

Debates de Política Pública

Estéfano Rubio

El diálogo de dos desafíos: Evolución y relación de la desigualdad y la escolaridad en Chile

Debates de Política Pública

Nº 15 / mayo 2016

La serie *Debates de Política Pública* es editada en formato digital por el Centro de Estudios Públicos (CEP). El director y representante legal del CEP es Harald Beyer Burgos. Dirección: Monseñor Sótero Sanz 162, Providencia, Santiago de Chile. Fono: 2 2328 2400. Fax: 2 2328 2440.

Sitio web: www.cepchile.cl. Email: escribanos@cepchile.cl.

Cada artículo es responsabilidad de su autor y no refleja necesariamente la opinión del CEP. Esta institución es una fundación de derecho privado, sin fines de lucro, cuyo objetivo es el análisis y difusión de los valores, principios e instituciones que sirven de base a una sociedad libre.

Edición gráfica: David Parra Arias.

El diálogo de dos desafíos: Evolución y relación de la desigualdad y la escolaridad en Chile

Estéfano Rubio

Ingeniero Comercial y Economista (Universidad de Chile). Magíster en Políticas Públicas (Universidad de Chile). Investigador del CEP. Email: erubio@cepchile.cl

* Agradezco los valiosos comentarios y sugerencias de Loreto Cox, Harald Beyer, Slaven Razmilic, Francisco Szederkenyi y Andrés Hernando.

Síntesis

A continuación se presenta un análisis de la relación entre escolaridad y desigualdad en Chile. Se analiza la evolución de la desigualdad de ingresos desde el año 1990 al 2013, período en que el país ha experimentado una caída en el índice de Gini desde 0,56 a 0,52. Este hecho se intenta explicar desde dos enfoques: por un lado se analizan cuáles han sido los factores que han favorecido esta disminución y, por el otro, se estudian posibles causas de por qué estas cifras han permanecido tan elevadas durante los años. En relación a lo primero se incluyen los siguientes factores: leve caída en los retornos a la educación superior, reducción en la dispersión de escolaridad de la población y efectos período favorables. Entre las razones que explican por qué estas cifras continúan siendo elevadas, están: el aumento en el promedio de la escolaridad de la población —sumado a que los retornos de la escolaridad crecen con el nivel educacional—; un elevado retorno de la educación superior (universitaria especialmente); elevada participación del salario en los ingresos totales; y factores de fondo que establecen una estructura social que limita la masificación de la escolaridad en grupos de menos recursos (una baja movilidad educacional intergeneracional sumada a elevados índices de emparejamiento selectivo u homofilia). Por último se simula un escenario contrafactual en el cual no hubiese existido homofilia por parte de los padres de cada individuo; se encuentra que el Gini del ingreso laboral (a nivel de individuos) caería de 0,48 a 0,43, aunque asociado con una reducción de 15% de los salarios promedio.

Introducción

Bastante se ha hablado acerca de uno de los principales desafíos pendientes de nuestro país: la desigualdad. Ya sea que se analice cómo ha evolucionado, sus componentes principales, sus consecuencias o los posibles mecanismos disponibles que hay para enfrentarla, la desigualdad es un tópico recurrente para los académicos y varios *policy makers*. Este trabajo aborda este tema estudiando su relación con uno de los principales factores que pueden permitir disminuirla: la escolaridad de la población (Beyer, 2000; 2011; OECD, 2011; Sapelli, 2011).

El presente documento otorga un panorama de la desigualdad entre los años 1990 y 2013, período en el cual Chile experimentó un elevado crecimiento económico y logró disminuir de forma considerable sus cifras de pobreza, desde 38,6% a 7,8% (MIDESO, 2013). Durante dicho lapso de tiempo, el índice de Gini¹ (calculado sobre el ingreso monetario) cayó desde 0,56 a 0,52². El avance que ha tenido nuestro país debe ponerse en contraste con el retroceso que han experimentado los países miembros de la OECD que, en promedio, pasaron de un Gini de 0,30 en 1990 a uno de 0,32 en 2012³. No obstante, y a pesar del positivo avance de nuestro país en las últimas décadas, todavía queda un largo camino por recorrer.

Para explicar el progreso que ha tenido Chile en estas cifras miraremos qué ha pasado con el promedio y la dispersión de escolaridad de la población, y con la evolución de los retornos de ésta. Para lo primero realizaremos un análisis de edad, año y cohorte utilizando un panel sintético creado a partir de las encuestas Casen de los años 1990 al 2013. Las cohortes estudiadas van desde 1926 hasta 1995. De esta manera observaremos el impacto que ha tenido el aumento en la escolaridad promedio y la disminución de su varianza en la desigualdad, controlando por los tres efectos mencionados (edad, año y cohorte). En paralelo mostraremos cómo han cambiado los retornos por los diferentes niveles educacionales, también para los mismos años en estudio; donde destaca una caída en los retornos de educación básica y media y un alza, seguida de una caída mayor, de los de educación superior. De todas formas, estos últimos siguen siendo relativamente elevados.

Los anteriores podríamos catalogarlos como los efectos directos de la escolaridad sobre la desigualdad. Sin embargo, también estudiaremos dos factores que definen una estructura sobre la cual la primera puede influir sobre la segunda. Uno de estos factores corresponde a lo que en la literatura se menciona como *assortative mating* (emparejamiento selectivo) u homofilia, es decir,

¹ Principal indicador utilizado para medir la desigualdad y cuyos valores van entre 0 (perfecta igualdad) y 1 (perfecta desigualdad).

² Datos estimados con la encuesta Casen aplicando la metodología tradicional, con ajuste por cuentas nacionales.

³ <http://www.oecd.org/els/soc/income-distribution-database.htm>. Estas cifras se refieren al Gini después de impuestos y transferencias monetarias, en donde en general en los países de la OCDE, a diferencia de Chile, se redistribuye mucho a través de los subsidios en dinero y principalmente (aunque no únicamente) en pensiones.

la tendencia de los individuos a relacionarse o formar parejas con personas de características similares a ellos. Una de éstas es la escolaridad. El segundo factor corresponde a la baja movilidad educacional intergeneracional. Ambos contribuyen a aumentar la desigualdad de ingresos.

El primero, por un lado, mediante el hecho que fomenta que se formen parejas de alta escolaridad y otras de baja escolaridad, lo que genera una concentración de ingresos en los hogares del primer tipo; en contraste con un escenario en donde hubiese núcleos familiares que mezclasen individuos con alta y baja escolaridad (y por lo mismo con altos y bajos ingresos, respectivamente).

Por otro lado, el segundo factor provoca una herencia en los hijos de los niveles educativos de los padres, lo cual produce una concentración intergeneracional de la riqueza y reduce las posibilidades de meritocracia en la sociedad.

Esto se ve intensificado debido al *assortative mating*, ya que al agruparse entre sí las personas educadas —sumado a que será su descendencia la que con mayor probabilidad heredará estos niveles educativos— se reducirían aún más las posibilidades de salir adelante para muchos de quienes actualmente tienen ambos padres con baja educación, provocando que la distribución de los niveles de escolaridad (y por ende de ingresos) sea desigual entre ciertos sectores socioeconómicos, limitando considerablemente los avances que pudieran hacerse en materia de desigualdad.

Panorama de la desigualdad en Chile: 1990-2013

Los indicadores usados con mayor frecuencia para medir la desigualdad son el índice de Gini y las razones de deciles. Estos últimos tienen el fin de comparar de forma directa los ingresos de cierta parte de la población, en contraste con el Gini, el cual es calculado tomando en cuenta los ingresos de todos los individuos. No obstante, estas razones permiten observar con más detención los extremos de la distribución de ingresos, mientras que el Gini pone mayor énfasis en los tramos medios⁴.

Para calcular la razón de deciles se ordenan los hogares de menor a mayor de acuerdo a su ingreso autónomo per cápita. Posteriormente se dividen en diez grupos iguales, conformando así los deciles. Dada la elevada concentración de ingresos de Chile, suele ser especialmente relevante la razón 10/10, puesto que permite observar cuánto más ingresos posee el decil 10 (mayores ingresos) con el decil 1 (menores ingresos).

A continuación presentamos en la tabla 1 la evolución de estos indicadores para todos los años del período en los cuales se realizó la encuesta Casen:

⁴ En efecto, por la fórmula de cálculo de este índice, es más sensible a cambios de ingresos en el medio de la distribución.

TABLA 1: Indicadores de Desigualdad⁵ para Ingresos Autónomos y Monetarios

	Ingreso Autónomo										
	1990	1992	1994	1996	1998	2000	2003	2006	2009	2011	2013
10/10	30,5	28,1	30,9	33,0	34,7	34,2	34,4	31,3	46,0	39,9	35,4
20/20	14,0	13,2	14,0	14,8	15,6	14,5	14,5	13,1	15,7	14,2	13,7
Gini	0,57	0,56	0,57	0,57	0,58	0,58	0,57	0,54	0,55	0,54	0,53
	Ingreso Monetario										
	1990	1992	1994	1996	1998	2000	2003	2006	2009	2011	2013
10/10	27,1	25,2	27,7	28,7	28,4	29,5	27,3	23,9	25,9	24,3	21,7
20/20	13,0	12,3	13,1	13,6	13,9	13,3	12,8	11,5	11,9	11,3	10,7
Gini	0,56	0,56	0,55	0,56	0,57	0,58	0,56	0,53	0,53	0,53	0,52

Fuente: Elaboración propia a partir de la Encuesta CASEN, años respectivos. Ministerio de Desarrollo Social.

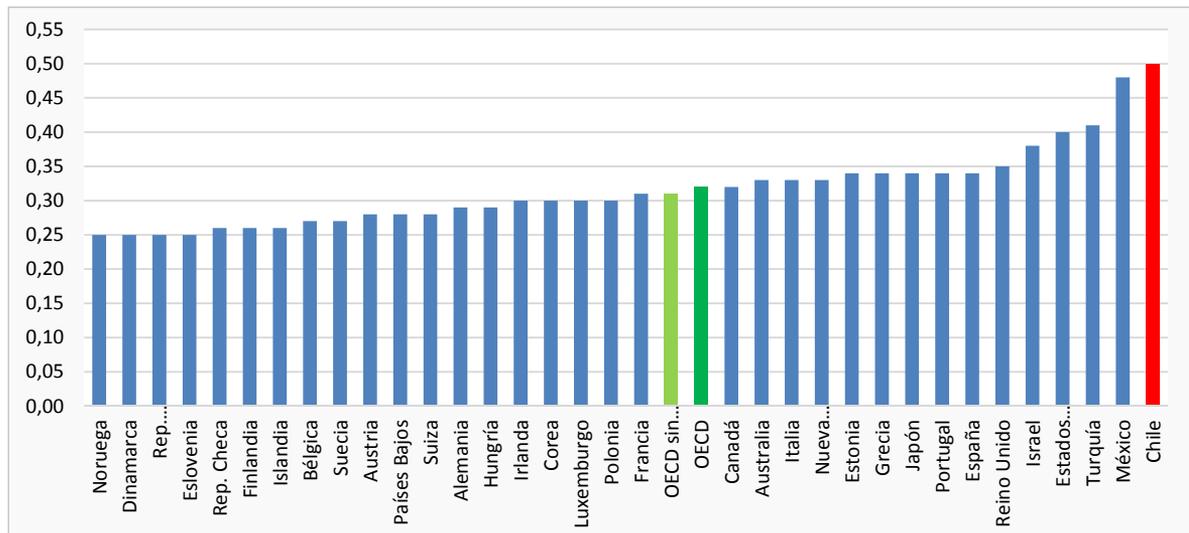
Podemos observar que, tanto para el ingreso autónomo como para el monetario, el coeficiente de Gini pasó de una etapa de estancamiento en los años 90 a una leve alza para finales de esa década, para finalmente instalarse en una senda de descenso desde el 2000 en adelante. Tan sólo notamos un aumento de un punto porcentual del Gini en el ingreso autónomo para el año 2009, en donde pudieran notarse las consecuencias de la crisis económica global de fines del año 2008.

Por último, al analizar la evolución de las razones de deciles 10/10 y 20/20 observamos que la tendencia seguida por el Gini también ocurre para estos otros índices (estancamiento en los años 90, alza para finales de la década, para finalmente caer desde el 2000, pasando por un alza en el 2009). Lo relevante de esto es que si bien estos indicadores han evolucionado positivamente para el ingreso monetario (el cual incluye las transferencias monetarias realizadas por el gobierno), no lo han sido para el ingreso autónomo. Para éste, el índice 10/10 indica que si en 1990 los hogares del 10% más rico ganaban 30,5 veces más que el 10% más pobre, ahora dicha diferencia aumentó a 35,4 veces. Este ratio fue especialmente alto para el 2009, lo que podría interpretarse como que quienes más salen perjudicados de las crisis económicas son las familias con menores capacidades para generar ingresos; los efectos en desigualdad de esto aún no pueden ser superados, ya que todavía no retornamos a los niveles previos a la crisis (31,3 en 2006). Por otro lado, la razón 20/20 permanece casi inalterada al comparar la primera y última cifra del período, lo que contrasta de manera importante con el ratio 10/10. Este es un claro indicador de que gran parte de las desigualdades de ingresos provienen por las diferencias entre los deciles más extremos.

⁵ Es importante que recordemos que los datos están estimados aplicando la metodología tradicional, con ajuste por cuentas nacionales.

