientes de nivel 2 muestran que el beneficio de haber completado la educación universitaria es mayor entre los hombres que tenían *menos* probabilidades de completar dicha educación (figuras 2 y 4).

Las mujeres del WLS muestran una pendiente descendente mucho más empinada a los 35 años, en comparación con las mujeres de la NLSY que se encuentran en la mitad de la tercera década, como resultado de rendimientos muy altos entre las mujeres del estrato 1. Las mujeres de contextos sociales desfavorecidos que obtuvieron títulos universitarios a comienzos de la década de 1960 fueron, posiblemente, especialmente selectivas, lo que generó rendimientos inusualmente altos. Posiblemente, estas mujeres fueron menos propensas a cumplir con los papeles familiares tradicionales, debido a que tenían menos probabilidades de estar casadas con hombres con recursos económicos suficientes para la especialización de funciones dentro de las familias, si se las compara con sus pares provenientes de contextos favorecidos.

Las figuras 2 a 5 demuestran el funcionamiento de un mecanismo de selección sistemático: los individuos con baja propensión a completar la educación universitaria (es decir, individuos con orígenes sociales más desfavorecidos y con menores capacidades cognitivas y logros) que realmente finalizan los estudios universitarios son los que obtienen mayores beneficios de este hecho. Las tablas A2 y A3, que figuran en el apéndice, muestran que la brecha salarial entre los grupos de tratamiento y control (coeficiente de nivel 1) es estadísticamente significativa dentro de varios (pero de ninguna manera todos) los estratos de *score* de propensión. Esta faceta de nuestros resultados es consistente con la hipótesis de selección negativa: 8 de cada 10 efectos estimados en el estrato 1 son estadísticamente significativos, en tanto que sólo 3 de cada 10 efectos estimados son significativos en el último estrato. Aún así, existen pocos coeficientes de pendiente de nivel 2 estadísticamente significativos. Además, los coeficientes de pendiente de nivel 2 están basados en muy pocos puntos de datos. No existe “población” de estratos de *scores* de propensión que pueda considerarse verdaderas unidades de análisis de nivel 2. Ajustamos el modelo lineal jerárquico para proporcionar un resumen general de un grado de libertad sobre la dirección del patrón de heterogeneidad del efecto como una función del *score* de propensión. La dirección es negativa en todos los casos. Cuando menos, podemos decir que los patrones de selección en las figuras 2 a 5 son claramente no positivos. Para cada caso presentado, la evidencia a favor de la hipótesis de selección negativa es sólo una sugerencia. No obstante, el descubrimiento nuclear de un patrón negativo es válido para dos fuentes de datos diferentes con medidas de calidad variable, en todos los períodos de la vida observados para las diferentes cohortes y tanto para los hombres como para las mujeres (un total de 10 pendientes negativas diferentes de nivel 2). Más aún, encontramos el mismo patrón en otra base de datos longitudinales basado en cohortes de los Estados Unidos (Brand y Xie 2007).<sup>25</sup> De todos modos, nuestros resultados deben considerarse descriptivos e indicadores, no definitivos.

## ANÁLISIS AUXILIARES

Dada la evidencia que sugiere selección negativa, consideramos entonces la cuestión de los mecanismos causales. Es plausible, incluso probable, que sean múltiples los mecanismos que expliquen el patrón observado. En primer término, evaluamos la idea de mecanismos diferenciales de selección por estrato de *score* de propensión con una medida del valor de una educación universitaria entre estudiantes de los últimos cursos de la escuela secundaria en el WLS. “Valor de la educación universitaria” es un promedio ponderado de los puntajes en respuesta a una serie de 18 afirmaciones relacionadas con las percepciones del valor de concurrir a la universidad (Amer 1964). Los ítems con ponderación más elevada son, por ejemplo: “Preferiría empezar a ganar dinero rápido y aprender en el trabajo”; “aprender en el trabajo es más práctico que la mayor parte de lo que se aprende en la escuela” y “para mí, ir a la universidad sería perder el tiempo”. (Véase el Memorandum 129 del WLS para mayores detalles sobre los ítems y la construcción de las variables.) En la tabla 5 examinamos valoraciones de educación universitaria por estrato de *score* de propensión y educación entre hombres del WLS. Encontramos un diferencial importante entre graduados y no graduados universitarios en los estratos de baja propensión: los primeros otorgan más valor a la educación universitaria que los segundos. La brecha disminuye gradualmente en todos los estratos del *score* de propensión hasta que prácticamente no existe diferencia entre los pertenecientes al estrato de propensión más alta. El valor inusualmente alto que los jóvenes desfavorecidos que concurren a la universidad realmente le otorgan a la educación universitaria contrasta marcadamente con el valor uniformemente alto (es decir, indiferenciado por estado real de finalización de la educación universitaria) que le otorgan los jóvenes de contextos favorecidos. Como planteamos en la hipótesis, este resultado parece indicar que los mecanismos que llevan al logro de la educación universitaria difieren según el contexto social.

Dado que el valor de la variable educación universitaria comprende tanto los incentivos económicos como los no económicos, utilizamos una variable adicional para determinar si los asistentes a la universidad de baja propensión estaban más determinados por cuestiones económicas que los de alta propensión, para quienes la concurrencia a la universidad constituye una expectativa cultural. Dado que el campo de estudio afecta los ingresos (Thomas y Zhang 2005), estudiamos las asignaturas universitarias específicas de estrato elegidas por los hombres con educación universitaria del WLS. La tabla 6 muestra que es más probable que los estudiantes con baja propensión se concentren en los negocios y la educación -asignaturas con rendimientos económicos inmediatos- y que los estudiantes con alta propensión estudien asignaturas de ciencias y humanidades, temas que requieren intereses académicos especiales y que, posiblemente, estén menos motivados por réditos económicos inmediatos.

Según nuestros resultados, la selección negativa plantea la siguiente pregunta: ¿por qué algunos estudios previos parecen indicar fundamento empírico para la selección positiva? Sospechamos que una explicación está vinculada a la elección, o disponibilidad, de las covariables del análisis. El fundamento empírico para la selección positiva a veces se basa en modelos que omiten variables clave como la capacidad cognitiva, el rendimiento académico en la escuela secundaria y el estímulo de los padres y maestros. Omitir estas importantes variables de confusión puede introducir una distorsión del patrón observado de selección negativa a positiva. Para evaluar esta posibilidad, decidimos actuar como si no pudiéramos acceder al conjunto completo de covariables que teníamos a nuestra disposición y las restringimos en el WLS a un conjunto comparable al utilizado por Carneiro, Heckman y Vytlačil (2001).<sup>27, 28</sup> Omitimos en nuestros modelos las variables referidas a la capacidad cognitiva, los estudios académicos y la religión, así como también las sociopsicológicas, factores muy significativos en el modelo expandido que mostramos en la tabla A1 (en el apéndice). Como se esperaba, encontramos diferencias medias importantes en estas covariables entre los encuestados que recibieron educación universitaria y los que no la recibieron dentro de cada estrato de *score* de propensión.

La figura 6 muestra los resultados revisados para los hombres de 35 años del WLS, con la omisión de las variables antes mencionadas. Cuando restringimos los modelos a un conjunto más limitado de covariables, encontramos evidencia de selección positiva. Esta figura debería compararse con la figura 4, dado que el análisis es paralelo para la misma muestra y la diferencia reside en la especificación de covariables para el modelo de propensión. La omisión de estas variables no sólo cambia el tamaño total del efecto de la educación universitaria (como en el caso en que se supone que los efectos son homogéneos), sino que también cambia la dirección de la asociación entre propensión al tratamiento y efectos del tratamiento. Con el conjunto completo de covariables a nuestra disposición, observamos un patrón moderado de selección negativa; cuando restringimos las covariables a un conjunto más limitado, observamos selección positiva. No podemos atribuir el cambio de dirección a la omisión de una covariable única; un análisis de descomposición detallado, elaborado a partir de Xie y Shauman (1988), muestra que no es posible considerar responsable a una covariable única.<sup>29</sup> Más bien, la omisión del conjunto completo de covariables (capacidad y logro académico, aspiraciones y apoyo de los adultos) induce el cambio observado de una pendiente positiva a una negativa. Este cambio es atribuible al hecho de que estas covariables adicionales ejercen mayor poder explicativo sobre la finalización de los estudios universitarios en los estratos de *score* de propensión más altos que en los más bajos. Esto es consistente con nuestra discusión previa de las ecuaciones 3 y 4, en la que afirmamos que los factores no económicos deberían desempeñar un papel más importante en la predicción de la educación universitaria en las personas con probabilidades de finalizar la universidad más que en sus pares, cuyas probabilidades de finalizarla son menores.