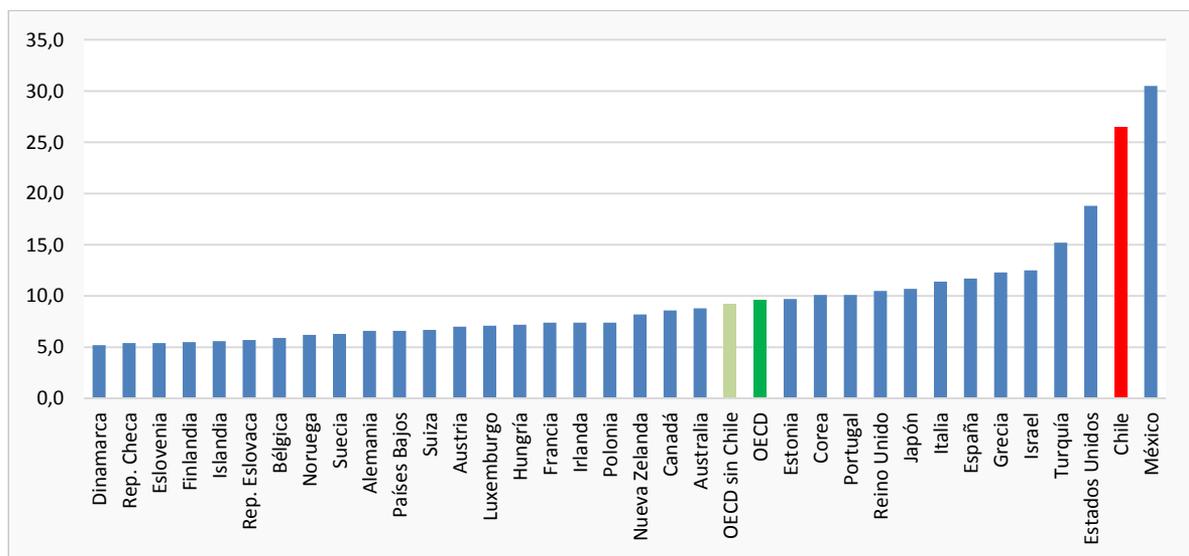
Es importante que ahora pongamos en perspectiva la realidad de nuestro país con la del resto del mundo. A continuación se presenta, en los gráficos 1 y 2, las últimas cifras⁶ oficiales de desigualdad la OECD para cada uno de sus países miembros. En el primero se muestra el índice de Gini mientras que en el segundo la razón de deciles 10/10.

GRÁFICO 1: Coeficiente de Gini del ingreso monetario para países de la OECD



Fuente: Elaboración propia a partir de datos OECD, Income Distribution and Poverty Data.

GRÁFICO 2: Razón de deciles 10/10 para países de la OECD



Fuente: Elaboración propia a partir de datos OECD, Income Distribution and Poverty Data.

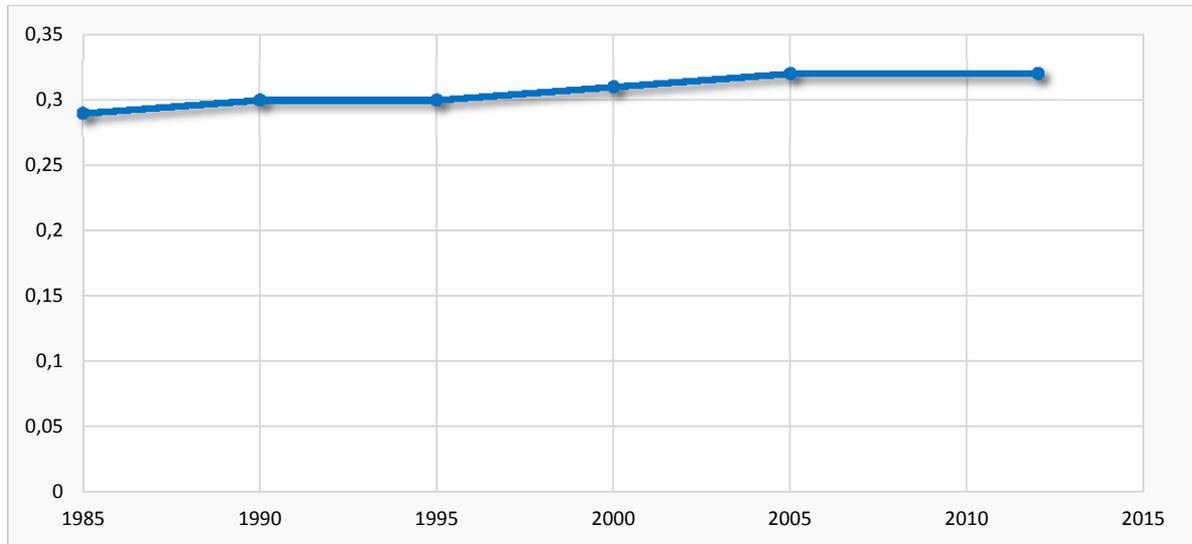
⁶ Cálculo del índice en base a ingreso por hogar equivalente (ajustado por tamaño del hogar) después de impuestos y transferencias.

Mientras en los países nórdicos y en algunos de Europa del este el índice de Gini es cercano al 0,25, para nuestro país dicha cifra, en 2011, alcanza el doble (0,50) y está muy lejana además del promedio de los demás países de la OECD que es de un 0,31. Chile es quien lidera el ranking en desigualdad para este conjunto de naciones.

Para el caso del gráfico 2, otro indicador de desigualdad, nuestro país está sólo detrás de México. La razón de deciles para Chile en el año 2011, según las cifras de la OECD, es de 26,5. Casi tres veces el promedio de los demás países de esta institución, el cual es de 9,2.

Si bien un análisis estático no favorece a nuestro país en relación a los demás miembros de la OECD, si miramos las tendencias de largo plazo ahí sí podemos esgrimir argumentos en favor de Chile. Vemos que en promedio el Gini de la OECD posee una tendencia al alza, desde un 0,30 en 1990 a un 0,32 en 2012. Esta alza en 2 puntos porcentuales contrasta con la caída de Chile de 4 puntos, desde 0,56 a 0,52⁷ para el período entre 1990 y el 2013. Sin embargo, por ningún motivo podemos ser conformistas en este aspecto. Nuestro país tiene un largo trecho por recorrer aun en esta materia. En el gráfico 3 podemos ver el progreso del Gini para la OECD entre 1985 y 2012.

GRÁFICO 3: Evolución del promedio OECD del Coeficiente de Gini



Fuente: Elaboración propia a partir de datos OECD.

De esta forma podemos extraer las siguientes cuatro principales conclusiones a partir de los datos presentados en esta sección:

⁷ Esta comparación es meramente ilustrativa de las tendencias de largo plazo, pero recordemos que las cifras de la OECD calcula el índice en base a ingreso por hogar equivalente (ajustado por tamaño del hogar) después de impuestos y transferencias.

1. La caída en las razones 10/10 y 20/20 para el ingreso monetario nos hablan de la efectividad que han tenido las políticas sociales chilenas en disminuir la desigualdad, más que compensando las desigualdades crecientes o permanentes del ingreso autónomo para estos mismos indicadores. Ello no sólo habla bien de las políticas públicas de nuestro país, sino que también de una efectiva focalización de éstas (lo que no quita que con una focalización más efectiva su impacto hubiese sido mayor).
2. ¿Qué ha pasado con los deciles extremos? El alza del ratio 10/10 y el estancamiento en el de 20/20 para los ingresos autónomos (comparando el inicio y el final del período analizado), nos indica que para los deciles 2 y 9 (el segundo decil más pobre y el segundo decil más rico, respectivamente), ha habido una evolución positiva en ámbitos de desigualdad. Es decir, los ingresos autónomos del decil 2 crecieron más que proporcionalmente que los del decil 9, debido a que el ratio 20/20 se mantuvo relativamente estable durante el período, en contraste con el 10/10, y tomando en cuenta que el primero también incluye los ingresos de los deciles 1 y 10, entonces para que se haya mantenido estable el indicador 20/20 implica necesariamente que los deciles 2 y 9 hayan compensado el alza en desigualdad del 1 y 10. Todo ello confirma que los únicos tramos donde la desigualdad no ha evolucionado de forma positiva es en los ingresos autogenerados de los deciles más extremos. Esto nos obliga a observar con mayor detalle las causas de este síntoma para poder sugerir un tratamiento apropiado del problema.
3. ¿Qué ha pasado con la clase media? Dado que el índice de Gini es más sensible a cambios en el medio de la distribución de ingresos, podemos observar que, para dicho sector especialmente, la evolución en desigualdad ha sido positiva, considerando tanto ingresos autónomos como monetarios, lo que habla de un crecimiento relativamente más alto del ingreso de la clase media.
4. A pesar de mejoras en el margen, nuestro país todavía está entre los países con mayores niveles de desigualdad dentro de la OCDE, con un coeficiente de Gini de 0,50, frente al 0,32 del promedio de los países miembros (aunque este es un problema que es transversal para los demás países de Latinoamérica). Si bien se aprecia, dadas las tendencias de largo plazo, que existe una convergencia favorable para nuestro país, todavía queda mucho por hacer al respecto.

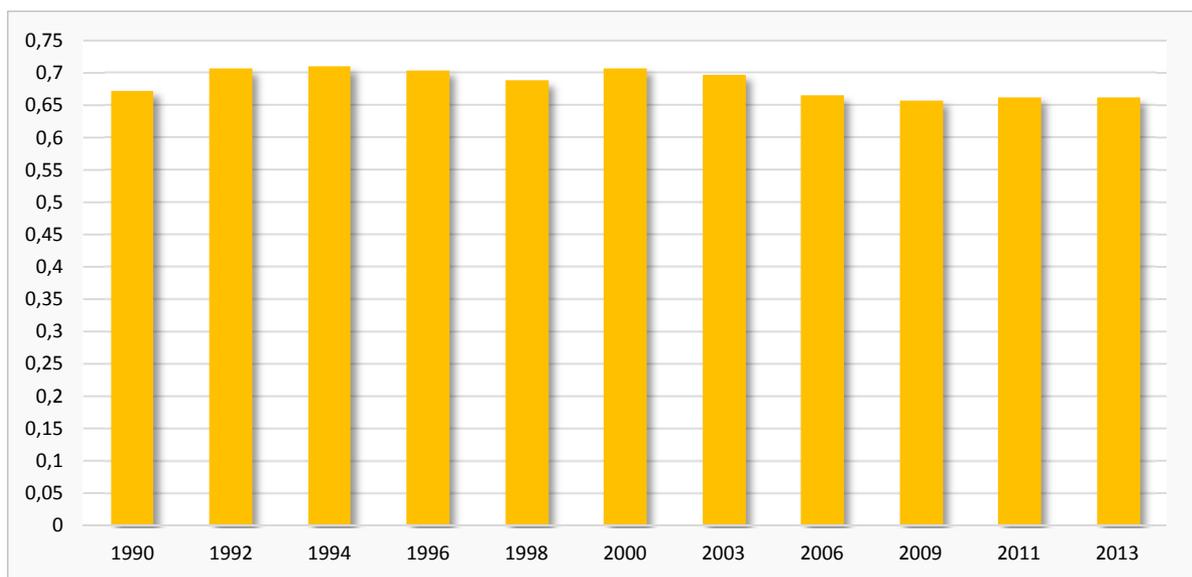
La relación entre escolaridad y desigualdad

Ya manifestábamos al final de la sección anterior lo preocupante de la elevada y creciente desigualdad existente entre los deciles más extremos de la distribución de ingreso autónomo. Ante ello cabe preguntarse cuál es la causa de este fenómeno, así como por qué, en general, se han mantenido elevados los indicadores de desigualdad.

Tal como diagnostican Beyer (2000, 2011), Contreras et al. (2013), Sapelli (2003, 2011), entre otros, una de las principales causas de la desigualdad es debido a la heterogeneidad de los retornos a los diferentes niveles educacionales. Esto ocurre debido al importante impacto que posee la escolaridad a la hora de explicar los salarios de las personas y del importante porcentaje que estos últimos representan del ingreso total de los hogares.

En el gráfico 4 se muestra la evolución que ha tenido la participación del ingreso de la ocupación principal del hogar como porcentaje del ingreso total de los hogares.

GRÁFICO 4: Porcentaje de participación del ingreso de la ocupación principal del hogar en el ingreso total del hogar



Fuente: Elaboración propia. Cálculos basados en micro-datos encuestas Casen, años respectivos.

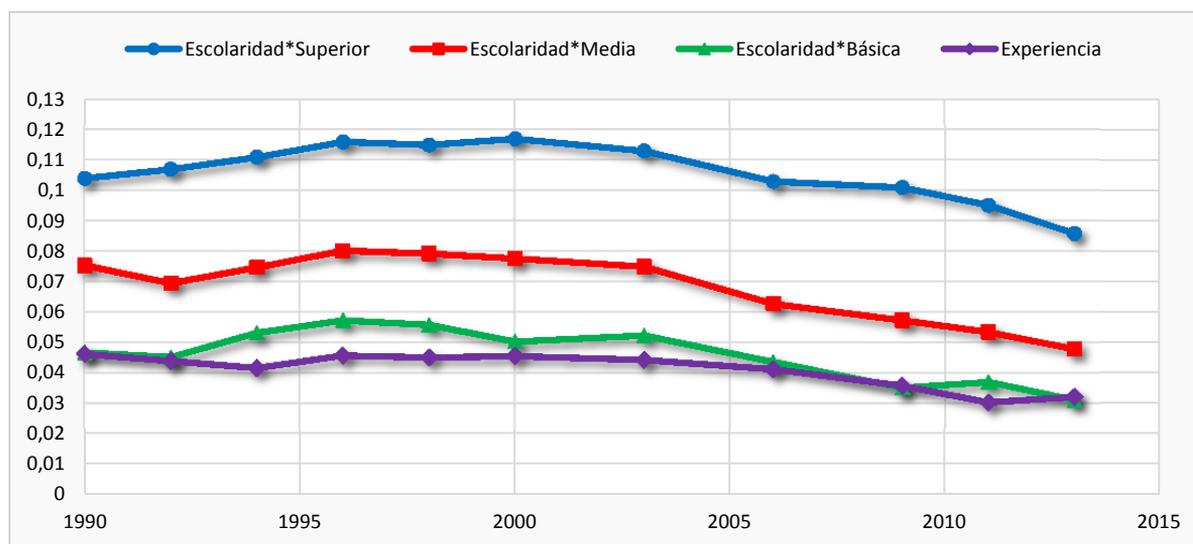
Lo que vemos a partir de estos datos es una tendencia similar a la que posee la razón de deciles y el Gini para el mismo período. Apreciamos un leve aumento para inicios del año 2000 y luego una caída durante la siguiente década. Esto nos habla de que períodos en donde ha habido mayor desigualdad coinciden con una alta participación de los salarios en los ingresos totales de los hogares y viceversa. De igual forma, tal como la desigualdad, que se ha mantenido elevada en todo el período, también lo ha hecho el porcentaje en cuestión, con un máximo de 71% en 1994 y un mínimo de 65,7% en 2009.

Evolución del premio a la educación

La relación entre evolución de la tasa de participación de los salarios en el ingreso total con el nivel de desigualdad confirma el hecho del importante rol que tendría la escolaridad para determinar este último. Es por esto que a continuación analizaremos cómo ha cambiado, para el

período en cuestión, el premio a la educación, desagregando por los distintos niveles escolares. Para ello procederemos a hacer una regresión de tipo *spline* sobre la típica ecuación de Mincer (1974)⁸. El gráfico 5 nos muestra los coeficientes obtenidos⁹ para dicha regresión para cada año en que fue tomada la encuesta Casen, desde 1990. Es importante mencionar que la estimación fue aplicada sólo para los hombres, ya que al tomar en cuenta la expansión en la tasa de participación laboral femenina durante el período (Beyer, 2011), el no desagregar los datos pudiera provocar que los coeficientes pudiesen estar capturando este fenómeno.

GRÁFICO 5: Evolución coeficientes regresión *spline* para hombres



Fuente: Elaboración propia. Cálculos basados en micro-datos encuestas Casen, años respectivos.

Las estimaciones de la regresión anterior se han hecho bajo el supuesto que los retornos de la educación difieren según el nivel educacional de la persona, en particular dependiendo de si posee educación básica, media o superior (Beyer, 2000; Sapelli, 2003). Bajo esta metodología, el retorno de la educación universitaria correspondería a la suma de los coeficientes de las variables de 'Escolaridad*Básica', 'Escolaridad*Media' y 'Escolaridad*Superior'. Así, para el año 2013, cada año adicional de educación de enseñanza básica aporta del orden de un 3% adicional en salarios, para la educación media el aporte marginal de un año adicional es cercano a un 8%, mientras que

⁸ Donde las variables independientes son la escolaridad por variables dicotómicas de niveles educacionales, la experiencia y la experiencia al cuadrado.

⁹ No se incluye la variable de experiencia al cuadrado en el gráfico. Todas las variables resultaron ser significativas al 0,1% en todos los períodos. Las variables Escolaridad*Básica, Escolaridad*Media, Escolaridad*Superior multiplican el nivel de escolaridad del individuo por una variable dicotómica que indica si posee dicho nivel educativo, respectivamente; el cual corresponde a la metodología de la regresión *spline*. Cabe mencionar también que los ingresos de cada año fueron ajustados por inflación, tomando como año base el 2013.

el aporte marginal de un año de educación superior es de un 17%, como podemos observar en detalle en la tabla 2.

TABLA 2: Retorno de la Escolaridad por nivel educacional alcanzado para hombres

Variables	Coefficientes
Escolaridad*Básica	0,0312
Escolaridad*Media	0,0479
Escolaridad*Superior	0,0859
Experiencia	0,0322
Experiencia ²	-0,0004
Constante	11,82
N	50.723
R ²	0,322

Nota: Todos los coeficientes son significativos al 0,1%

Fuente: Elaboración propia a partir de Casen 2013.

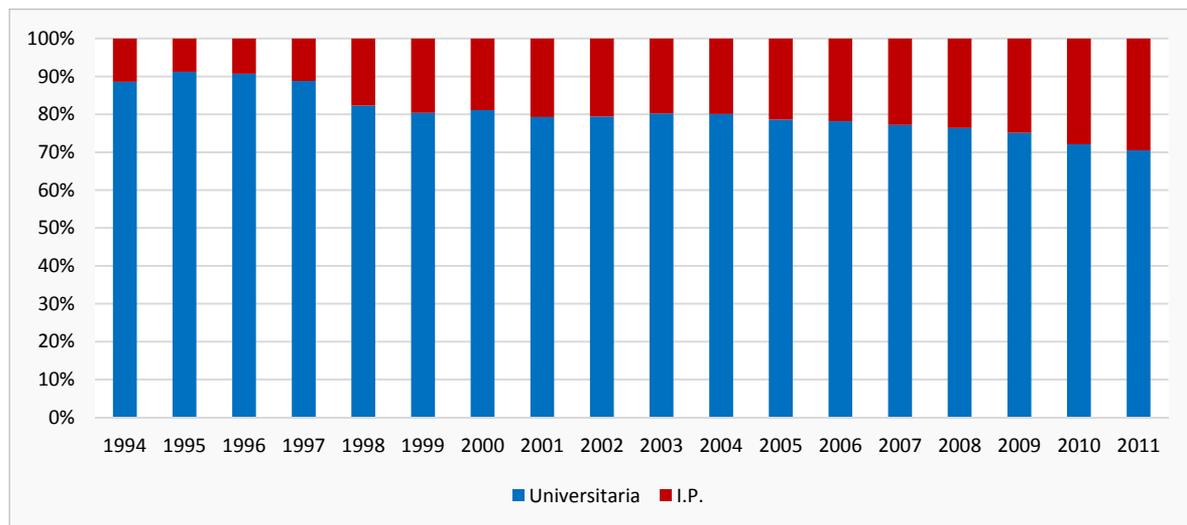
A partir de los datos del gráfico 5 podemos notar que el retorno a la educación ha caído para el período estudiado, en especial para los niveles de básica y media. La caída en el retorno de la educación superior ha sido menor en términos relativos y ha seguido un patrón muy similar al de los indicadores de desigualdad (presentando un alza entre fines de los años 90 e inicios de los 2000, para luego caer en la siguiente década). Este factor es clave para explicar los altos índices de desigualdad, pues tomando en cuenta la gran injerencia que tienen los ingresos laborales en el ingreso total, es clara la ventaja que adquieren, en estos términos, las personas con educación superior.

Chile, durante las últimas dos décadas, se ha convertido en un país globalizado. A partir del proceso de globalización se puede justificar los altos retornos al capital humano de mayor calificación (OCDE, 2011). Sin embargo, lo paradójico es que se hayan mantenido tan elevados los retornos a la educación superior a pesar de que la oferta laboral de los grupos más calificados se ha expandido considerablemente. Esto lo podemos observar al analizar la evolución de las matrículas en instituciones de educación superior. La tabla del gráfico 6 nos muestra los datos provistos por el Consejo Nacional de Educación que van desde 1994 al 2011. En la primera observación, la matrícula total ascendía a 149.689, mientras que en la última llegaba a 854.503, casi seis veces más. En contraste, el premio de un año adicional de educación superior en 1994 era de 22,0%, mientras que en 2011 bajó solamente a un 18,4%; teniendo presente que la calidad de esta educación posee bastantes reparos¹⁰, lo cual no explicaría tampoco que la caída se hubiese aminorado por esta razón.

¹⁰ O, al menos, ésta (la calidad) parece estar reservada para un grupo pequeño de instituciones.

La ‘paradoja’ anterior se resuelve al analizar por separado la evolución de las matrículas para universidades y para institutos profesionales (IPs) (gráfico 6,) y después al comparar lo anterior con las tendencias presentes en los premios a educación universitaria y educación técnica (gráfico 7).

GRÁFICO 6: Evolución de la matrícula total por tipo de institución de educación superior

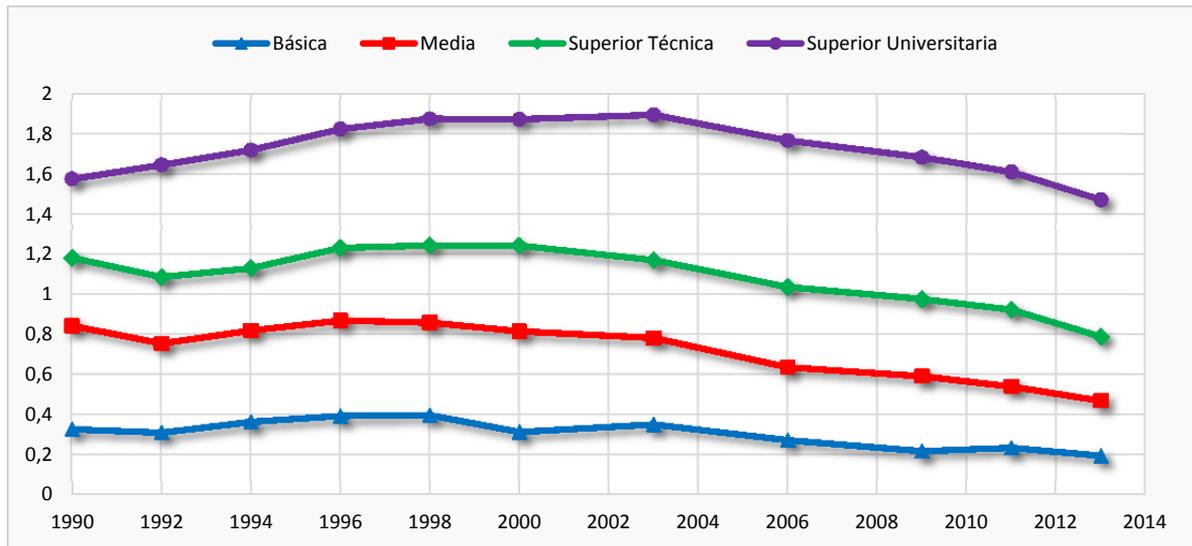


	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
Universitaria	132.624	172.537	183.907	193.334	217.307	223.953	266.365	289.749	330.293	371.298	400.745	436.732	459.484	482.062	503.752	533.819	561.782	602.425
I.P.	17.065	16.648	18.857	24.384	46.736	54.596	62.075	75.634	85.693	91.667	99.561	118.907	128.092	142.370	155.609	176.992	217.388	252.078
Total	149.689	189.185	202.764	217.718	264.043	278.549	328.440	365.383	415.986	462.965	500.306	555.639	587.576	624.432	659.361	710.811	779.170	854.503

Fuente: Elaboración propia en base a datos del Consejo Nacional de Educación (2011).

Podemos extraer de la tabla del gráfico 6 que la matrícula para los IPs creció desde 17.065 a 252.078, casi 15 veces. En tanto, para la educación universitaria aumentó en casi 5 veces, pasando de 132.624 en 1994 a 602.425 en 2011. Por su parte, el gráfico 7 nos muestra los coeficientes asociados a variables dicotómicas que toman el valor 1 si es que la persona posee el nivel educacional correspondiente y 0 si no. Esta metodología es algo diferente a la regresión *spline*, ya que no analiza cuál es el aporte marginal de un año adicional de escolaridad por nivel educacional, sino que intenta hacerse cargo del llamado *sheepskin effect* (Ferrer y Ridell, 2001; Hungerford y Solon, 1996; Jaeger y Page, 1996), el cual señala que el premio a la escolaridad está asociado a tener completo el correspondiente nivel educacional, y no existiría un real aporte marginal por los años intermedios entre la concreción de cada nivel. En suma, si suponemos que la demanda laboral se ha expandido de forma similar tanto para educación superior técnica como para la universitaria, entonces efectivamente notamos que una mayor oferta laboral se condice con una caída en el premio a educación superior ya que así como la matrícula de educación técnica ha crecido más que la universitaria, también el premio a este nivel ha caído en una proporción mayor, como distinguimos del gráfico 7.

GRÁFICO 7: Evolución coeficientes de variables dicotómicas por nivel educacional para hombres y mujeres



Fuente: Elaboración propia a partir de datos de la encuesta Casen, años correspondientes.

Es interesante también analizar en detalle los resultados de las regresiones¹¹ expuestos en el último gráfico. Aquí, la interpretación de los coeficientes es directa, no hay que sumarlos, e indican qué porcentaje más de ingresos gana una persona en el nivel correspondiente en relación a alguien sin educación o con educación básica incompleta. De los resultados podemos notar una tendencia similar a la observada por los coeficientes de la regresión *spline*; no obstante, al desagregar la educación superior en técnica y profesional podemos notar que esta última posee un premio mucho más elevado que todas las demás, incluso el aporte marginal de poseer un título universitario en comparación con uno técnico es considerablemente mayor que el de tener uno técnico en relación a educación media completa.

Ahora si observamos en qué porcentaje cayeron los premios a cada nivel (en comparación con un individuo sin educación) en 2013, con respecto a su valor en 1990, notamos que los de educación básica, media, técnica y universitaria se redujeron en 40,4%, 44,4%, 33,3% y 6,6%; respectivamente. Esto nos permite ver con facilidad por qué no ha disminuido de forma importante la desigualdad en nuestro país. De igual modo, podemos inferir que la razón por la cual no ha caído el premio a la educación universitaria (ahora podemos decir con propiedad que se trata de la educación superior profesional la que ha permanecido principalmente elevada, ya que la técnica ha disminuido), a pesar de una considerable expansión en la oferta laboral, sería debido a que estaría siendo acompañada de una importante expansión también de la demanda por profesionales.

¹¹ Tal como en la regresión *spline* se incorporaron las variables de experiencia y experiencia al cuadrado.

Por último, podemos mencionar que si bien el premio a la educación universitaria se ha mantenido elevado en todo el período, con una pequeña variación entre la primera y última observación, se distingue una tendencia importante a la baja, en términos absolutos, en los últimos 10 años. No obstante, vemos que esta tendencia es compartida por los demás niveles educacionales, por lo que en términos relativos la diferencia sigue siendo elevada entre los más y los menos educados.

Efectos edad, período y cohorte

A continuación mostraremos cómo han afectado en la desigualdad el aumento en la escolaridad media de la población y la disminución de la dispersión de la misma, controlando por efectos edad, período y cohorte. La estimación de estos tres efectos es problemática debido a que la edad, el año del período y el cohorte de nacimiento son linealmente dependientes entre sí¹²; y por otro lado, cualquier análisis que estime sólo dos de las tres variables está sujeto a resultados espurios (Mason et al., 1973). Por ende, antes de mostrar los resultados plantearemos una discusión metodológica acerca de la manera óptima de estimación de dichos efectos.