## DISCUSIÓN Y CONCLUSIONES

La heterogeneidad en respuesta a un tratamiento común es una norma, no una excepción. Los individuos difieren no sólo en sus atributos contextuales, sino en cómo responden a un tratamiento en particular. Una tarea importante de la investigación sociológica consiste en sintetizar patrones sistemáticos en la variabilidad de las poblaciones, una tradición demográfica de larga data que Xie (2007) atribuye a Otis Dudley Duncan. En este artículo, consideramos la heterogeneidad de la población en los rendimientos económicos de la educación, al examinar los efectos de la finalización de la educación universitaria por estrato de *score* de propensión en un modelo lineal jerárquico. En primer término, calculamos los efectos de la educación universitaria para diferentes grupos, sobre la base de la probabilidad de lograr una educación universitaria; luego, examinamos la heterogeneidad sistemática en los efectos específicos de grupo. Nuestro análisis describe si los patrones de heterogeneidad de la población reflejan una selección positiva o negativa, es decir, si los beneficios económicos de la educación universitaria son mayores en personas que tienen las mayores probabilidades de completar la educación universitaria o en personas cuyas probabilidades son las menores. Nuestra evidencia parece indicar selección negativa: los individuos que tienen las mayores probabilidades de obtener beneficios de la educación universitaria son los que tienen las *menores* probabilidades de acceder a ella. Este resultado se encontró tanto en hombres como en mujeres de todas las franjas etarias observadas a lo largo de la vida y en dos cohortes diferentes.

Los patrones empíricos en general son consistentes con el efecto de interacción que planteamos como hipótesis y que presentamos en la figura 1. La demanda cada vez mayor de trabajadores educados, combinada con la demanda cada vez menor de trabajadores con bajo nivel educativo, ha generado un aumento en el diferencial de ingresos entre trabajadores con y sin educación (Farley 1966). Sin embargo, encontramos que este diferencial es especialmente amplio entre individuos con baja propensión a finalizar

la educación universitaria. Por consiguiente, una razón importante para explicar los rendimientos económicos relativamente importantes de los trabajadores de baja propensión con educación universitaria, es que su posición social está marcada por una desventaja sustancial. Si no cuentan con un título universitario, los hombres y mujeres con baja propensión tienen un capital humano, cultural y social limitado y, por consiguiente, perspectivas especialmente limitadas en el mercado laboral. En cambio, aun cuando carezcan de un título universitario, los individuos que provienen de contextos sociales más favorecidos pueden contar con recursos y capacidades superiores. El patrón de selección negativa no emerge porque los asistentes a la universidad de propensión baja ganen más que los de propensión alta, porque esto no sucede. Más bien, el patrón emerge porque los de baja propensión que no concurren a la universidad ganan muy poco.

Comprendimos que el uso del *score* de propensión para identificar efectos de tratamiento tiene limitaciones, porque depende del supuesto de ignorabilidad y no tiene en cuenta la heterogeneidad debida a variables no observadas. La plausibilidad del supuesto de ignorabilidad es específica de cada marco de investigación y depende de la riqueza de las covariables observadas. Apelamos al supuesto de ignorabilidad porque deseamos saber qué pueden revelar los datos observados por sí solos, sabiendo que tenemos un conjunto de covariables ricas. Sin embargo, siempre existen mecanismos o factores causales no observables que persisten. No obstante, centrarse en la heterogeneidad observable de los efectos del tratamiento tiene varios beneficios. Aunque la heterogeneidad del efecto del tratamiento es potencialmente observable, rara vez se estudia en la investigación sociológica empírica. Al centrarnos en la heterogeneidad observable, descubrimos un resultado importante: los individuos más desfavorecidos en relación con el contexto social observado, los logros y las capacidades, son los que más posibilidades tienen de beneficiarse con la educación universitaria. Los análisis auxiliares fortalecen la hipótesis de selección negativa: los individuos que provienen de contextos sociales desfavorecidos y concurren a la universidad pueden usar la educación como un medio de movilidad económica, en tanto que los individuos que provienen de contextos sociales favorecidos, para quienes la educación universitaria constituye una norma cultural, pueden verse menos impulsados por las razones económicas. Más aún, consideramos que en las investigaciones anteriores, el apoyo empírico a favor de la selección positiva puede explicarse por la omisión de algunas variables clave.

Tenemos varios planes de investigación para el futuro. En primer término, en este trabajo nos centramos sólo en la brecha entre los ingresos de los individuos que finalizaron su educación universitaria y los que sólo completaron la escuela secundaria, lo que nos permite tomar prestados conceptos de la bibliografía metodológica sobre inferencia causal. Aun cuando existe una diferencia bien documentada entre ambos grupos en el mercado laboral (Grubb 1993; Kane y Rouse 1995; Mare 1981), claramente resulta una simplificación considerar que la educación es un tratamiento dicotómico. En estudios futuros mediremos la educación superior con mayor precisión, por monto, calidad y asignaturas principales. En segundo término, en este estudio encontramos algunas diferencias interesantes entre los hombres y las mujeres caracterizados por estratos de *score* de propensión durante la vida que, conjeturamos, reflejan las diferencias en la intermitencia de la mujer en el mercado laboral entre los distintos estratos. Las próximas investigaciones analizarán con mayor profundidad los efectos heterogéneos de la educación superior sobre los patrones de formación de la familia en las mujeres.

La creencia difundida en los rendimientos socioeconómicos de la educación superior ha impulsado cambios en las políticas públicas para ampliar las oportunidades educativas para todos los estadounidenses. Si bien los encargados de diseñar políticas suponen la homogeneidad de los rendimientos económicos de la educación, la heterogeneidad potencial de los rendimientos recibe ahora mayor atención, dado que muchos países están experimentando la rápida expansión de la matrícula universitaria. Esto ha llevado a que algunos cuestionen los costos y beneficios relativos de la educación superior para las personas que antes no la recibían. Pero, en presencia de los efectos heterogéneos del tratamiento, no puede invocarse una simple declaración resumida sobre el beneficio de completar la educación universitaria, tanto para los individuos que ya habían accedido a la educación superior como para aquellos que tenían posibilidades de beneficiarse con la expansión de la educación. El beneficio promedio depende de la composición que tenga en un momento determinado el grupo de estudiantes que ha finalizado la universidad. Una interpretación de nuestros resultados es que una educación universitaria puede ser especialmente beneficiosa para los grupos a quienes se han dirigido los esfuerzos de expansión de la educación, es decir, individuos que de otro modo posiblemente no hubieran concurrido a la universidad, sobre la base de sus características observadas.

A causa de nuestro supuesto metodológico simplificador, la conclusión anterior es sólo tentativa. El verdadero patrón de los efectos heterogéneos del tratamiento de la educación universitaria sobre los ingresos según la propensión a finalizar los estudios universitarios sugiere la instauración de un mecanismo de selección no observado: los individuos de contextos sociales desfavorecidos, para los cuales la educación universitaria no es un resultado culturalmente esperado, superan importantes dificultades para concurrir a la universidad y pueden estar especialmente motivados por razones económicas. Así, si la expansión de la educación dio como resultado una cantidad mayor de asistentes a la universidad que de otro modo no hubiesen podido hacerlo, podría disminuir la selectividad no observada debida a la motivación económica. Esto podría equiparar la selectividad no observada a lo largo del espectro de propensión a completar la educación universitaria y cambiar el patrón general negativo que observamos en este estudio por un patrón plano. Por lo tanto, una interpretación alternativa de nuestros resultados es que el patrón observado de selección negativa se debe a la selectividad diferencial, en el que las personas de baja propensión a completar la educación universitaria son más selectivas que las de alta propensión. Si bien no podemos decidir entre estas dos interpretaciones alternativas, hemos encontrado un importante resultado empírico: los individuos que tienen menos posibilidades de lograr una educación universitaria son los que obtienen de ella mayores beneficios.

**Jennie E. Brand** es Profesora Adjunta de Sociología en la University of California-Los Ángeles. También es miembro del cuerpo docente en el California Center for Population Research. Sus investigaciones se centran en la relación entre el contexto socioeconómico, el logro educativo y las condiciones laborales y el logro socioeconómico y el bienestar durante la vida. A la investigación de estos temas de relevancia se agrega un énfasis metodológico en la inferencia causal y en la aplicación e innovación de los modelos estadísticos para datos de panel. Sus actuales proyectos de investigación comprenden la evaluación de los efectos heterogéneos de los programas de asistencia a los estudiantes desfavorecidos y de la educación superior sobre los resultados socioeconómicos, y las consecuencias sociales de los desplazamientos laborales.

**Yu Xie** es Profesor Universitario Distinguido Otis Dudley Duncan de Sociología y Estadística en la University of Michigan. También es Profesor Investigador en el Population Studies Center y Survey Research Center del Institute for Social Research y miembro adjunto del cuerpo docente en el Center for Chinese Studies. Sus principales áreas de interés son la estratificación social, la demografía, los métodos estadísticos, los estudios sobre China y la sociología de la ciencia. Sus trabajos publicados recientemente incluyen: *Women in Science: Career Processes and Outcomes* (Harvard University Press 2003, con Kimberlee Shauman, y *Marriage and Cohabitation* (University of Chicago Press 2007), con Arland Thornton y William Axinn.