Una de las soluciones más típicas para abordar este problema es mediante la metodología de coeficientes restringidos de Modelos Lineales Generalizados Convencionales (CGLIM), la cual impone restricciones de igualdad entre dos o más de los coeficientes de tal manera que se dejan como categoría base a las dos variables que se asumen de igual coeficiente. Sin embargo, este método también implica varios problemas¹³ (Yang et al., 2004).

Por otra parte, Sapelli (2011) realizó un análisis similar al que mostramos en este trabajo. No obstante, nos diferenciamos de éste (además de las bases de datos¹⁴) en que él sigue la metodología de Deaton (1997), mientras que aquí usaremos una alternativa que se hace cargo de los problemas de la anterior y también por los planteados por la estimación mediante CGLIM del *Intrinsic Estimator* (Yang et al., 2004; 2008).

La metodología de Deaton para identificar estos tres efectos asume que los del período (los años) no tienen tendencia y que por lo tanto “suman cero”. Sin embargo, como el mismo autor sugiere, este procedimiento es peligroso de emplear cuando hay pocas encuestas donde es

¹² Existe colinealidad perfecta entre edad, cohorte y período debido a que: $\text{edad} + \text{cohorte} = \text{período}$.

¹³ (i) El investigador puede desear observar los efectos individuales de cada uno de los coeficientes de las tres categorías, (ii) se requiere de un fundamento teórico sólido acerca de porque asumir dicha restricción de igualdad, (iii) diferentes elecciones de la identificación de restricciones pueden producir distintos patrones de estimación entre edad, período y cohorte, y (iv) esta elección de qué restricción asumir no es fácil de llevar a cabo porque todos los modelos CGLIM anteriores otorgan una bondad de ajuste idéntica.

¹⁴ Sapelli usa la Encuesta de Ocupación y Desocupación en el Gran Santiago de la Universidad de Chile, mientras en este trabajo se usaron las Casen.

complejo separar tendencias de *shocks* transitorios. Por lo tanto solamente cuando se cuente con suficientes años para separar tendencia y ciclo es posible utilizar esta metodología con confianza (Deaton, 1997). Y dado que en nuestro caso contamos con sólo 11 efectos período, debido a que tenemos 11 encuestas Casen, entonces dicho procedimiento no es el óptimo.

El *intrinsic estimator*, en tanto, es un nuevo método de estimación¹⁵ que entrega una solución única al modelo de regresión lineal que incluye variables dicotómicas por año, cohorte y edad. Logra la identificación del modelo con un mínimo de supuestos.

El análisis llevado a cabo conllevó a la realización de un panel sintético con cohortes de nacimiento que van desde 1926 hasta 1995. Siguiendo a Sapelli (2011), la variable dependiente fue el Gini medido por *celda*. Cada una de estas celdas constituye una observación por cohorte en un año respectivo. Ahora, dado que por construcción debe cumplirse que la cantidad de variables de cohorte sea igual a la suma de variables de edad y período menos uno, para poder aprovechar la información de todas las encuestas Casen (es decir, poder tener 11 variables de período) y tomando en cuenta la frecuencia con que ésta es tomada, se optó por construir tramos de edades (y por ende también de cohortes) de a dos años¹⁶. De esta manera, por ejemplo, una celda (u “observación”) del Gini podría corresponder a éste medido para los que tienen entre 32 y 33 años en el año 1996 (y que corresponden a los cohortes nacidos entre 1964 y 1963). Así obtuvimos una observación para cada tramo de dos cohortes para todos los años en estudio¹⁷. Por último es importante señalar que, para cada año, se incluyó en la muestra sólo a personas entre 18 y 65 años de edad.

En relación a la interpretación a desprender de cada efecto podemos esperar que la de las cohortes nos muestre cómo ha evolucionado la desigualdad para las distintas generaciones. Es de esperar que las variables de cohorte reflejen cambios en los niveles de escolaridad de la población, así como los impactos de cualquier otra política pública que fuese afectando a cada una de las

¹⁵ Su utilidad radica en que remueve la influencia de la matriz de diseño (que está fija por el número de grupos de edad y período y no está relacionada con la observaciones de la variable de resultados) en las estimaciones de los coeficientes. Esta restricción produce estimaciones que poseen propiedades estadísticas deseables, tales como: estimabilidad, insesgamiento, eficiencia relativa y consistencia asintótica (Yang et al., 2004; 2008).

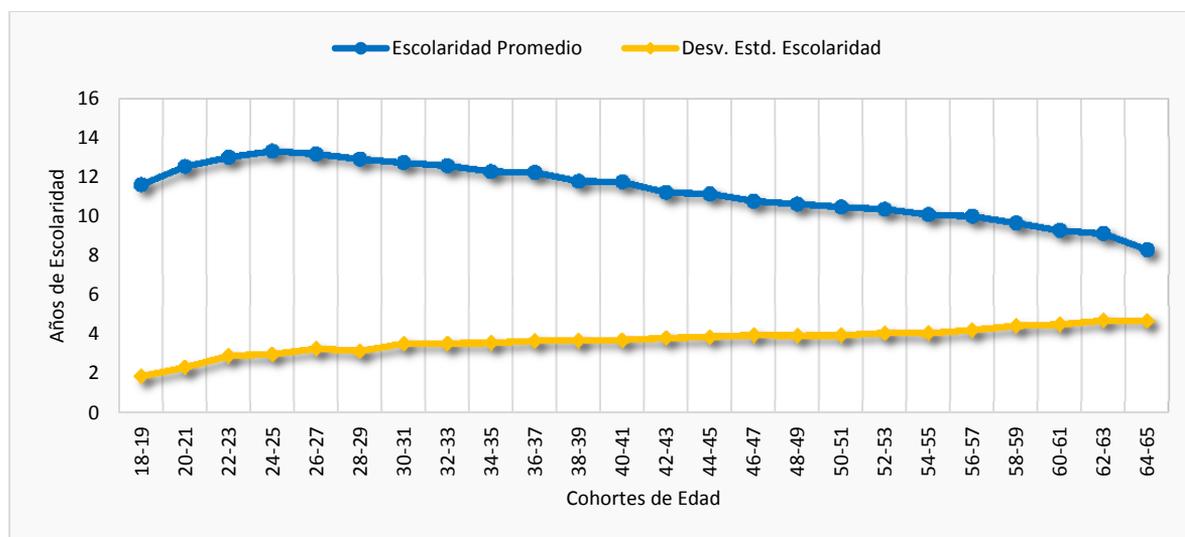
¹⁶ Es importante señalar que dado que a partir del 2003 y hasta 2009 la Casen fue tomada cada 3 años y luego desde 2009 hasta 2013 cada dos (pero durante años impares y no pares como se estaba haciendo hasta el 2000), entonces para construir las observaciones correspondientes a los años 2002, 2004, ..., 2012. Se imputaron sus resultados a partir de los datos ofrecidos por las encuestas. Por ejemplo, para construir el valor de la desviación estándar (DS) de la escolaridad para las personas entre 46 y 47 años del año 2002 (cohortes 1956 y 1955) se realizó lo siguiente: $DS_{2002, j} = DS_{2000, j} + 2(DS_{2003, j} - DS_{2000, j})/3$. Donde j corresponde a las personas entre 46 y 47 años para el año correspondiente. Para los períodos correspondientes a sólo un año posterior a una Casen (como el 2004) se hizo: $DS_{2004, j} = DS_{2003, j} + (DS_{2006, j} - DS_{2004, j})/3$. En tanto que para imputar las observaciones para las celdas correspondientes a los años 2010 y 2012 se calcularon los promedios de cada variable para los años 2009 y 2011, y 2011 y 2013, respectivamente.

¹⁷ A excepción claro de los cohortes más viejos y jóvenes que no se observan en las encuestas más recientes y más antiguas, respectivamente.

nuevas cohortes. Por otro lado, el efecto período muestra cómo varía la desigualdad ante variaciones en el nivel de crecimiento de la economía, lo que puede afectar los niveles de empleo, la inflación, entre otros factores que son propios del año en cuestión que se está analizando. Por último, el efecto edad es de esperar que muestre que la desigualdad aumente conforme se avance en el ciclo de vida, pero a tasas decrecientes (Sapelli, 2011), lo que refleja principalmente cómo van evolucionando las diferencias salariales entre personas de diferentes niveles de calificación.

Antes de presentar los resultados de las estimaciones observemos cómo evolucionaron el promedio y la desviación estándar de la escolaridad para el período analizado. El gráfico 8 nos muestra que las cohortes más jóvenes¹⁸ poseen mayores niveles educativos que las más viejas, lo que habla de una mayor masificación de la educación. De igual forma, las generaciones más recientes poseen menor dispersión en la escolaridad, lo que señala que ésta se ha expandido a más sectores de la población, reduciéndose así las brechas. Un hecho que contrasta con lo que ocurría antes, en donde una pequeña porción de las personas poseía alta escolaridad y un porcentaje grande exhibía bajos niveles de ésta.

GRÁFICO 8: Promedio y Dispersión de la Escolaridad por Cohortes de Edad



Fuente: Elaboración propia en base a datos Encuesta Casen 2013.

En la tabla 3 podemos observar los resultados de las estimaciones usando la metodología del *intrinsic estimator*. Es importante mencionar que el índice de Gini, que es usado como variable dependiente, está medido con respecto al ingreso autónomo de los hogares. Por otro lado, dado

¹⁸ El hecho que el *peak* se encuentre en el tramo de 24-25 años obedece a que en general a esa edad las personas ya han finalizado de acumular capital humano, por lo que no puede atribuirse a que los cohortes entre 1990 y 1995 (personas entre 23 y 18 años, respectivamente) posean menos escolaridad en promedio que quienes, por ejemplo, nacieron entre 1988 y 1989 (25 y 24 años, respectivamente).

que la estimación mediante este método no obliga a dejar una variable como categoría base, entonces para interpretar los coeficientes se debe tener en cuenta que la suma de todos ellos, para cada categoría (edad/año/cohorte), suman cero, por ende el promedio también. De esta manera, cada coeficiente indica el impacto relativo en desigualdad de la edad/año/cohorte en cuestión con respecto al promedio respectivo de la categoría. Por ejemplo, si para el tramo de edades de 58 y 59 años su coeficiente fuera 0,05, ello indica que para las personas de ese tramo el Gini aumentaría en 0,05 con respecto al promedio de las demás edades. Luego, si el coeficiente para individuos entre 60 y 61 años fuera 0,035, significaría que ese grupo posee un Gini mayor que el promedio, pero que existiría una tendencia a la baja de 0,015 en relación al grupo de 58 y 59 años.

TABLA 3: Estimación del Índice de Gini mediante la metodología del *intrinsic estimator*

VARIABLES	(1)	(2)	(3)	(4)	(5)
Promedio Escolaridad		0.0251**		0,0084	0.0744**
Desv. Estdr. Escolaridad			0.0725***	0.0645***	0.2277***
Promedio Esc * D.E. Esc					-0.0169**
Edad 18-19	-0.0226***	-0.0374***	0.0738***	0.0583*	0,0476
Edad 20-21	-0.0247***	-0.0549***	0.0405**	0,0232	0,0072
Edad 22-23	-0.0355***	-0.0710***	0,0016	-0,0144	-0,0225
Edad 24-25	-0.0184***	-0.0579***	0,0034	-0,0123	-0,0159
Edad 26-27	0,0032	-0.0330**	0.0186**	0,0048	0,0045
Edad 28-29	0.0119*	-0,0191	0.0235***	0,0118	0,0123
Edad 30-31	0.0154**	-0,0112	0.0202***	0,0108	0,0140
Edad 32-33	0.0140**	-0,0076	0.0191***	0,0113	0,0134
Edad 34-35	0.0198***	0,0003	0.0219***	0,0152	0,0175
Edad 36-37	0,0088	-0,0054	0.0140**	0,0087	0,0097
Edad 38-39	0,0037	-0,0067	0,0040	0,0005	0,0026
Edad 40-41	0,0082	-0,0005	0,0038	0,0014	0,0045
Edad 42-43	0,0028	0,0008	-0,0037	-0,0036	-0,0004
Edad 44-45	-0,0042	-0,0039	-0.0141**	-0.0129*	-0,0093
Edad 46-47	-0,0024	0,0015	-0.0112*	-0,0089	-0,0049
Edad 48-49	-0.0147**	-0,0047	-0.0298***	-0.0248**	-0.0204**
Edad 50-51	-0,0098	-0,0003	-0.0323***	-0.0267**	-0.0216*
Edad 52-53	0,0010	0.0230**	-0.0201**	-0,0104	-0,0080
Edad 54-55	0,0095	0.0364***	-0.0148*	-0,0031	-0,0016
Edad 56-57	0,0086	0.0402***	-0.0169*	-0,0034	-0,0010
Edad 58-59	0.0167***	0.0554***	-0,0058	0,0097	0,0122
Edad 60-61	-0,0027	0.0356**	-0.0400***	-0,0231	-0,0243

Edad 62-63	-0.0126*	0.0376*	-0.0436***	-0,0233	-0,0232
Edad 64-65	0.0240***	0.0824***	-0,0123	0,0113	0,0077
Período 1990	0.0126***	0.0339***	-0,0003	0,0083	0,0100
Período 1992	0,0052	0.0253***	-0,0020	0,0055	0,0069
Período 1994	0.0109**	0.0272***	0,0060	0,0120	0,0122
Período 1996	0,0019	0.0129**	-0,0016	0,0024	0,0034
Período 1998	0.0153***	0.0205***	0.00895**	0.0114**	0.0115**
Período 2000	0.0162***	0.0151***	0.0152***	0.0150***	0.0154***
Período 2002	0.00930**	0,0044	0.0105**	0.00871*	0.0081*
Período 2004	-0,0023	-0.00926*	0,0026	-0,0003	-0,0019
Período 2006	-0.0192***	-0.0267***	-0.00937*	-0.0130*	-0.0155**
Período 2008	-0.0125***	-0.0254***	-0,0072	-0,0122	-0,0126
Período 2010	-0.0151***	-0.0324***	-0.00921**	-0,0157	-0,0153
Período 2012	-0.0224***	-0.0454***	-0.0135***	-0.0222*	-0.0223*
Cohorte 1926-1927	-0,0170	-0,0039	-0,0113	-0,0075	-0,0193
Cohorte 1928-1929	-0,0008	0,0124	-0,0074	-0,0023	-0,0191
Cohorte 1930-1931	0,0173	0.0344**	0.0211*	0.0264*	0,0206
Cohorte 1932-1933	0,0133	0.0288**	0,0172	0.0219*	0,0189
Cohorte 1934-1935	0,0033	0.0196*	-0,0075	-0,0008	-0,0108
Cohorte 1936-1937	-0,0102	0,0131	-0.0164*	-0,0079	-0,0148
Cohorte 1938-1939	0,0046	0.0276**	-0,0069	0,0021	-0,0045
Cohorte 1940-1941	0,0032	0.0261**	-0,0146	-0,0049	-0,0116
Cohorte 1942-1943	0,0034	0.0244**	-0.0184**	-0,0089	-0,0135
Cohorte 1944-1945	0.0212***	0.0365***	-0,0117	-0,0029	-0,0061
Cohorte 1946-1947	0,0039	0.0175**	-0.0293***	-0,0211	-0,0208
Cohorte 1948-1949	0,0020	0,0128	-0.0284***	-0,0215	-0,0159
Cohorte 1950-1951	0.0114*	0.0197***	-0.0216**	-0,0152	-0,0064
Cohorte 1952-1953	0,0062	0,0069	-0.0248**	-0.0212*	-0,0061
Cohorte 1954-1955	0,0070	0,0015	-0.0152*	-0.0146*	0,0045
Cohorte 1956-1957	0,0034	-0,0024	-0,0050	-0,0060	0,0133
Cohorte 1958-1959	-0,0013	-0,0082	-0,0031	-0,0052	0,0140
Cohorte 1960-1961	-0,0003	-0,0042	0,0007	-0,0007	0,0178
Cohorte 1962-1963	-0,0079	-0.0114*	-0,0056	-0,0070	0,0110
Cohorte 1964-1965	-0,0094	-0.0123*	-0,0069	-0,0081	0,0097
Cohorte 1966-1967	-0,0067	-0,0099	-0,0039	-0,0053	0,0123
Cohorte 1968-1969	-0.00009	-0,0047	0,0038	0,0018	0.0186*
Cohorte 1970-1971	-0,0026	-0,0103	0,0012	-0,0018	0,0142

Cohorte 1972-1973	0,0034	-0,0087	0,0088	0,0042	0,0175
Cohorte 1974-1975	-0,0011	-0.0141*	0,0055	0,0004	0,0113
Cohorte 1976-1977	-0,0036	-0.0164*	0,0027	-0,0023	0,0077
Cohorte 1978-1979	0,0020	-0,0121	0,0082	0,0028	0,0099
Cohorte 1980-1981	0,0005	-0,0148	0,0109	0,0047	0,0072
Cohorte 1982-1983	-0,0027	-0.0196*	0,0127	0,0054	0,0016
Cohorte 1984-1985	0,0002	-0,0191	0.0225**	0,0136	0,0007
Cohorte 1986-1987	-0,0039	-0.0236*	0.0227*	0,0132	-0,0067
Cohorte 1988-1989	-0,0042	-0.0226*	0.0283**	0,0186	-0,0085
Cohorte 1990-1991	-0,0090	-0.0236*	0.0250*	0,0164	-0,0138
Cohorte 1992-1993	-0,0111	-0,0212	0,0255	0,0181	-0,0162
Cohorte 1994-1995	-0,0142	-0,0183	0,0211	0,0159	-0,0168
Constante	0.556***	0.307***	0.274***	0.221**	-0,4342
N	288	288	288	288	288
*** p<0.01, ** p<0.05, * p<0.1					

Fuente: Elaboración propia a partir de las Encuestas Casen 1990-2013.

La columna 1 de la tabla 3 muestra los efectos edad, período y cohorte sin incluir el promedio y la dispersión de la escolaridad. De aquí se puede apreciar que entre los 18 y 25 años existe menos desigualdad en comparación con el resto de las edades, con un nivel de significancia del 1%, lo cual es lógico pues coincide con la época en que los individuos aún se encuentran adquiriendo capital humano en la educación superior. Posteriormente, a lo largo de la adultez, no parece haber diferencias significativas entre las demás edades, salvo algunos casos aislados pero que no muestran un patrón definido. En relación al efecto período notamos que los años 90 e inicios de los 2000 contribuyeron positivamente a incrementar la desigualdad, mientras que a partir del año 2006 se observa una tendencia a la baja en el índice de Gini. Estos resultados coinciden con los de la literatura, pues tal como indica Sapelli (2011) efectivamente en los años donde la economía creció más (durante los 90) hubo mayor contribución de dichos períodos a la desigualdad. Por último vemos que existe una tendencia que muestra que las cohortes más jóvenes van poseyendo menor desigualdad de ingresos; sin embargo, salvo dos tramos de cohortes aislados, ninguno de los demás coeficientes resultó ser significativos al 10%.

Una posible explicación al hecho de que no exista una diferencia estadísticamente significativa en la desigualdad entre las cohortes es que el efecto de éstas incluye otros dos que van en direcciones opuestas. Y es que tal como se observa en el gráfico 8, las cohortes más jóvenes han

ido incrementando sus niveles promedio de escolaridad y al mismo tiempo disminuyendo la desviación estándar de ésta.

Teniendo en cuenta que la entrega de enseñanza básica y media se universalizó a medida que avanzó el siglo XX, los aumentos en el promedio de escolaridad indican en una primera instancia la concreción de este proceso (la masificación de la básica y media), pero principalmente el creciente aumento que ha habido últimamente en el acceso a la educación superior, lo cual ha sido un factor clave en el aumento del promedio de escolaridad de la población. En gran parte del siglo XX sólo un grupo muy reducido tenía acceso a la educación superior, en tanto que el resto de la población no la poseía, lo cual provocaba que las diferencias de ingresos entre la gran mayoría de individuos fuesen bajas. Conforme se ha ido masificando la educación superior, facilitando su acceso a personas de clase media, ello ha permitido que ahora un porcentaje considerable de la población sí tenga acceso a educación universitaria y por ende los mayores ingresos que ahora obtiene dicho grupo más numeroso de personas genera mayores índices de desigualdad¹⁹. Por último, el hecho de que los retornos de la escolaridad crezcan a tasas crecientes es otro factor fundamental (y quizás el más importante) del por qué a mayor escolaridad promedio de la población aumenta más la desigualdad.

Por otra parte, dado que los niveles de escolaridad presentan una suerte de ‘techo’ debido a que, salvo casos excepcionales de individuos que realizan posgrados, la educación universitaria culmina alrededor de los 17 años de escolaridad de la persona, los incrementos en el promedio de la escolaridad estarán asociados a una disminución en la desviación estándar de los niveles educacionales, ya que indica que se provoca un cierre en las brechas de escolaridad de la población. Todos estarían alcanzando el ‘techo’. Esto a su vez se traduciría en un cierre en las brechas de ingresos. En suma la caída en la varianza de la escolaridad provoca que, en conjunto con el aumento en el promedio de ésta, no sólo que un grupo adquiera niveles educacionales mayores, sino que están siendo alcanzados por más personas, lo cual permite reducir el Gini.

Además podemos apreciar que el análisis anterior es coherente con lo que teóricamente uno debiese esperar. La ecuación (1) muestra una ecuación de Mincer (1974) sólo con escolaridad como variable explicativa:

$$(1) \quad \ln Y = \alpha + \beta * S + \mu$$

¹⁹ Pongamos como ejemplo una sociedad de 100 personas que en el año 0 sólo 2 iban a la universidad y ganaban salarios muy elevados, si se mide el Gini de esa economía se vería que es menor que el de la misma en el año 1 en donde ahora 20 personas asisten a la universidad y todas ellas poseen ingresos más elevados que los otros 80. Nótese que en el primer caso los 2 que tenían educación superior actuaban como *outliers*, dejando que el 98% de las personas tuviesen niveles de ingresos muy similares, en tanto en el segundo caso se muestra que existe un grupo de privilegiados y otro sin privilegios.

Donde Y es el ingreso, α es la constante, β el retorno de la escolaridad S y μ es el error. Ahora, asumiendo que la escolaridad no está correlacionada con el error y aplicando varianza a la ecuación (1), tenemos:

$$(2) \quad V(\ln Y) = E(\beta)^2 V(S) + E(S)^2 V(\beta) + V(\mu)$$

De aquí se desprende que las derivadas parciales de la varianza del ingreso con respecto al promedio de la escolaridad y la varianza de ésta son positivas. Por lo tanto, tal como señala Sapelli (2011), desde un punto de vista de la estructura de Mincer acerca del capital humano; durante el siglo XX existieron dos fuerzas actuando en direcciones opuestas que afectaban la distribución de ingresos, el aumento en el promedio de la escolaridad (que aumentó la desigualdad) y la reducción en la dispersión de ésta (que disminuyó la desigualdad).

Entendiendo que cada nueva cohorte tiene mayores niveles educacionales y una menor dispersión de la escolaridad; el análisis anterior explicaría por qué el efecto cohorte no es estadísticamente significativo y distinto de cero en el modelo 1. De hecho en las columnas 2, 3, 4 y 5 de la tabla 3 se confirma lo recién expuesto, pues se incluyen las variables de promedio de la escolaridad y desviación estándar de la misma, medidas para cada *celda*.

En el modelo de la columna 2 se añade sólo el promedio de la escolaridad, en el 3 sólo la desviación estándar de ésta, en el 4 se incluyen ambos y en el 5 se añade también una variable interactiva entre ambas. En los primeros dos vemos que un número importante de tramos de edad y de cohorte ahora sí se tornan estadísticamente significativos y distintos de cero, y tal como era de esperar, los signos de los coeficientes pasan de positivos a negativos y viceversa al controlar sólo por escolaridad promedio o por su dispersión, respectivamente. Esto ocurre ya que los efectos edad y cohorte adquieren el efecto de la variable omitida correspondiente.

En el modelo 4, que incluye tanto el promedio como la dispersión de la escolaridad, podemos apreciar que el primero ahora deja de ser significativo, mientras que la segunda ahora lo es al 1% e indica que una disminución de la desviación estándar en 1 año de escolaridad provoca una disminución en el Gini de 6,45 puntos porcentuales. Por otro lado, ocurre que casi ningún efecto edad o cohorte resulta ser estadísticamente significativo siquiera al 10%, aunque sí se aprecia que las cohortes más antiguas tienen coeficientes que indican menor desigualdad que las más nuevas, aunque no son estadísticamente distintas de cero.

Tal como fue señalado anteriormente el problema del techo está presente en estas estimaciones, en donde a partir de un determinado promedio si éste aumenta entonces sólo puede reducirse la dispersión. Por lo tanto, en el modelo 5 se optó por incluir una variable interactiva entre el promedio y la desviación estándar de la escolaridad con el fin de poder capturar la heterogeneidad del efecto de cada una ante cambios en la otra. Los resultados muestran que nuevamente la variable de dispersión es positiva y significativa (al 1%), pero ahora además la del promedio también pasa a ser significativa (al 5%) manteniendo su signo positivo, en tanto que la variable inter-

activa resultó ser negativa y significativa (al 5%). Esto quiere decir, por ejemplo, que si en los últimos años ha aumentado el promedio de escolaridad y disminuido la dispersión, comparado con el inicio del período analizado, entonces mayor escolaridad provocará mayor desigualdad (por un efecto nivel) pero además habrá aún menos desigualdad en términos relativos con los períodos anteriores pues el efecto interactivo que es negativo estaría siendo multiplicado por una desviación estándar de menor tamaño.

Por último, podemos apreciar que, en general, los efectos período se muestran bastante robustos ante cambios en el modelo y mantienen sus tendencias en comparación al modelo 1.

Una herencia no compartida

Hasta aquí hemos revisado la relación directa que existe entre desigualdad y escolaridad, en donde el aumento del promedio de la última, sumado a los altos retornos que se han mantenido para la educación superior, han sido factores relevantes para explicar por qué ha continuado tan elevada la desigualdad. En tanto que la disminución en la dispersión de la escolaridad ha permitido disminuirla paulatinamente.

Si bien los gobiernos realizan constantes esfuerzos para poder masificar y distribuir más igualitariamente la educación en la población, en especial en aquellos grupos con menores oportunidades, también existen otros factores determinantes que poseen una naturaleza diferente y que pueden limitar seriamente dichos esfuerzos.

Uno de estos problemas de fondo es el llamado emparejamiento selectivo (*assortative mating* en inglés) u homofilia. Éste consiste en la tendencia que poseen los individuos en relacionarse o formar pareja con personas que poseen características similares a las de ellos mismos, tales como coeficiente intelectual, estatura, atractivo físico, color de piel, origen étnico, religión, y por supuesto, educación, entre otras (Becker, 1973).

Este fenómeno tiene fuertes implicancias en explicar las desigualdades de ingresos entre los diversos hogares de un país (Breen y Hald, 2012; Eika et al., 2014; Fernández et al., 2005), pues mientras mayor *assortative mating* (AM) haya más pronunciadas serán las diferencias de ingresos de la sociedad, ya que existirán hogares conformados por parejas de alta educación o de baja educación, generando así dos fuentes de altos ingresos en los primeros y dos fuentes de bajos ingresos en los últimos, lo que contrastaría con un escenario contrafactual en el cual hubiese más familias compuestas por parejas con niveles educacionales heterogéneos²⁰. Es importante señalar que

²⁰ Por ejemplo, pensemos en un país que se compone de dos familias. Donde quienes tienen alta escolaridad ganan 100 y quienes tienen baja ganan 50. Si hubiese homofilia, entonces habría una familia ganando 200 y la otra 100. El contrafactual sin homofilia sería que hubiese dos familias ganando 150. Claramente el último escenario refleja mayor igualdad de ingresos que el primero.

si bien no se puede rechazar *a priori* una causalidad inversa (mayor desigualdad genere mayores índices de homofilia), de todas formas el efecto del emparejamiento selectivo sobre la desigualdad de los ingresos de los hogares es provocado de forma directa (casi por construcción si se quiere), por lo que en el peor de los casos se podría hablar de la existencia de un círculo vicioso de este fenómeno, mas no se puede obviar su impacto.