## REFERENCIAS

- Altonji, Joseph G. and Thomas A. Dunn. 1996. "The Effects of Family Characteristics on the Return to Education." *The Review of Economics and Statistics* 78:692–704.
- Angrist, Joshua D., Guido W. Imbens, and Donald B. Rubin. 1996. "Identification of Causal Effect Using Instrumental Variables." *Journal of the American Statistical Association* 91(434):444–55.
- Angrist, Joshua D. and Alan B. Krueger. 1999. "Empirical Strategies in Labor Economics." Pp. 1277–1366 in *Handbook of Labor Economics*, Vol. 3A, edited by O. Ashenfelter and D. Card. Amsterdam: Elsevier.
- Armer, J. Michael. 1964. *Community and School Environments and College Plans of Public High School Seniors*. PhD dissertation, University of Wisconsin, Madison, WI.
- Ashenfelter, Orley and Alan B. Krueger. 1994. "Estimates of the Economic Return to Schooling from a New Sample of Twins." *American Economic Review* 84:1157–73.
- Ashenfelter, Orley and Cecilia Rouse. 1998. "Income, Schooling, and Ability: Evidence from a New Sample of Identical Twins." *Quarterly Journal of Economics* 113:253–84.
- Averett, Susan L. and Mark L. Burton. 1996. "College Attendance and the College Wage Premium: Differences by Gender." *Economics of Education Review* 15:37–49.

- Barrow, Lisa and Cecilia E. Rouse. 2005. "Do Returns to Schooling Differ by Race and Ethnicity?" *American Economic Review* 95:83–87.
- Beattie, Irene. 2002. "Are all Adolescent Econometricians Created Equal? Racial, Class, and Gender Differences in College Enrollment." *Sociology of Education* 75:19–43.
- Becker, Gary S. 1964. *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*. New York: Columbia University Press.
- . 1991. *A Treatise on the Family*. Cambridge, MA: Harvard University Press.
- Becker, Sascha and Andrea Ichino. 2002. "Estimation of Average Treatment Effects Based on Propensity Scores." *Stata Journal* 2:358–77.
- Bianchi, Suzanne. 1995. "Changing Economic Roles of Women and Men." Pp. 107–154 in *State of the Union: America in the 1990s; Volume One: Economic Trends*, edited by R. Farley. New York: Russell Sage Foundation.
- Bjorklund, Anders and Robert Moffitt. 1987. "The Estimation of Wage Gains and Welfare Gains in Self-Selection Models." *Review of Economics and Statistics* 69:42–49.
- Blake, Judith. 1981. "Family Size and the Quality of Children." *Demography* 18:421–42.
- Blau, Peter and Otis Dudley Duncan. 1967. *The American Occupational Structure*. New York: John Wiley & Sons.
- Boudon, Raymond. 1974. *Education, Opportunity and Social Inequality*. New York: John Wiley & Sons.
- Bourdieu, Pierre. 1977. "Cultural Reproduction and Social Reproduction." Pp. 487–510 in *Power and Ideology in Education*, edited by J. Karabel and A. H. Halsey. New York: Oxford University Press.
- Bowles, Samuel and Herbert Gintis. 1976. *Schooling in Capitalist America: Educational Reform and the Contradictions of Economic Life*. New York: Basic Books.
- Brand, Jennie E. and Charles N. Halaby. 2006. "Regression and Matching Estimates of the Effects of Elite College Attendance on Educational and Career Achievement." *Social Science Research* 35:749–70.
- Brand, Jennie E. and Yu Xie. 2007. "Who Benefits Most From College? Evidence for Negative Selection in Heterogeneous Economic Returns to Higher Education." California Center for Population Research, University of California-Los Angeles CCPR 033-07.
- Breen, Richard and John H. Goldthorpe. 1997. "Explaining Educational Differentials: Towards a Formal Rational Action Theory." *Rationality and Society* 9:275–305.
- Brewer, Dominic, Eric Eide, and Ronald Ehrenberg. 1999. "Does it Pay to Attend an Elite Private College? Cross-Cohort Evidence on the Effects of College Type on Earnings." *Journal of Human Resources* 34(1):104–123.
- Bryk, Anthony, Valerie Lee, and Peter Holland. 1993. *Catholic Schools and the Common Good*. Cambridge, MA: Harvard University Press.
- Cameron, Stephen V. and James J. Heckman. 2001. "The Dynamics of Educational Attainment for Black, Hispanic, and White Males." *Journal of Political Economy* 109:455–99.
- Card, David. 1995. "Using Geographic Variation in College Proximity to Estimate the Return to Schooling." Pp. 201–222 in *Aspects of Labour Market Behavior: Essays in Honour of John Vanderkamp*, edited by L. Christofides, E. Kenneth Grant, and R. Swidinsky. Toronto, Canada: University of Toronto Press.

- . 1999. “The Causal Effect of Education on Earnings.” Pp. 1801–1863 in *Handbook of Labor Economics*, Vol. 3A, edited by O. Ashenfelter and D. Card. Amsterdam: Elsevier.
- . 2001. “Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems.” *Econometrica* 69:1127–60.
- Carneiro, Pedro, Karsten Hansen, and James J. Heckman. 2003. “Estimating Distributions of Treatment Effects with an Application to the Returns to Schooling and Measurement of the Effects of Uncertainty on College Choice.” *International Economic Review* 44:361–422.
- Carneiro, Pedro and James J. Heckman. 2002. “The Evidence on Credit Constraints in Post-Secondary Schooling.” *Economic Journal* 112:705–34.
- Carneiro, Pedro, James J. Heckman, and Edward Vytlačil. 2001. “Estimating the Return to Education When it Varies Among Individuals.” Department of Economics, University College London. Unpublished manuscript.
- . 2007. “Estimating Marginal and Average Returns to Education.” Department of Economics, University College London. Unpublished manuscript.
- Cawley, John, Karen Conneely, James Heckman, and Edward Vytlačil. 1997. “Cognitive Ability, Wages, and Meritocracy.” Pp. 179–92 in *Intelligence, Genes, and Success: Scientists Respond to the Bell Curve*, edited by B. Devlin, S. E. Feinberg, D. Resnick, and K. Roeder. New York: Springer.
- Coleman, James A. 1988. “Social Capital and the Creation of Human Capital.” *American Journal of Sociology* 94:S95–S120.
- Dale, Stacy B. and Alan B. Krueger. 2002. “Estimating the Payoff to Attending a More Selective College: An Application of Selection on Observables and Unobservables.” *Quarterly Journal of Economics* 117(4):1491–1527.
- DiMaggio, Paul. 1982. “Cultural Capital and School Success: The Impact of Status Culture Participation on the Grades of U.S. High School Students.” *American Sociological Review* 47:189–201.
- DiPrete, Thomas A. and Henriette Engelhardt. 2004. “Estimating Causal Effects with Matching Methods in the Presence and Absence of Bias Cancellation.” *Sociological Methods and Research* 32:501–528.
- Farley, Reynolds. 1996. *The New American Reality: Who We Are, How We Got Here, Where We Are Going*. New York: Russell Sage Foundation.
- Fischer, Claude S. and Michael Hout. 2006. *Century of Difference: How America Changed in the Last One Hundred Years*. New York: Russell Sage Foundation.
- Goldthorpe, John and Michelle Jackson. 2008. “Education-Based Meritocracy: The Barriers to its Realization.” Pp. 93–117 in *Social Class: How Does it Work?* edited by A. Lareau and D. Conley. New York: Russell Sage Foundation.
- Griliches, Zvi. 1977. “Estimating the Returns to Schooling: Some Econometric Problems.” *Econometrica* 45:1–22.
- Grubb, W. Norton. 1993. “The Varied Economic Returns to Post-Secondary Education: New Evidence from the Class of 1972.” *The Journal of Human Resources* 28:365–82.
- Hauser, Robert M. 1973. “Socioeconomic Background and Differential Returns to Education.” Pp. 129–45 in *Does College Matter? Some Evidence on the Impacts of Higher Education*, edited by L. C. Solomon and P. J. Taubman. New York: Academic Press.
- . 2004. “Progress in Schooling.” Pp. 271–318 in *Social Inequality*, edited by K. Neckerman. New York: Russell Sage Foundation.