Otro factor relevante, que eventualmente podría afectar la desigualdad de ingresos y que además menoscaba las oportunidades de distintos grupos de la sociedad, es el de la movilidad educacional intergeneracional (MEI). Mientras menor sea ésta, quiere decir que el nivel educacional alcanzado por los padres es un determinante importante del que pueden llegar a alcanzar sus hijos. Esto implica que si el padre y la madre de alguien tienen baja escolaridad, entonces es probable que su hijo también la tenga, a pesar que pudiese ser alguien con mucho potencial, es decir, es un fenómeno que fomenta el *statu quo* de las posiciones socioeconómicas de tal forma que el efecto de una baja MEI va en la dirección contraria a las políticas de un gobierno en pos de otorgar mayor educación a los sectores con menos recursos, pues si es que los padres poseían baja escolaridad, ello provocaría que les sea difícil a sus hijos acceder a niveles mayores de ésta. No obstante, lo inverso también puede ser cierto²¹ y, por ende, no es claro el efecto que pudiese tener la MEI por sí misma en la desigualdad.

Junto con lo anterior, si una misma sociedad posee baja MEI y elevado AM, entonces ello podría provocar mayor desigualdad aun de lo que podría generar un alto nivel de homofilia por sí sólo. Ello dependerá del tipo de relación existente entre la propensión a poseer cierto nivel educacional y el nivel de escolaridad de los padres, es decir, si ésta es linealmente positiva o creciente a tasas decrecientes o creciente a tasa crecientes. Por ejemplo, la propensión de una persona a tener educación superior pudiera crecer más que proporcionalmente al tener sus dos padres con educación superior, versus un escenario en donde sólo uno de ellos la poseyera, en cuyo caso la relación sería creciente a tasas crecientes.

Sin embargo un caso como el anterior suele ser descartado en la literatura, ya que para analizar la MEI, por ejemplo, sólo se observa al progenitor con el mayor nivel educacional (Gaviria, 2007; González y Mackenna, 2015), omitiendo la información del otro. Esto asume implícitamente que el aporte marginal del padre con menor escolaridad no debiese influir en el nivel educacional del individuo. En caso que este supuesto no se cumpla, querría decir que en un escenario de baja MEI y bajo AM, disminuiría el porcentaje de gente con educación superior y sin educación, y aumentaría la proporción en niveles intermedios, como gente con educación básica y media, haciendo más leptokúrtica la distribución. Ahora, si este fenómeno ocurre de manera similar en ambas colas de esta última, entonces al pasar gente de un nivel educacional al siguiente, debe tenerse en

²¹ Es decir, si existe una baja MEI, ello permite que mientras más padres posean, por ejemplo, educación superior, entonces sus hijos serán más propensos a tenerla.

cuenta que las ganancias en ingresos, producidas por el aumento de escolaridad de quienes pasan de no tener educación a tener básica, son sustancialmente menores de lo que dejarían de percibir quienes pasan de tener educación superior a educación media. Si este es el caso, entonces la distribución de ingresos, por su parte, se tornaría más simétrica (disminuyendo la asimetría positiva que posee en el caso chileno), provocándose que disminuya la desigualdad. Al mismo tiempo que, dado los parámetros de la distribución para Chile, disminuirían los ingresos promedio.

En las siguientes secciones develaremos si se cumple o no el supuesto en cuestión. De no cumplirse, querría decir que en el agregado la segregación social pudiese tener costos adicionales en términos de desigualdad, pero ganancias en términos de escolaridad²² e ingresos medios, mediante el canal del emparejamiento selectivo de las parejas.

A continuación revisaremos el nivel de AM y MEI en Chile, y lo pondremos en perspectiva con algunas mediciones de algunos países desarrollados.

Assortative Mating en Chile

Como adelantamos, la homofilia puede terminar siendo un medio de reproducción intergeneracional de la desigualdad, donde mientras mayor sea la primera es más probable que se reproduzca la misma estructura social. A partir de ello se hace necesario poder cuantificar este fenómeno. Existen diferentes formas de medir el emparejamiento selectivo. Nosotros presentaremos dos de éstas. La primera, que es la más directa y simple, es mediante la correlación de años de escolaridad entre las personas que forman pareja²³. En el gráfico 9 podemos observar la evolución de ésta desde 1990 hasta 2013.

Podemos apreciar nuevamente que los datos siguen una tendencia similar a la de la desigualdad y a la evolución de los retornos de la educación superior, con un alza a inicios de los 2000 y una caída hacia finales del período observado²⁴. Con variaciones de pequeña magnitud, claro está. Estas cifras, si bien han mostrado un leve descenso en el período analizado (pasando de 0,75 en 1990 a 0,72 en 2013), son altas en comparación con otros países. Para Estados Unidos, Warren (1966) encontró valores cercanos a 0,60, en tanto que Becker (1973) mostró una correlación de 0,53 para blancos y 0,56 para negros. Latinoamérica Valenzuela y Duryea (2011) presentan correlaciones de 0,66 para Argentina en 2002, 0,65 para Brasil y Costa Rica en 2003, 0,67 para

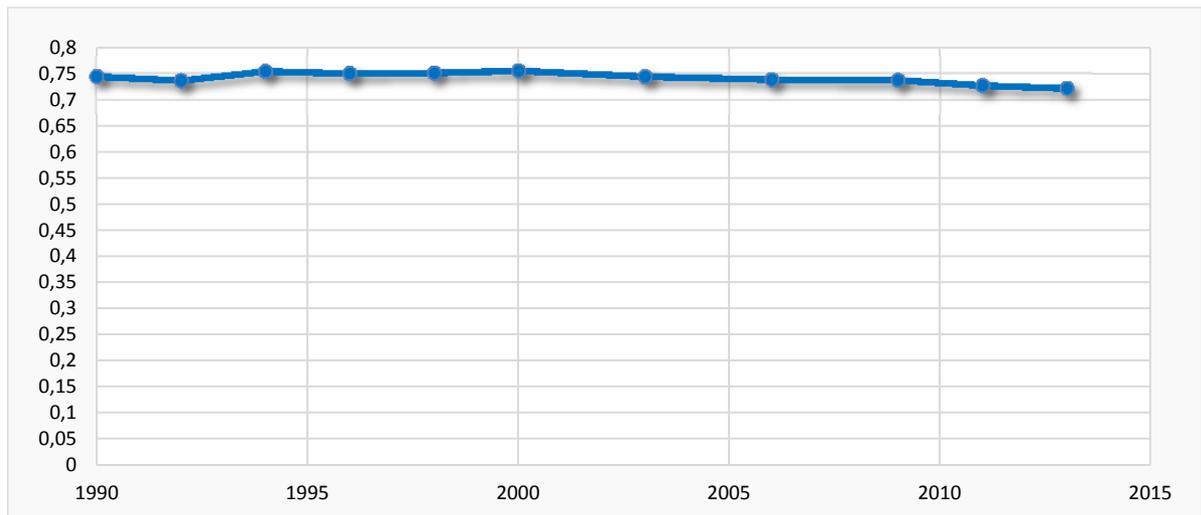
²² Ello podría traducirse en que el capital cultural de los hogares no crecería a tasas decrecientes, sino que el aporte de ambos padres es igualmente trascendente para la educación de los hijos, y no sólo el del más educado.

²³ La medición realizada en este trabajo es aplicada sólo a núcleos dentro de un hogar cuyo jefe de hogar esté casado o tenga pareja.

²⁴ Es importante tener presente para el análisis que este fenómeno coincide con la reforma constitucional del año 2003 que establece la educación media obligatoria, lo cual es de esperar que sesgue los resultados de las correlaciones al alza a partir de ese año, pues ello tendería a haber reducido la varianza de la escolaridad.

Uruguay en el mismo año, y 0,70 para México en 2004. Buss (1985) señala que variables como la edad poseen una correlación entre 0,70 y 0,90, y que valores mayores a 0,50 ya pueden considerarse como elevados. Por lo tanto puede decirse con seguridad que la correlación de los años de escolaridad de las parejas en Chile es alta.

GRÁFICO 9: Correlación de escolaridad de la pareja



Fuente: Elaboración propia a partir de datos de la Encuesta Casen, años respectivos.

Valenzuela y Duryea (2011) también observan estas correlaciones para Chile en áreas urbanas en los años 1992 y 2003 y registran correlaciones de 0,717 y 0,714²⁵, respectivamente. No obstante, tal como señala Torche (2010), existen muchas limitaciones al usar una simple medida agregada de emparejamiento selectivo (como la correlación de años de escolaridad) para capturar la relación entre variación en homofilia y su relación con desigualdad. Es por ello que a continuación presentamos otra metodología de cálculo de emparejamiento selectivo que cumple con no ser una medida agregada, pero que no se ha aplicado anteriormente para Chile²⁶ y que nos ofrece novedosos resultados que nos permitirán extender el análisis sobre este tópico y sus implicancias.

Una manera alternativa de medir el grado de AM de una sociedad es mediante la forma realizada por Eika et al. (2014), quienes miden la clasificación marital entre los niveles educacionales i y j como la probabilidad observada de que un hombre con nivel educacional j esté casado (o emparejado) con una mujer con nivel educacional i , relativo a la probabilidad bajo emparejamiento aleatorio con respecto a la educación:

$$(3) \quad C_{ij} = \frac{\Pr(\text{Mujer}=i \cap \text{Hombre}=j)}{\Pr(\text{Mujer}=i) * \Pr(\text{Hombre}=j)}$$

²⁵ Resultados que coinciden con los nuestros al restringir la muestra sólo a la zona urbana.

²⁶ Difiere con la empleada por Torche (2010).

Donde el numerador corresponde a la probabilidad conjunta observada de que una mujer con nivel educacional i esté emparejada con un hombre de nivel j , y el denominador corresponde al producto entre las probabilidades marginales observadas de la mujer poseyendo nivel i y el hombre nivel j .

Si hombres y mujeres con igual nivel de educación se casan (o emparejan) más frecuentemente de lo que sería esperado bajo un patrón aleatorio en términos de educación, ello quiere decir que habría un AM positivo. Esto se reflejaría en un parámetro de clasificación marital c_{ij} mayor a 1 cuando $i = j$; y viceversa.

La tabla 4 muestra los valores de estos parámetros para cada combinación de nivel educacional para hombres y mujeres que forman una pareja²⁷, para los años 1990, 2000 y 2013:

TABLA 4: Evolución del *Assortative Mating* por nivel educacional

1990	Educación del Esposo			
Educación de la Esposa	Sin Educación	Básica	Media	Superior
Sin Educación	2,40	0,81	0,19	0,07
Básica	0,65	1,61	0,76	0,23
Media	0,14	0,51	2,21	1,93
Superior	0,09	0,27	1,15	6,22
2000	Educación del Esposo			
Educación de la Esposa	Sin Educación	Básica	Media	Superior
Sin Educación	2,98	0,87	0,21	0,03
Básica	0,78	1,70	0,76	0,23
Media	0,18	0,66	1,88	1,09
Superior	0,02	0,19	0,89	4,66
2013	Educación del Esposo			
Educación de la Esposa	Sin Educación	Básica	Media	Superior
Sin Educación	3,41	1,06	0,34	0,05
Básica	1,02	1,82	0,76	0,16
Media	0,25	0,74	1,61	0,90
Superior	0,05	0,17	0,75	3,61

Fuente: Elaboración propia a partir de datos de la Encuesta Casen, años respectivos.

²⁷ Es importante mencionar que, para poder identificar a las parejas existentes, se consideró sólo a los núcleos de hogares que reportaban estar casados o estar en pareja. Es decir el análisis no incluye a posibles parejas compuestas por dos integrantes de núcleos diferentes. Lo cual también es positivo, pues toma en cuenta sólo a relaciones ya establecidas, es decir, se controla por aquellas que pudiesen tener horizontes de corto plazo que no llegan a determinar la composición de ingresos de hogares.

Podemos comparar estos resultados con los de Eika et al. (2014) que se muestran en la tabla 5. Ellos aplicaron el mismo análisis²⁸ para EE.UU. y Noruega para los años 1980 y 2007. De ese trabajo podemos notar que ambos países tienen un AM positivo, aunque considerablemente menor que Chile (a excepción de EE.UU. para la categoría de “educación media incompleta”²⁹, aunque no es directamente comparable con las categorías analizadas para nuestro país). En especial merece observarse con detención la categoría $i = j =$ “educación superior”. Ésta muestra un parámetro c_{ij} cercano a 2 para EE.UU. y Noruega, mientras que en Chile dicho valor es casi el doble. Esta cifra puede tener consecuencias muy grandes a la hora de mirar la desigualdad de nuestro país, en especial tomando en cuenta el alto retorno a la educación superior, relativo a otras naciones. Sin embargo es interesante notar que para el período de 27 años estudiado por los autores, en comparación con el de 23 que presentamos en la tabla 4, existe una tendencia muy similar en el AM: se observa una reducción para los índices de clasificaciones de mayor escolaridad (“media” y “superior”) y un incremento en los de menores niveles educacionales (“sin educación” y “básica”). Tal como fue señalado anteriormente en un pie de página, es posible que esta tendencia se relacione con la política que establece la educación media obligatoria.