- Hauser, Robert M. and Thomas N. Daymount. 1977. "Schooling, Ability, and Earnings: Cross-Sectional Findings 8 to 14 Years after High School Graduation." *Sociology of Education* 50:182–206.
- Hauser, Robert M., Shu-Ling Tsai, and William H. Sewell. 1983. "A Model of Stratification with Response Error in Social and Psychological Variables." *Sociology of Education* 56:20–46.
- Heckman, James J. 1978. "Dummy Endogenous Variables in a Simultaneous Equation System." *Econometrica* 46(4):931–59.
- . 2005. "The Scientific Model of Causality." *Sociological Methodology* 35:1–98.
- . 2007. "Schools, Skills and Synapses." Lecture at Peking University, China. Retrieved August 29, 2007 ([http://jenni.uchicago.edu/papers/pku\\_2007/](http://jenni.uchicago.edu/papers/pku_2007/)).
- Heckman, James J. and Bo E. Honoré. 1990. "The Empirical Content of the Roy Model." *Econometrica* 58:1121–49.
- Heckman, James J., Justin L. Tobias, and Edward Vytlačil. 2001. "Four Parameters of Interest in the Evaluation of Social Programs." *Southern Economic Journal* 68:210–223.
- Heckman, James, Sergio Urzua, and Edward Vytlačil. 2006. "Understanding Instrumental Variables in Models with Essential Heterogeneity." *The Review of Economics and Statistics* 88:389–432.
- Hill, C. Russell and Frank P. Stafford. 1974. "Allocation of Time to Preschool Children and Educational Opportunity." *Journal of Human Resources* 9:323–41.
- Hoffer, Thomas, Andrew Greeley, and James Coleman. 1985. "Achievement Growth in Public and Catholic Schools." *Sociology of Education* 58(2):74–97.
- Holland, Paul W. 1986. "Statistics and Causal Inference" (with discussion). *Journal of American Statistical Association* 81:945–70.
- Hout, Michael. 1984. "Status, Autonomy, and Training in Occupational Mobility." *American Journal of Sociology* 89:379–409.
- . 1988. "More Universalism, Less Structural Mobility: The American Occupational Structure in the 1980s." *American Journal of Sociology* 93:1358–1400.
- Hout, Michael and Thomas A. DiPrete. 2006. "What We Have Learned: RC28's Contributions to Knowledge about Social Stratification." *Research in Social Stratification and Mobility* 24:1–20.
- Imbens, Guido W. and Joshua Angrist. 1994. "Identification and Estimation of Local Average Treatment Effects." *Econometrica* 62:467–76.
- Jann, Ben, Jennie E. Brand, and Yu Xie. 2008. –hte– Stata module to perform heterogeneous treatment effect analysis, available from (<http://ideas.repec.org/>).
- Jencks, Christopher, Marshall Smith, Henry Acland, Mary Jo Bane, David Cohen, Herbert Gintis, Barbara Heyns, and Stephan Michelson. 1972. *Inequality: A Reassessment of the Effect of Family and Schooling in America*. New York: Basic Books.
- Kane, Thomas and Cecilia Rouse. 1995. "Labor-Market Returns to Two- and Four-Year College." *American Economic Review* 85(3):600–614.
- Kao, Grace and Jennifer S. Thompson. 2003. "Racial and Ethnic Stratification in Educational Achievement and Attainment." *Annual Review of Sociology* 29:417–42.
- Lareau, Annette. 2003. *Unequal Childhoods*. Princeton, NJ: Princeton University Press.

- Lucas, Samuel. 2001. "Effectively Maintained Inequality: Education Transitions, Track Mobility, and Social Background Effects." *American Journal of Sociology* 106:1642–90.
- MacLeod, Jay. 1989. *Ain't No Makin' It: Aspirations and Attainment in a Low-Income Neighborhood*. Boulder, CO: Westview Press.
- Manski, Charles F. 1990. "Adolescent Econometricians: How Do Youth Infer the Returns to Schooling?" Pp. 43–60 in *Studies of Supply and Demand in Higher Education*, edited by C. T. Clotfetter and M. Rothschild. Chicago, IL: University of Chicago Press.
- . 1995. *Identification Problems in the Social Sciences*. Boston, MA: Harvard University Press.
- Manski, Charles F. and David A. Wise. 1983. *College Choice in America*. Cambridge, MA: Harvard University Press.
- Mare, Robert D. 1981. "Change and Stability in Educational Stratification." *American Sociological Review* 46:72–87.
- McLanahan, Sara and Gary Sandefur. 1994. *Growing up with a Single Parent: What Hurts, What Helps*. Cambridge, MA: Harvard University Press.
- Mincer, Jacob. 1974. *Schooling, Experience, and Earnings*. New York: Columbia University Press.
- Mincer, Jacob and Solomon Polachek. 1974. "Family Investments in Human Capital: Earnings of Women." *Journal of Political Economy* 82:S76–S108.
- Morgan, Stephen. 2001. "Counterfactuals, Causal Effect Heterogeneity, and the Catholic School Effect on Learning." *Sociology of Education* 74(4):341–74.
- . 2005. *On the Edge of Commitment: Educational Attainment and Race in the United States*. Stanford, CA: Stanford University Press.
- Morgan, Stephen L. and Jennifer J. Todd. 2008. "A Diagnostic Routine for the Detection of Consequential Heterogeneity of Causal Effects." *Sociological Methodology* 38:231–81.
- Morgan, Stephen and Christopher Winship. 2007. *Counterfactuals and Causal Inference: Methods and Principles for Social Research*. Cambridge, UK: Cambridge University Press.
- Neyman, J. 1923. "On the Application of Probability Theory to Agricultural Experiments: Essay on Principles, Section 9." *Statistical Science* 5(4):465–80.
- Olneck, Michael. 1979. "The Effects of Education." Pp. 159–60 in *Who Gets Ahead?* edited by C. Jencks. New York: Basic Books.
- Orfield, Gary, S. E. Eaton, and E. R. Jones. 1996. *Dismantling Desegregation: The Quiet Reversal of Brown v. Board of Education*. New York: New Press.
- Perna, Laura. 2005. "The Benefits of Higher Education: Sex, Racial/Ethnic, and Socioeconomic Group Differences." *Review of Higher Education* 29(1):23–52.
- Raftery, Adrian and Michael Hout. 1993. "Maximally Maintained Inequality: Expansion, Reform, and Opportunity in Irish Education, 1921–75." *Sociology of Education* 66:41–62.
- Rosenbaum, James. 2001. *Beyond College for All*. New York: Russell Sage Foundation Press.
- Rosenbaum, Paul R. and Donald B. Rubin. 1983. "The Central Role of the Propensity Score in Observational Studies for Causal Effects." *Biometrika* 70:41–55.
- . 1984. "Reducing Bias in Observational Studies Using Subclassification on the Propensity Score." *Journal of the American Statistical Association* 79:516–24.

- Rouse, Cecilia. 2004. "Low-Income Students and College Attendance: An Exploration of Income Expectations." *Social Science Quarterly* 85(5):1299–1317.
- Roy, Andrew D. 1951. "Some Thoughts on the Distribution of Earnings." *Oxford Economic Paper* 3:135–46.
- Rubin, Donald B. 1974. "Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies." *Journal of Educational Psychology* 66:688–701.
- . 1997. "Estimating Causal Effects from Large Data Sets Using Propensity Scores." *Annals of Internal Medicine* 127:757–63.
- Sewell, William, Archibald O. Haller, and G. W. Ohlendorf. 1970. "Educational and Early Occupational Status Attainment Process: Replication and Revision." *American Sociological Review* 35:1014–1027.
- Sewell, William H., Archibald O. Haller, and Alejandro Portes. 1969. "The Educational and Early Occupational Attainment Process." *American Sociological Review* 34:82–92.
- Sewell, William H. and Robert M. Hauser. 1975. *Education, Occupation, and Earnings: Achievement in the Early Career*. New York: Academic Press.
- Sheridan, Jennifer. 2001. *Occupational Attainment across the Life Course: Sources of Stability and Change in Three Occupational Characteristics*. PhD Dissertation, Department of Sociology, University of Wisconsin, Madison, WI.
- Smith, Herbert L. and Brian Powell. 1990. "Great Expectations: Variations in Income Expectations among College Seniors." *Sociology of Education* 63:194–207.
- Sobel, Michael E. 2000. "Causal Inference in the Social Sciences." *Journal of the American Statistical Association* 95:647–51.
- Thomas, Scott. 2003. "Longer-Term Economic Effects of College Selectivity and Control." *Research in Higher Education* 44(3):263–99.
- Thomas, Scott and Liang Zhang. 2005. "Post-baccalaureate Wage Growth within Four Years of Graduation: The Effects of College Quality and College Major." *Research in Higher Education* 46(4):437–59.
- Train, Kenneth. 2003. *Discrete Choice Methods with Simulation*. Cambridge, UK: Cambridge University Press.
- Tsai, Shu-Ling and Yu Xie. 2008. "Changes in Earnings Returns to Higher Education in Taiwan since the 1990s." *Population Review* 47:1–20.
- U.S. Census Bureau. 2007. "Earnings Gap Highlighted by Census Bureau Data on Educational Attainment." *U.S. Census Bureau News*, March 15, 2007, Press Release.
- Welch, Finis. 1973. "Black-White Differences in Returns to Schooling." *American Economic Review* 63:893–907.
- Willis, Robert J. and Sherwin Rosen. 1979. "Education and Self-Selection." *Journal of Political Economy* 87:S7–36.
- Winship, Christopher and Stephen L. Morgan. 1999. "The Estimation of Causal Effects from Observational Data." *Annual Review of Sociology* 25:659–707.
- Xie, Yu. 2007. "Otis Dudley Duncan's Legacy: The Demographic Approach to Quantitative Reasoning in Social Science." *Research in Social Stratification and Mobility* 25(2):141–56.