TABLA 5: *Assortative Mating* en EE.UU. y Noruega por nivel educacional en los años 1980 y 2007

Husbands' Education \ Wives Education	U.S.				Norway			
	1980				1980			
	No high school degree	High school graduates	Some college	College graduates	No high school degree	High school graduates	Some college	College graduates
No high school degree	2.74	0.80	0.38	0.09	1.60	0.88	0.44	0.16
High school graduates	0.83	1.46	1.02	0.50	0.71	1.21	1.21	0.82
Some college	0.29	0.67	1.75	1.54	0.14	0.67	2.48	3.72
College graduates	0.08	0.30	0.83	2.98	0.19	0.66	1.71	4.35
	2007				2007			
No high school degree	6.37	0.97	0.34	0.09	1.96	1.15	0.62	0.27
High school graduates	1.12	1.89	0.83	0.35	1.10	1.25	0.91	0.49
Some college	0.46	0.92	1.68	0.72	0.56	0.78	1.63	1.34
College graduates	0.12	0.36	0.75	1.94	0.35	0.63	1.17	2.07

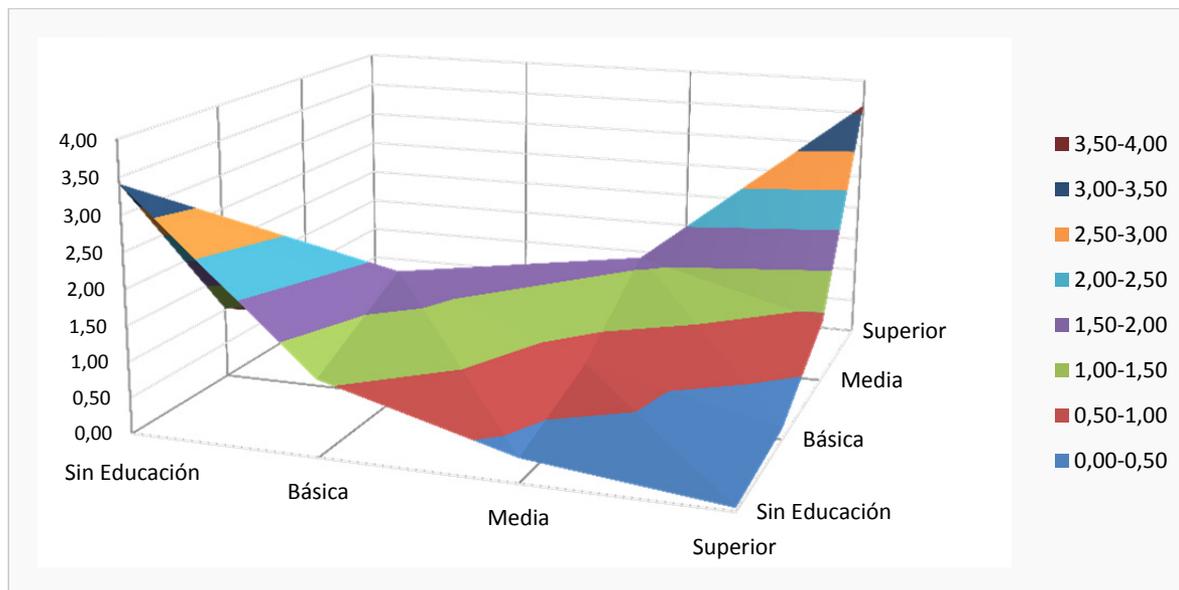
Fuente: Eika et al., (2014).

²⁸ Las categorías de los niveles educacionales es en lo único en que levemente difiere el análisis. Ellos utilizan “educación media incompleta”, “educación media completa”, “educación superior incompleta” y “educación superior completa”.

²⁹ “No high school degree”.

En el gráfico 10 podemos observar con mayor claridad la homofilia existente para el año 2013 en Chile, donde destacan los *peaks* para las categorías donde $i = j = \text{Sin Educación}$ e $i = j = \text{Educación Superior}$.

GRÁFICO 10: Assortative Mating por niveles educacionales para el año 2013



Fuente: Elaboración propia a partir de datos de la Encuesta Casen 2013.

Para poder explicar esta tendencia existen dos hipótesis. Por un lado puede que se haya incrementado el nivel de homofilia o *assortative mating*, o puede deberse a cambios en la distribución de educación. Observando los resultados expuestos en la tabla 6 puede apreciarse que es probable que no se deba a la primera hipótesis, puesto que el nivel de AM ponderado³⁰ se ha mantenido bastante similar en el período analizado, tal como ocurre con EE.UU. y Noruega, aunque se aprecia una leve disminución por parte de nuestro país y un leve aumento por parte de los últimos. En síntesis, la tendencia a que disminuya el AM de los individuos más educados y aumente el de los menos se debería a cambios en la distribución de educación de hombres y mujeres, tal como lo confirman Eika et al. (2014).

³⁰ Calculado a partir de sumar la ponderación entre $c_{i=j}$ y la cantidad de personas presentes en dicha categoría y posteriormente el resultado final dividirlo por el total de individuos que hay en la diagonal.

TABLA 6: Medida agregada de *assortative mating*

	EE.UU.		Noruega		Chile	
	1980	2007	1980	2007	1990	2013
Promedio ponderado de parámetros de clasificación (C_{ij}) de la diagonal	1,93	1,97	1,45	1,55	2,37	2,27

Fuente: Eika et al. (2014) para Estados Unidos y Noruega; elaboración propia a partir de la Encuesta Casen, años respectivos, para Chile.

Es importante resaltar que, dada la forma en que están contruidos los parámetros de clasificación c_{ij} , dicha disminución en los niveles superiores (y alza en los más bajos) no ocurre sólo porque conforme surgen nuevas cohortes éstas poseen promedios de escolaridad mayores (lo cual incrementa el numerador de la c_{ij}), sino que toma en cuenta que es más probable encontrar tanto hombres como mujeres en dichos niveles (el denominador está compuesto por el producto de las probabilidades marginales). De esta manera podemos indicar con seguridad que esta medida de AM nos muestra una tendencia real de lo que está ocurriendo en relación a la homofilia, ya que se hace cargo de la evolución natural de los niveles de educación.

A pesar de los elevados índices de emparejamiento selectivo que posee nuestro país, y que son un diagnóstico interesante acerca del por qué no ha progresado más rápido el proceso de disminución de la desigualdad, Chile ha mostrado una disminución importante en el período analizado para algunas de las clasificaciones. Disminuyó su AM para la de superior-superior en un 42%, mejor al de EE.UU. de 34,9%, aunque un poco inferior al de Noruega de 52,4%; pero lográndolo en un período más breve. Sin embargo es preocupante el aumento en el parámetro de clasificación c_{ij} de parejas con bajos niveles educacionales, más aún si es que en el país existen bajos niveles de movilidad educacional intergeneracional. A continuación en la siguiente sección revisaremos este tema.

Movilidad Educacional Intergeneracional en Chile

Existen diversas formas de medir la movilidad intergeneracional de los individuos, donde quizás las más usadas sean la de ingresos y la educacional. Para Chile existen estudios relativos principalmente a la primera (Núñez y Miranda, 2006; Torche, 2005a), sin embargo es más escasa para la segunda.

Para analizar la MEI estaremos por MCO una regresión donde la variable dependiente será el nivel educacional³¹ de la persona y como variable explicativa relevante estará el nivel educacional más alto alcanzado por los padres. Esta metodología también es aplicada por Gaviria (2007)

³¹ Los niveles utilizados son: sin educación, básica, media, técnica nivel superior y universitaria nivel superior.

y González y Mackena (2015)³². También se agregaron algunos controles, que se pueden observar tabla 7, la cual muestra los resultados de la estimación.

TABLA 7: Estimación del nivel educacional de cada individuo con respecto a la de sus padres

Variables	2006	2009	2011	2013
Nivel educacional de los padres	0,4706	0,5160	0,5196	0,5094
Sexo (Hombre = 1)	0,1092	0,0976	0,1060	0,0871
Edad	0,0180	0,0328	0,0304	0,0357
Edad*Edad	-0,0003	-0,0005	-0,0004	-0,0005
Tiene pareja (Sí = 1)	0,0404	0,0583	0,0861	0,0737
Zona (Urbano = 1)	0,4899	0,4912	0,4996	0,4865
Número de personas en el hogar	-0,0522	-0,0553	-0,0494	-0,0589
N	98.480	82.897	65.074	68.701
R2	0,346	0,357	0,355	0,364

Fuente: Elaboración propia a partir de la Encuesta Casen, años respectivos.

Nota: La variable dependiente es el nivel educacional del individuo. Todos los coeficientes son significativos al 0,1%.

Según Gaviria (2007), para países desarrollados, la correlación entre el nivel de escolaridad de la persona y el de sus padres se ubica entre 0,3 y 0,2; en tanto que en sus estimaciones para Latinoamérica, encuentra un coeficiente de 0,442. Por su parte, González y Mackena (2015)³³ encuentran un coeficiente de 0,622 para Chile (sin incluir controles en su estimación). Replicando la estimación con los datos de la Casen 2013, y sin incluir controles, se pudo observar que el coeficiente alcanza un valor de 0,587, el cual es un valor bastante similar al hallado por estos últimos autores.

A partir de las estimaciones³⁴ presentadas en la tabla 7, observamos que a fines de la década de los 2000 disminuyó la MEI (aumentó el coeficiente), posteriormente en 2013 se aprecia un pequeño aumento de la movilidad, pero que se mantiene en niveles inferiores a los registrados en la primera observación en 2006.

³² Una diferencia entre la metodología aplicada en este trabajo con respecto a la de estos autores es que aquí el análisis se hizo en relación al nivel educacional de los padres y no con respecto a los años de escolaridad como lo hicieron ellos. La razón de usar niveles es debido a que los individuos recuerdan (y por ende reportan) con mayor precisión su escolaridad en niveles en vez de años específicos.

³³ Usando datos de la Encuesta del Centro de Estudios Públicos del año 2014.

³⁴ Las estimaciones presentadas en este trabajo fueron hechas desde el año 2006 en adelante, debido a que sólo desde ese año se comenzó a preguntar por la educación de los padres del entrevistado en la Encuesta Casen.

Estos resultados sugieren que, en una perspectiva internacional, Chile posee una baja MEI, lo cual es un síntoma preocupante que provoca que existan limitadas posibilidades de alcanzar niveles educacionales elevados para alguien con padres de baja escolaridad, y viceversa. Ello no sólo puede considerarse como un escenario que menoscaba la meritocracia, sino que también es ineficiente pues impide que personas con talento salgan adelante. Por otro lado, además de la desigualdad de oportunidades generada entre distintos sectores socioeconómicos de la población, esto también genera un escenario adverso para poder disminuir las desigualdades de ingresos de la sociedad al sumarlo con los elevados índices de AM que posee el país.

Es relevante destacar que el análisis de la MEI puede enfrentar un problema de “techo”, lo cual se produciría porque, asumiendo que cada nueva cohorte de personas sería más educada que la anterior, entonces en un escenario extremo, hipotético, habrá una sociedad compuesta sólo por individuos con educación superior y los hijos de esas personas también lo serían, por lo que no existiría movilidad. Esto podría afectar las estimaciones, en particular se sobrestimarían los coeficientes de correlación conforme pasen los años. En años posteriores tendería a alcanzarse el techo educacional por los padres y por los hijos, generando una mayor correlación de la que sólo reflejase la MEI por sí misma. Esto significaría que el hecho de que la correlación de países desarrollados sea menor que la nuestra, a pesar del problema del techo (el cual ellos estarían más próximos a alcanzar), estaría subestimando las diferencias en MEI que tenemos con ellos, haciendo aún más preocupante la situación chilena.

Simulando un contrafactual

A continuación se presenta un análisis novedoso que, al menos hasta donde alcanza mi conocimiento, no ha sido realizado previamente en la literatura, acerca de cómo cambiaría el índice de Gini³⁵ si se simulase un escenario en el cual no hubiese existido emparejamiento selectivo u homofilia (con respecto a la escolaridad) por parte de los padres de los individuos de la muestra. Ello nos permitirá estimar un nuevo nivel de escolaridad para cada persona, con el cual estimaremos nuevos ingresos. A partir de este tratamiento, generamos una nueva distribución desde la cual compararemos los niveles de desigualdad. Los resultados de interés finalmente será la nueva distribución por niveles educacionales, el nuevo ingreso promedio, el Gini y el ratio 10/10.

Esta simulación resulta ser similar en espíritu a la realizada por Greenwood et al. (2014), pues también ellos intentan comparar la desigualdad actual con un escenario de emparejamiento aleatorio, sin embargo las metodologías entre ese y este trabajo son diferentes. En particular, destaca en el presente análisis el hecho de hacer la simulación con respecto a los padres de los actua-

³⁵ En esta parte se trabajara con el Gini a nivel individual y no de hogares, más adelante se comprenderán las razones de ello.

les individuos y no sobre los individuos mismos, porque ello nos permitirá dos cosas: (i) podremos hacer un análisis representativo de toda la población, pues todos los individuos necesariamente tuvieron padres y por ende habrá una expresión de AM a partir de cada persona. Si, en cambio, intentásemos simular cómo sería una situación sin homofilia para la generación actual (parejas actuales), como lo hacen Greenwood et al. (2014), ello tendría el problema que otorgaría resultados sesgados, relativos al efecto que provocaría el *assortative mating* en la distribución de ingresos y de escolaridad, debido a que no todas las personas se encuentran en pareja. De esta forma sólo podríamos observar cómo cambian los ingresos del hogar de aquellos núcleos que actualmente se constituyen de parejas, dejando sin representación a solteros, viudos, separados, etc., ya sea que vivan solos o con alguien más. Es decir, podremos observar el efecto que tendría este escenario sin AM sobre toda la sociedad y no sólo sobre una parte de ésta.

Además hacer la simulación con respecto a los padres nos permite (ii) observar la validez de la hipótesis de que la propensión de los individuos a tener mayores niveles educacionales crece a una tasa marcadamente decreciente con el nivel educacional del otro padre. Ésta se validaría en caso que evidenciemos un aumento de la gente con mayores niveles educacionales, ya que en dicho caso querría decir que sería más beneficioso en el agregado *distribuir* a los padres más educados con parejas menos educadas, puesto que en el agregado se ganaría mayor propensión a tener más escolaridad, si es que un padre educado pasa, de estar emparejado con alguien igualmente educado, a alguien con menor educación.

Antes de explicar en detalle el procedimiento de la simulación, es importante señalar que tal como ya vimos, incluso países desarrollados poseen aunque sea un poco de AM, por lo que estos resultados simplemente muestran un escenario o nivel de desigualdad potencial en el que nuestro país podría haber estado en caso de haber sido una sociedad menos segregada y más inclusiva. De esta forma se busca mostrar la enorme relevancia de este tópico.