Xie, Yu and Kimberlee A. Shauman. 1998. "Sex Differences in Research Productivity Revisited: New Evidence about an Old Puzzle." *American Sociological Review* 63:847–70.

Xie, Yu and Xiaogang Wu. 2005. "Market Premium, Social Process, and Statisticism." *American Sociological Review* 70:865–70.

Zhang, Liang. 2005. *Does Quality Matter? Benefits of Attending a High-Cost, Prestigious College*. New York: Rutledge.

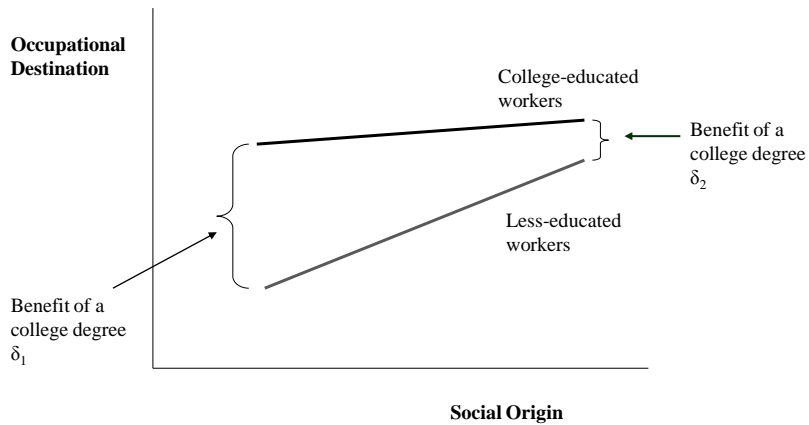


Figura 1. Modelo hipotético: origen, educación y resultados finales

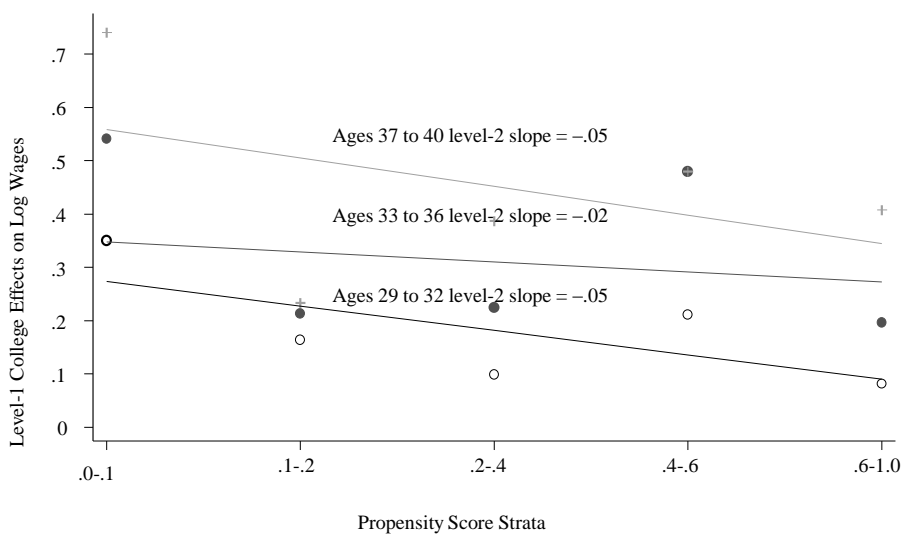


Figura 2. Modelo lineal jerárquico de los rendimientos económicos de la educación universitaria; hombres de la NLSY

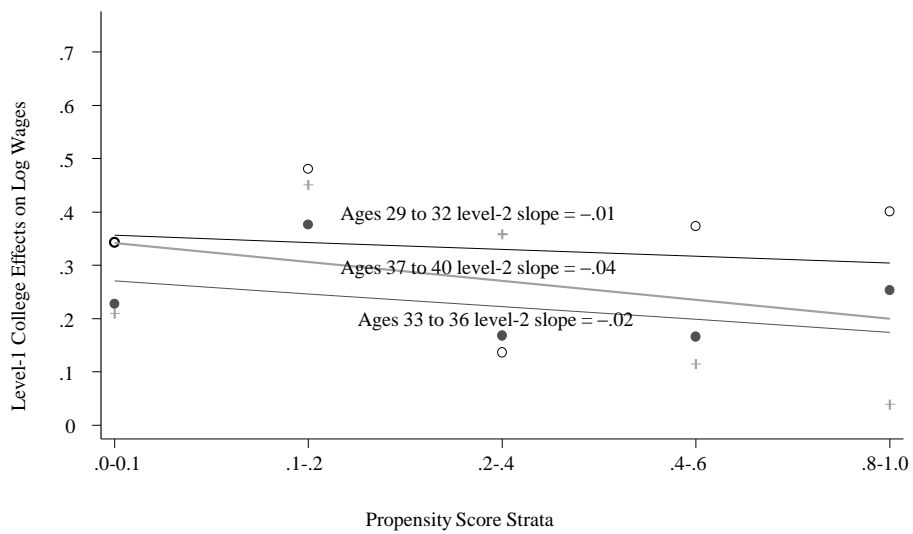


Figura 3. Modelo lineal jerárquico de los rendimientos económicos de la educación universitaria; mujeres de la NLSY

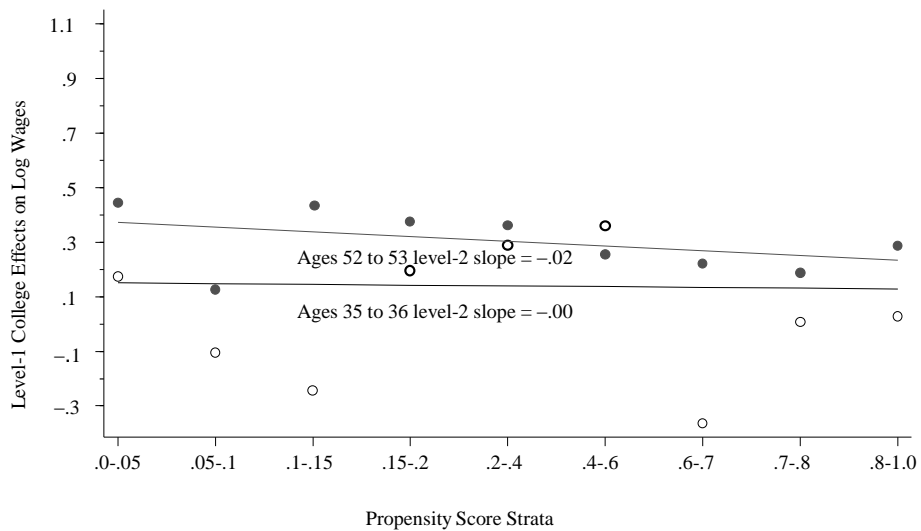


Figura 4. Modelo lineal jerárquico de los rendimientos económicos de la educación universitaria; hombres del WLS

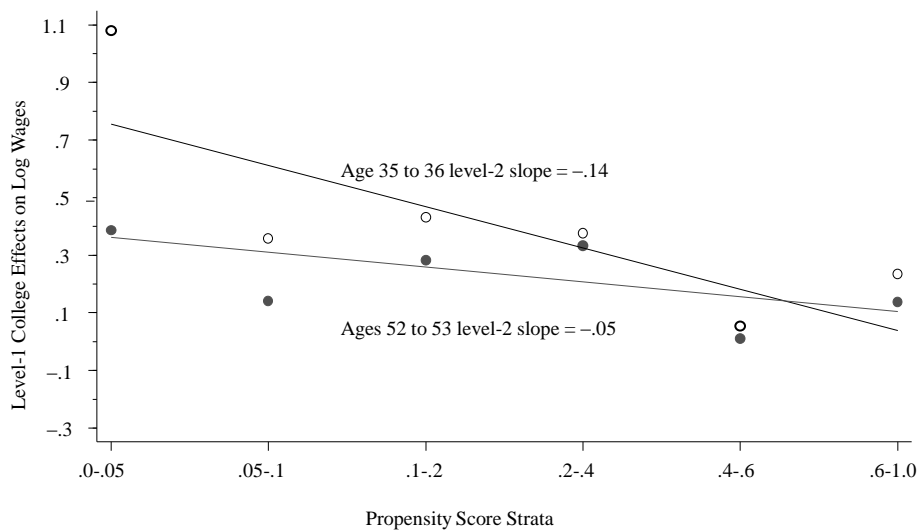


Figura 5. Modelo lineal jerárquico de los rendimientos económicos de la educación universitaria; mujeres del WLS

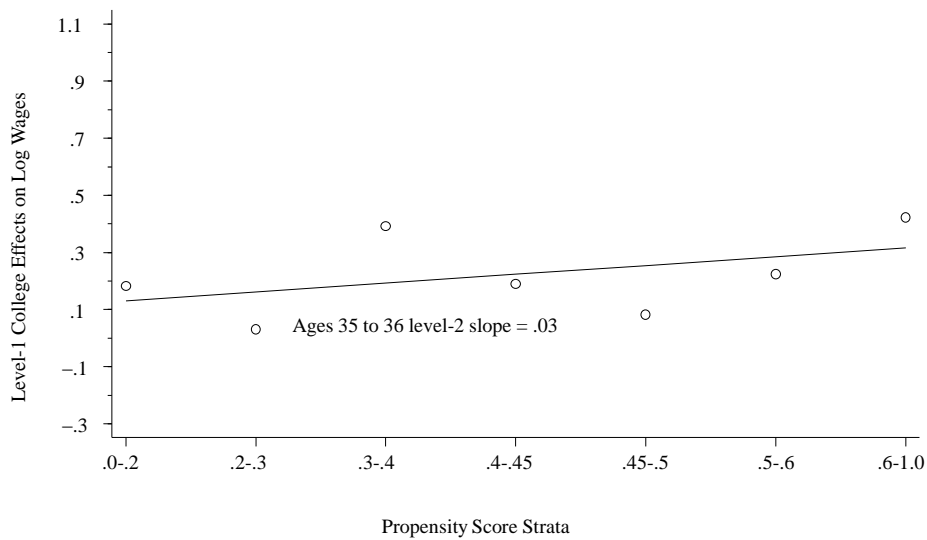


Figura 6. Modelo lineal jerárquico de los rendimientos económicos de la educación universitaria; pequeño conjunto de covariables-hombres del WLS

La correspondencia deberá dirigirse a Jennie E. Brand, Department of Sociology, University of California-Los Angeles, 264 Haines Hall, Los Angeles, CA 90095-1551 (brand@soc.ucla.edu). Esta investigación ha sido financiada con aportes de los National Institutes of Health, con el subsidio 1 R21 NR010856-01. Se presentaron versiones anteriores de este artículo en las siguientes instituciones y/o eventos: Institute for Research on Poverty de la University of Wisconsin-Madison, Department of Sociology de la Yale University, Department of Sociology de la University of California-Los Angeles, Carolina Population Center de la University of North Carolina-Chapel Hill, Center for Poverty and Inequality de la Stanford University, 2008 Annual Meeting of the Population Association of America y el 2008 Annual Meeting of the American Sociological Association. Agradecemos a Sara Goldrick-Rab, Eric Grodsky, Robert Hauser, James Heckman, Michael Hout, Robert Mare, Stephen Morgan, Kimberlee Shauman, Jeff Smith, Hung-Hua Tien, Florencia Torche, Donald Treiman y varios revisores anónimos de la *ASR* por los útiles comentarios y sugerencias. Agradecemos a Ben Jann por su ayuda en la programación del STATA. Las ideas expresadas aquí son las de los autores.

<sup>1</sup> Hay ciertas excepciones a esta afirmación que son dignas de mención. Por ejemplo, Heckman (2007) enfatizó la importancia de las aptitudes socioemocionales –como la motivación, la sociabilidad, la autoestima y la salud– para el logro educativo.

<sup>2</sup> Desafortunadamente, no es posible calcular la variación de los rendimientos de la educación superior a nivel individual (Holland 1986).

<sup>3</sup> Un enfoque de interacción total también agotaría rápidamente los grados de libertad.

<sup>4</sup> Con ponderaciones apropiadas (Morgan y Todd 2008), pudimos obtener efectos del tratamiento promedio para poblaciones específicas, como el efecto tratamiento promedio, el efecto del tratamiento sobre los que recibieron el tratamiento y el efecto del tratamiento sobre los que no lo recibieron (Brand y Halaby 2006). Otros enfoques que estudian los efectos heterogéneos del tratamiento comprenden el modelo de regresión de *switching* (Heckman 1978; Roy 1951; Willis y Rosen 1979), los efectos locales promedio del tratamiento (Angrist, Imbens y Rubin 1996; Angrist y Krueger 1999; Imbens y Angrist 1994) y los efectos marginales del tratamiento (Bjorklund y Moffitt 1987; Heckman et al. 2006).

<sup>5</sup> Por ejemplo, los estudios que usan modelos de regresión de *switching* apelan a un supuesto paramétrico y requieren una teoría fuerte. Los estudios que utilizan variables instrumentales (en el caso del promedio local y de los efectos marginales del tratamiento) enfrentan la dificultad de encontrar una variable instrumental significativa que afecte directamente la asignación del tratamiento, pero que afecte el resultado sólo indirectamente, a través del tratamiento.

<sup>6</sup> También los economistas reconocen el papel relevante que juegan en el logro educativo el contexto familiar (Ashenfelter y Rouse 1998), las expectativas juveniles (Rouse 2004), las capacidades (Cameron y Heckman 2001) y las aptitudes socioemocionales (Heckman 2007).

<sup>7</sup> Las características del barrio y de la escuela, influidas por la segregación económica y racial, también afectan el logro académico de la juventud (MacLeod 1989; Orfield, Eaton y Jones 1996).

<sup>8</sup> De todos modos, hay una cantidad de trabajos bien pagos que no requieren educación universitaria y algunos estudiosos se preguntan cuántos graduados universitarios necesita realmente la fuerza laboral de los Estados Unidos (Rosenbaum 2001).

<sup>9</sup> En otras palabras, las estimaciones de la variable instrumental pueden interpretarse como efectos locales promedio del tratamiento, efectos que pertenecen a unidades inducidas por el instrumento. Las estimaciones de las variables instrumentales también podrían exceder los estimadores de MCO si los instrumentos se correlacionan con la capacidad (véase Carneiro y Heckman 2002).

<sup>10</sup> Notar que en la ecuación 1  $\alpha$  puede considerarse heterogénea, es decir,  $\alpha_i$ , porque no puede separarse del término de error  $U_i$ . Formulamos en detalle la ordenada al origen heterogénea explícitamente en la ecuación 7.

<sup>11</sup> Rosenbaum y Rubin (1984) utilizaron estratos de *score* de propensión, aunque no investigaron la variación de los efectos del tratamiento como una función del *score* de propensión.

<sup>12</sup> Para facilitar la implementación de nuestro método, desarrollamos un nuevo módulo del programa Stata: "hte" (Jann, Brand y Xie 2008), que está disponible para su uso público.

<sup>13</sup> La NLSY79 fue financiada por la Oficina de Estadísticas Laborales del Departamento de Trabajo de los Estados Unidos. La encuesta se llevó a cabo mediante un contrato con el Center for Human Resource Research y el National Opinion Research Center. El National Institute of Child Health and Human Development y el National Institute on Drug Abuse proporcionaron fondos adicionales.

<sup>14</sup> Desde 1991, la mayoría de los fondos para la realización del WLS provinieron del National Institute on Aging (AG-9775 y AG-21079), con fondos adicionales de: Vilas Estate Trust, National Science Foundation, Spencer Foundation y Graduate School of the University of Wisconsin-Madison.

<sup>15</sup> Las muestras del análisis final de la NLSY y del WLS presentan en general más ventajas que las muestras totales, aunque las diferencias son pequeñas.

<sup>16</sup> La forma logarítmica mide diferencias de ingresos proporcionales más que diferencias brutas en dólares. En estratos de *scores* de propensión altos (en comparación con los *scores* bajos), es probable encontrar diferencias brutas mayores que las diferencias logarítmicas en los ingresos entre los individuos que han completado la educación universitaria y los que no lo han hecho. Aunque podríamos apelar a argumentos de peso a favor de las diferencias proporcionales, no encontramos que esta distinción sea significativa. Utilizamos la forma logarítmica para la comparabilidad con estudios previos y para facilitar la interpretación.

<sup>17</sup> En el WLS, no hay medidas de paga por hora en 1975, que puedan compararse con 1993. También analizamos los ingresos de los encuestados de la NLSY que estaban a comienzos de la segunda década de vida y los del WLS que estaban a comienzos de la sexta década. Sin embargo, dados los altos niveles de desempleo (y el desempleo diferencial a través de los estratos) como resultado de transiciones tempranas entre carreras en la NLSY y transiciones por la jubilación en el WLS, nos centramos sólo en los resultados correspondientes a las edades laborales principales.

<sup>18</sup> También probamos todos los modelos donde se mantenían los trabajadores desempleados y les asignamos ingresos nulos. Las conclusiones principales son análogas a las presentadas aquí. Existe poca diferencia en las tasas de participación en la fuerza laboral a través de los estratos durante los años en que los hombres perciben sus mayores ingresos. Sin embargo, es más probable que las mujeres con educación universitaria que se encuentran en los estratos de *score* de propensión alto (en comparación con las de *scores* de propensión bajos) salgan de la fuerza laboral durante los años reproductivos.

<sup>19</sup> Los autores ponen a disposición de los interesados que las soliciten, las estimaciones de las variables de control.

<sup>20</sup> En la NLSY, ajustamos por un indicador de matrimonio y la presencia de hijos a los 25 años.

<sup>21</sup> Excluimos a 238 hombres y 74 mujeres de la NLSY y a 16 hombres y 590 mujeres del WLS, que no cumplían con este requerimiento.

<sup>22</sup> Para presentar menos tablas, mostramos resultados sólo para los hombres de la NLSY. Los resultados, que los autores ponen a disposición de los interesados que los soliciten, son comparables a los de las mujeres de la NLSY y del WLS.

<sup>23</sup> La cantidad de graduados no universitarios de la NLSY y de mujeres y hombres del WLS en nuestro estrato final, y de graduados universitarios mujeres en nuestro primer estrato del WLS, es muy pequeña. Este patrón no debería sorprender, dado que estos individuos completaron o no la educación universitaria contra las posibilidades esperadas. De modo que colapsamos estratos en estas instancias y ajustamos para el *score* de propensión estimado.

<sup>24</sup> Expandimos el rango del eje y para los resultados del WLS para dar cabida al impacto menor (para los hombres) y mayor (para las mujeres) de la educación universitaria en relación con los de la NLSY.

<sup>25</sup> Utilizamos datos del Estudio Longitudinal Nacional de la Clase de *High School* de 1972 (NLS) y estudiamos los efectos de la finalización de los estudios universitarios sobre los ingresos por rango de *score* de propensión en 1986 (32 años). Estos datos representan una tercera cohorte, ubicada entre las cohortes del WLS y las de la NLSY. Los resultados del NLS proporcionan apoyo adicional a la teoría de la selección negativa (Brand y Xie, 2007). Dadas las limitaciones de espacio, no presentamos aquí estos resultados.

<sup>26</sup> En Brand y Xie (2007), investigamos esta idea de manera más explícita, al examinar un cociente de la importancia de factores monetarios y no monetarios en la selección de una carrera a través de estratos de *score* de propensión, para hombres con educación universitaria del Estudio Longitudinal Nacional de la Clase de *High School* de 1972 (NLS). Encontramos que era más probable que los hombres que se hallaban en los estratos de *score* de propensión bajo (en comparación con los de estratos de *score* de propensión alto) dijeran que los factores monetarios eran más importantes que los no monetarios. Si bien todas las mujeres afirman que los factores no monetarios son más importantes que los monetarios, es más probable que las mujeres con *score* de propensión bajo afirmen que los factores monetarios son importantes, lo que establece un patrón de *score* de propensión de los resultados comparable al de hombres.

<sup>27</sup> Utilizamos los datos del WLS para este análisis porque teníamos a nuestra disposición un conjunto rico de covariables. Para la comparación con estudios previos, restringimos el análisis de los ingresos a los hombres de 35 años.

<sup>28</sup> Carneiro, Heckman y Vytlačil (2001) no aceptan el supuesto de ignorabilidad. Por consiguiente, su enfoque difiere del nuestro, lo que proporciona otra explicación posible de la diferencia en los resultados. No nos ocupamos aquí de esta posibilidad.

<sup>29</sup> Los autores ponen los resultados a disposición de los interesados que los soliciten.

**Table A1.**

Logit Models Predicting College Completion for the Generation of Estimated Propensity Scores

Variables	NLSY		WLS	
	Men	Women	Men	Women
Black	-.651* (.268)	-.208 (.256)		
Hispanic	-.792* (.335)	-1.100** (.326)		
Parents' income	.000 (.000)	.000 (.000)	.000*** (.000)	.000* (.000)
Parents' income <sup>2</sup>	.000 (.000)	.000 (.000)	.000* (.000)	.000 (.000)
Mother's education	-.363* (.162)	-.407** (.137)	-.334*** (.092)	-.200 (.109)
Mother's education <sup>2</sup>	.020** (.007)	.021*** (.006)	.017*** (.004)	.018* (.005)
Father's education	.124** (.037)	.072* (.036)	.088*** (.017)	.055** (.017)
Intact family	.281 (.236)	.507* (.230)	.073 (.169)	.025 (.185)
Number of siblings	-.089 (.051)	-.123* (.050)	-.018 (.021)	-.067** (.025)
Rural residence	.066 (.251)	-.098 (.225)	-.099 (.143)	.667*** (.150)
Availability of college	-.419 (.239)	-.065 (.228)	-.092 (.099)	.092 (.106)
Jewish	1.999 (1.032)		1.280* (.609)	1.387** (.409)
Class rank			.028*** (.002)	.027*** (.003)
Mental ability	1.984*** (.298)	1.514*** (.239)	.023*** (.004)	.018*** (.005)
Mental ability <sup>2</sup>	-.386 (.218)	-.305 (.176)		
College track	.603** (.196)	.514* (.199)	.618*** (.132)	.693*** (.003)
Teachers' enc.			.438*** (.107)	.565*** (.117)
Parents' enc.			.996*** (.130)	1.323*** (.146)
Friends' plans	.992*** (.209)	.645** (.204)	.946*** (.101)	.723*** (.116)
LR Chi-Sq.	445.23	304.11	1765.50	1429.10
Prob. > Chi-Sq.	.00	.00	.00	.00
Sample size	1,265	1,203	3,690	4,215

Note: Numbers in parentheses are standard errors.

\*  $p < .05$ ; \*\*  $p < .01$ ; \*\*\*  $p < .001$  (two-tailed tests).

**Table A2.**

Effects of College Completion on Log Wages by Propensity Score Strata: NLSY

	Level-1 Slopes					Level-2 Slopes
	Stratum 1	Stratum 2	Stratum 3	Stratum 4	Stratum 5	
<b>Men</b>						
1994 Earnings (ages 29 to 32)	.351** (.123)	.164 (.121)	.099 (.107)	.212* (.104)	.082 (.082)	-.046 (.038)
1998 Earnings (ages 33 to 36)	.542*** (.150)	.214 (.108)	.225* (.099)	.479*** (.120)	.197 (.164)	-.019 (.046)
2002 Earnings (ages 37 to 40)	.740*** (.163)	.234 (.191)	.387** (.133)	.479* (.188)	.408* (.178)	-.053 (.055)
<b>Women</b>						
1994 Earnings (ages 29 to 32)	.343** (.119)	.480*** (.109)	.137 (.107)	.373 (.142)	.401* (.168)	-.013 (.044)
1998 Earnings (ages 33 to 36)	.228* (.103)	.376** (.121)	.169 (.114)	.167 (.140)	.253 (.265)	-.024 (.047)
2002 Earnings (ages 37 to 40)	.210 (.160)	.452* (.176)	.359** (.134)	.116 (.193)	.039 (.368)	-.035 (.068)

Note: Numbers in parentheses are standard errors.

\*  $p < .05$ ; \*\*  $p < .01$ ; \*\*\*  $p < .001$  (two-tailed tests).

**Table A3.**  
Effects of College Completion on Log Wages by Propensity Score Strata: WLS

	Level-1 Slopes									Level-2 Slopes
	Stratum 1	Stratum 2	Stratum 3	Stratum 4	Stratum 5	Stratum 6	Stratum 7	Stratum 8	Stratum 9	
<b>Men</b>										
1975 Earnings	.175	-.104	-.243	.195	.288*	.360*	-.364	.007	.029	-.003
(age 35)	(.268)	(.269)	(.277)	(.225)	(.132)	(.146)	(.244)	(.261)	(.249)	(.033)
1993 Earnings	.444***	.126	.428*	.374**	.360***	.255***	.221	.188	.285*	-.017
(age 52)	(.118)	(.133)	(.168)	(.112)	(.080)	(.073)	(.127)	(.151)	(.132)	(.017)
<b>Women</b>										
1975 Earnings	1.079*	.358	.433	.376	.053	.235				-.143
(age 35)	(.462)	(.403)	(.321)	(.208)	(.223)	(.329)				(.087)
1993 Earnings	.385**	.141	.283**	.333***	.011	.138				-.052
(age 52)	(.147)	(.137)	(.106)	(.076)	(.087)	(.103)				(.029)
<b>WLS Men, Small Set of Covariates</b>										
1975 Earnings	.181	.031	.392**	.189	.082	.224	.423			.031
(age 35)	(.117)	(.099)	(.119)	(.204)	(.202)	(.219)	(.246)			(.032)

Note: Numbers in parentheses are standard errors.

\*  $p < .05$ ; \*\*  $p < .01$ ; \*\*\*  $p < .001$  (two-tailed tests).

**Table 1.**  
Descriptive Statistics of Precollege Covariates

Variables	NLSY Means				WLS Means			
	Men (N = 1,265)		Women (N = 1,209)		Men (N = 3,690)		Women (N = 4,215)	
	Non-College Graduate	College Graduate	Non-College Graduate	College Graduate	Non-College Graduate	College Graduate	Non-College Graduate	College Graduate
Race								
Black	.18	.07	.15	.07				
Hispanic	.07	.03	.07	.03				
Social Background								
Parents' income	17870	26538	18174	25991	5605	8123	5622	9262
Mother's education	11.26	13.32	11.18	13.37	10.15	11.56	9.94	12.02
Father's education	11.23	14.39	11.16	14.14	9.10	11.37	9.21	11.79
Intact family (0–1)	.72	.83	.67	.85	.90	.92	.90	.92
Number of siblings	3.29	2.34	3.40	2.45	3.45	2.61	3.51	2.40
Rural residence (0–1)	.25	.19	.24	.21	.22	.12	.20	.16
Urban res. / prox. to college	.77	.78	.75	.80	.42	.50	.50	.53
Jewish (0–1)	.00	.03	.00	.04	.00	.02	.00	.03
Ability and Academics								
Class rank					35.76	65.49	53.78	79.51
Mental ability (IQ)	–.09	.69	–.04	.64	97.03	111.75	98.67	112.00
College-prep (0–1)	.23	.59	.23	.49	.54	.91	.46	.89
Social-Psychological								
Teachers' encouragement					.35	.75	.36	.77
Parents' encouragement					.47	.91	.39	.90
Friends' college plans	.42	.79	.48	.81	.22	.66	.30	.76
Weighted Sample Proportion	.76	.24	.77	.23	.69	.31	.82	.18

*Notes:* Parents' income is measured as total net family income in 1979 dollars in the NLSY and in 1957 dollars in the WLS. Urban residency / proximity to college indicates whether a respondent lived in an SMSA in the NLSY and whether a respondent's high school was within 15 miles of a college or university in the WLS. Mental ability is measured with a scale of standardized residuals of the ASVAB in the NLSY and with the Henmon-Nelson IQ test in the WLS. College-prep indicates whether a student was enrolled in a college-preparatory curriculum in the NLSY or whether a student completed the requirements for UW-Madison in the WLS.

**Table 2.**  
Effects of College Completion on Log Wages under the Assumption of Homogeneity

	Men	Women
NLSY		
1994 Wages (ages 29 to 32)	.180*** (.047)	.276*** (.051)
1998 Wages (ages 33 to 36)	.296*** (.054)	.188*** (.052)
2002 Wages (ages 37 to 40)	.410*** (.069)	.216** (.075)
WLS		
1975 Earnings (age 35)	.124 (.067)	.380** (.113)
1993 Wages (age 53)	.302*** (.034)	.225*** (.038)

*Note:* Numbers in parentheses are standard errors. Treatment effects are conditional on the set of covariates for each data source described in Table 1. NLSY estimates further condition on age at baseline. NLSY estimates for women also condition on an indicator for married with children at age 25. All outcome variables are current hourly wages, except for WLS 1975 earnings, which are current yearly earnings. Unemployed workers are omitted.

\*  $p < .05$ ; \*\*  $p < .01$ ; \*\*\*  $p < .001$  (two-tailed tests).

**Table 3.**  
Mean Covariate Values by Propensity Score Strata: NLSY Men

Variables	Propensity Score Strata									
	[.0-.1)		[.1-.2)		[.2-.4)		[.4-.6)		[.6-1.0)	
	Non-Coll. Grad.	Coll. Grad.	Non-Coll. Grad.	Coll. Grad.	Non-Coll. Grad.	Coll. Grad.	Non-Coll. Grad.	Coll. Grad.	Non-Coll. Grad.	Coll. Grad.
Black	.37	.25	.21	.40	.23	.23	.19	.12	.15	.07
Hispanic	.18	.30	.12	.08	.12	.09	.12	.09	.07	.05
Parents' income	13381	12253	17614	18482	19324	18422	23062	21348	23469	34702
Mother's edu.	10.31	10.05	11.67	12.16	11.98	12.21	12.71	12.54	13.67	14.79
Father's edu.	10.17	9.95	11.79	10.72	12.08	12.53	13.33	13.97	15.11	16.30
Intact family	.63	.55	.63	.80	.74	.67	.85	.80	.85	.91
Num. of siblings	3.84	4.05	3.04	3.04	2.64	2.47	2.88	2.46	2.04	2.17
Rural resident	.21	.30	.26	.20	.21	.21	.19	.12	.11	.20
Availability coll.	.76	.70	.80	.84	.75	.77	.73	.77	.81	.78
Jewish	.00	.00	.00	.00	.01	.00	.00	.02	.04	.08
Mental ability	-.14	-.01	.31	.48	.62	.57	.79	.76	.90	1.05
College track	.17	.16	.32	.37	.41	.52	.57	.55	.83	.73
Friends' plans	.35	.55	.61	.52	.66	.74	.90	.85	.93	.93

**Table 4.**  
Frequency Counts per Propensity Score Stratum

P-Score	NLSY						WLS					
	Men			Women			Men			Women		
	$d=0$	$d=1$	P-Score	$d=0$	$d=1$	P-Score	$d=0$	$d=1$	P-Score	$d=0$	$d=1$	
[.00, .10)	454	20	[.00, .05)	573	12	[.00, .05)	931	28	[.00, .05)	1367	27	
[.10, .20)	135	25	[.05, .10)	181	17	[.05, .10)	418	33	[.05, .10)	441	38	
[.20, .40)	130	43	[.10, .20)	156	28	[.10, .15)	255	25	[.10, .20)	367	67	
[.40, .60)	52	65	[.20, .40)	147	47	[.15, .20)	155	45	[.20, .40)	391	172	
[.60, 1.00)	27	76	[.40, .60)	37	48	[.20, .40)	386	149	[.40, .60)	204	185	
			[.60, 1.00)	19	55	[.40, .60)	208	200	[.60, 1.00)	101	265	
						[.60, .70)	72	122				
						[.70, .80)	46	173				
						[.80, 1.00)	48	380				

**Table 5.**  
Mean “Value of College” by Propensity Score Strata: WLS Men

	Propensity Score Strata								
	[.0-.05)	[.05-.1)	[.1-.15)	[.15-.2)	[.2-.4)	[.4-.6)	[.6-.7)	[.7-.8)	[.8-1.0)
Non-college grad.	45.12	54.55	63.29	66.09	74.82	82.52	84.90	88.73	85.36
College grad.	58.86	70.59	81.00	82.77	85.26	86.79	89.13	88.48	88.92

*Note:* The variable “value of college” is a weighted average of the scores in response to a series of 18 statements regarding perceptions of the value of going to college.

**Table 6.**  
Proportion of College Majors for College-Educated Men by Propensity Score Strata: WLS Men

College Major	Propensity Score Strata								
	[.0-.05)	[.05-.1)	[.1-.15)	[.15-.2)	[.2-.4)	[.4-.6)	[.6-.7)	[.7-.8)	[.8-1.0)
Physical science	.00	.06	.04	.02	.03	.05	.05	.04	.05
Math	.00	.06	.04	.02	.06	.09	.08	.04	.05
Biological science	.11	.03	.04	.02	.09	.09	.11	.07	.12
Engineering	.04	.06	.13	.12	.06	.14	.13	.23	.22
Pre-professional	.00	.00	.00	.00	.00	.01	.01	.01	.02
Computer science	.04	.00	.04	.00	.01	.02	.01	.01	.01
Business	.19	.27	.17	.19	.16	.15	.10	.11	.10
Social science	.15	.15	.25	.17	.18	.19	.10	.22	.21
Humanities	.04	.03	.00	.10	.13	.08	.13	.11	.10
Art and music	.11	.09	.04	.07	.04	.05	.05	.01	.05
Education	.22	.18	.21	.14	.15	.08	.07	.06	.05
Communications	.04	.03	.00	.02	.06	.01	.01	.04	.01
Agriculture	.04	.00	.00	.02	.01	.01	.02	.04	.01
Other	.04	.03	.04	.10	.02	.03	.03	.04	.02
Number	27	33	24	42	145	196	120	171	375