Metodología

Para entender con mayor claridad la metodología utilizada en este trabajo, la pondremos en contraste con el análisis de Greenwood et al. (2014), en donde señalan que g_k denota la fracción de hogares que son del tipo k , donde los puede haber compuestos por solteros de distintos niveles educacionales y por parejas de distintas combinaciones de dichos niveles, y w_k corresponde al ingreso de dicho tipo de hogar. Luego, si P representa el conjunto de parejas y S el de solteros, entonces, el experimento de estos autores se trata de reemplazar los $\{g_k\}$ observados, para cada $k \in P$, con el conjunto que se obtendría de un emparejamiento aleatorio, denotado por $\{\tilde{g}_k\}$, $\forall k \in P$. Así el Gini contrafactual queda definido por $Gini(\{g'_k\}, \{w_k\})$, donde $\{g'_k\} \equiv \{\tilde{g}_k\}_P \cup \{g_k\}_S$.

Dado que nuestro objetivo es dimensionar el alcance y las consecuencias que tiene el emparejamiento selectivo sobre toda la sociedad (además de también observar cómo intermedia entre el nivel de desigualdad y la escolaridad de la población), entonces el análisis realizado por los anteriores autores es limitado, pues $|Gini(\{g'_k\}, \{w_k\}) - Gini(\{g_k\}, \{w_k\})|$ estaría siendo subestimado debido a que el subconjunto $\{g_k\}_S$ permanece inalterado.

Por lo tanto, dado que por definición $\{g_k\}_S$ permanecerá invariante bajo el ejercicio de aleatorizar las parejas, entonces si mantenemos constante los $\{g_k\}$ observados, podemos proceder a estimar un nuevo conjunto de ingresos $\{\widehat{w}_k\}$, tal que sabemos que $w_k(esc_{kl}(escp_{klm}))$; donde esc_{kl} corresponde a la escolaridad del individuo l del tipo de hogar k y $escp_{klm}$ es la escolaridad del progenitor m (padre y madre) del individuo l del tipo de hogar k . Luego, si definimos a f_k como la fracción de hogares de los padres de los individuos que son del tipo k , tal que $\{f_k\} \equiv \{f_k\}_P$. De esta forma, si aplicamos el emparejamiento aleatorio sobre los padres, obtendremos $\{\tilde{f}_k\} \forall k$. Con ello obtendremos un nuevo conjunto $\{\widehat{escp}_{klm}\}$ lo que nos permitirá estimar nuevos niveles de escolaridad y salarios $\widehat{w}_k(\widehat{esc}_{kl}) \forall k, l$. Con lo que el Gini contrafactual queda definido por $Gini(\{g_k\}, \{\widehat{w}_k\})$, donde $\{\widehat{w}_k\} \equiv \{\widehat{w}_k\}_P \cup \{\widehat{w}_k\}_S$.

Para realizar el ejercicio descrito se utilizaron los datos de la Encuesta Casen del año 2013. Ésta incluye preguntas acerca de la educación de los padres de las personas encuestadas³⁶. No obstante sólo 68.947 de un total de 218.491 individuos reportaban el nivel educacional de ambos de sus progenitores. Ello no hubiese sido un inconveniente de no ser porque la distribución de ingresos y de niveles educacionales de los individuos era considerablemente distinta entre la submuestra y la muestra total. Ello puede verse en la tabla 8.

Se aprecia con claridad los contrastes existentes entre la submuestra y la muestra total. Es presumible que las diferencias en la distribución de ingresos sean producto de las presentes en la distribución por niveles educacionales. Se observa que existe una sobrerrepresentación de las personas con mayor educación y una subrepresentación de quienes no poseen educación y educación básica incompleta, en la submuestra. Por lo tanto se procedió a realizar imputaciones.

³⁶ Nótese que las preguntas relativas al nivel educacional de los padres no fueron realizadas solamente al jefe de hogar.

TABLA 8: Comparación de distribución de ingresos y por niveles educacionales de muestra original y muestra restringida

		Valores Originales Muestra	Muestra Restringida
Ingreso de la Ocupación Principal	Gini	0,48	0,51
	Ratio 10/10	6,25	7,333
	Promedio	501.020	592.710
	Desviación Estándar	775.985	935.981
	Mínimo	2.270	2.270
	Máximo	43.500.000	43.500.000
	Sin Educación (Frec)	5.417.480	839.650
	Sin Educación (%)	31,49%	15,21%
	Ed. Básica (Frec.)	4.198.088	1.395.572
	Ed. Básica (%)	24,40%	25,29%
	Ed. Media (Frec.)	5.463.920	2.093.977
	Ed. Media (%)	31,76%	37,94%
	Ed. Superior (Frec.)	2.125.105	1.189.654
	Ed. Superior (%)	12,35%	21,56%

Dicho proceso fue llevado a cabo minuciosamente, procurando conservar la misma distribución del nivel educacional de los individuos. Para lograrlo se procedió de la siguiente forma: primero se analizó la distribución de las clasificaciones educacionales c_{ij} (las mismas expuestas en la sección anterior) de los padres de los encuestados, por cada nivel educacional de estos últimos. Es decir, por ejemplo, se tomaron a todos los individuos con educación media y se observó el porcentaje de personas que tenían sus padres con cada nivel de clasificación ij . Luego, para todos los individuos que también tenían educación media pero que no reportaban el nivel educacional de alguno de sus padres, se procedió a imputárseles aleatoriamente un nivel de escolaridad de sus progenitores de tal forma que se conservasen los porcentajes de representatividad de cada clasificación ij de acuerdo a los niveles originales; es decir, se conservaron las distribuciones intra nivel educacional³⁷. Este procedimiento fue realizado por separado para individuos que reportaban sólo la escolaridad del padre, sólo la de la madre y para quienes no reportaban ninguna, conservando

³⁷ Retomando el ejemplo mencionado, podía tenerse que a partir de los datos observados se veía que un 6% de las personas con educación media poseían un padre sin educación y una madre con educación básica; mientras que para alguien con educación superior dicho porcentaje correspondía sólo al 4%. De esta forma, si había un individuo con educación media pero que no reportaba la educación de alguno de sus progenitores, entonces tendría un 6% de probabilidades de asignársele en el nivel de clasificación en cuestión. Mientras que en las mismas condiciones alguien con educación superior tenía un 4% de probabilidades de ser asignado de dicha manera.

en todos los casos, y también a nivel global, las proporciones originales por nivel educacional de los encuestados.

De haberse optado sólo por mantener las proporciones globales de las clasificaciones ij originales, hubiese habido mayores probabilidades de sesgar los datos en las imputaciones, ya que dichas proporciones provenían de una submuestra que sobrerrepresentaba los niveles altos de educación y subrepresentaba los bajos. En tanto, al mantener las proporciones *dentro de cada nivel educacional* de los encuestados, entonces el supuesto de identificación es menos restrictivo, pues se asume que las proporciones de quienes reportaron la educación de ambos padres serán las mismas que las proporciones poblacionales, pero controlando por el nivel educacional de cada individuo. De todas formas, más adelante se presentarán simulaciones con y sin imputaciones.

Posterior a las imputaciones se pudo pasar a analizar cómo se distribuían las parejas de padres clasificados por sus niveles educacionales, lo que se puede observar en la tabla 9. Con ello se calcularon los parámetros de clasificación c_{ij} como medida de AM. Estos últimos resultados se pueden ver en la tabla 10. Es importante mencionar, que conservar las proporciones intra nivel educacional, no garantiza que se conserven los mismos niveles de *assortative mating* globales para los padres. En efecto, cambian. En las tablas A.1 y A.2, en Anexos, se pueden observar las mismas tablas pero para la muestra restringida sin imputaciones.

TABLA 9: Distribución porcentual por nivel educacional de la educación de los padres de los encuestados³⁸, post imputaciones

	Educación del Padre				
Educación de la Madre	Sin Educación	Básica	Media	Superior	Total
Sin Educación	53,20%	6,68%	1,99%	0,24%	62,10%
Básica	4,67%	14,63%	3,28%	0,60%	23,19%
Media	1,19%	2,13%	7,93%	1,25%	12,50%
Superior	0,07%	0,21%	0,55%	1,37%	2,20%
Total	59,13%	23,65%	13,75%	3,47%	100,00%

Fuente: Elaboración propia en base a Encuesta Casen 2013.

³⁸ Estos datos no incluyen factores de expansión, debido a que su utilidad se centra en la construcción del escenario contrafactual.

TABLA 10: *Assortative Mating* de los padres de los encuestados³⁹, post imputaciones

	Educación del Padre			
Educación de la Madre	Sin Educación	Básica	Media	Superior
Sin Educación	1,45	0,45	0,23	0,11
Básica	0,34	2,67	1,03	0,75
Media	0,16	0,72	4,61	2,89
Superior	0,05	0,41	1,81	18,00

Fuente: Elaboración propia en base a Encuesta Casen 2013.

Tal como indicamos con anterioridad, si $c_{ij} = 1$ indica que no existe AM, es decir, hombres y mujeres de igual nivel de educación no están ni más, ni menos frecuentemente emparejados de lo que lo estarían mediante un patrón que fuese aleatorio en términos de educación, el interés está en poder conocer cuál sería el valor de $\Pr(\text{Mujer} = i \cap \text{Hombre} = j)$ para todas las combinaciones de i y j , ya que con esas proporciones, al ponderarlas por el número total de observaciones de parejas de padres, es posible saber cómo sería la distribución por nivel educacional de los padres que sea acorde a un escenario sin homofilia o uno de emparejamiento aleatorio.

Ahora, dado que sabemos que:

$$c_{ij} = \frac{\Pr(\text{Madre}=i \cap \text{Padre}=j)}{\Pr(\text{Madre}=i) * \Pr(\text{Padre}=j)}$$

Entonces si $c_{ij} = 1$:

$$\Pr(\text{Madre} = i \cap \text{Padre} = j) = \Pr(\text{Madre} = i) * \Pr(\text{Padre} = j)$$

Y dado que las probabilidades marginales para el padre y la madre son conocidas a partir de los datos de la tabla 9, entonces podemos construir una distribución *benchmark*, sobre la cual poder comparar la distribución generada por “el emparejamiento aleatorio de los padres”⁴⁰. Así, mientras más pequeñas sean las diferencias entre ambas, más podremos confiar en la aleatorización realizada. La distribución *benchmark*⁴¹ se presenta en la tabla 11, en donde podemos apreciar que los porcentajes totales o marginales coinciden con los de la tabla 9, sólo cambia la distribución entre cada una de las clasificaciones de los niveles educacionales.

³⁹ Resulta sumamente interesante revisar cómo cambió el nivel de AM desde esta generación de “padres” en comparación con los niveles de los individuos/entrevistados/hijos, en donde se aprecia que se mantiene la tendencia que se venía observando desde 1990: un alto grado de homofilia para las personas con mayores niveles educacionales y menor para los niveles de poca escolaridad; con una tendencia a la baja por parte del primer grupo, y al alza del segundo.

⁴⁰ En concreto aleatorizamos, sin reemplazo, los niveles educacionales entre los padres.

⁴¹ Si se calculara nuevamente el AM para los resultados de la tabla 10 obtendríamos una matriz de llena de unos.

TABLA 11: Distribución por nivel educacional de los padres de los encuestados del escenario *benchmark*

	Educación del Padre				
Educación de la Madre	Sin Educación	Básica	Media	Superior	Total
Sin Educación	36,72%	14,69%	8,54%	2,15%	62,10%
Básica	13,72%	5,49%	3,19%	0,80%	23,19%
Media	7,39%	2,96%	1,72%	0,43%	12,50%
Superior	1,30%	0,52%	0,30%	0,08%	2,20%
Total	59,13%	23,65%	13,75%	3,47%	100,00%

Fuente: Elaboración propia a partir de la Encuesta Casen 2013.

El siguiente paso de la simulación era “aleatorizar a los padres de los individuos” (en la práctica se aleatorizó el nivel educacional de éstos). Se optó por aleatorizar los niveles educacionales de ambos progenitores, aunque de haberlo hecho sólo con uno de ellos los resultados no cambiaban sustancialmente. A partir de dicha aleatorización se obtuvo la distribución por clasificaciones *ij* expuestas en la tabla 12.

TABLA 12: Distribución, post aleatorización, del nivel educacional de los *nuevos* padres de los encuestados

	Educación del Padre				
Educación de la Madre	Sin Educación	Básica	Media	Superior	Total
Sin Educación	36,69%	14,68%	8,59%	2,15%	62,10%
Básica	13,69%	5,52%	3,17%	0,82%	23,20%
Media	7,44%	2,95%	1,69%	0,42%	12,51%
Superior	1,31%	0,50%	0,31%	0,08%	2,19%
Total	59,13%	23,65%	13,75%	3,47%	100,00%

Fuente: Elaboración propia a partir de la Encuesta Casen 2013.

Como podemos ver, dado el gran tamaño muestral de la Encuesta Casen, el proceso de aleatorización garantizó mínimas diferencias (recién al segundo decimal), entre la nueva distribución de niveles educacionales simulados y la *benchmark* de la tabla 11.

En la etapa siguiente de la simulación se requería estimar un nuevo nivel educacional de los encuestados. Se optó por esto último y no por calcular la escolaridad de éstos como variable continua, pues con ello sólo sabríamos la cantidad de años que hubiese estado estudiando el individuo, pero no si hubiera completado ciertos niveles educacionales. Por ejemplo, si se predecía

que el individuo tenía 14 años de escolaridad, ¿correspondería eso a educación superior de nivel técnico completo o educación superior profesional incompleta? Y dado que, como fue analizado previamente, existe un *sheepskin effect* notorio en el mercado laboral chileno, entonces se optó por predecir, mediante un Probit Ordenado, el nivel educacional de las personas.

Para esto último, primero se utilizaron las mismas variables expuestas en la tabla 7, pero en vez de usar el nivel educacional de uno de los padres (que sólo permitía observar el efecto de la MEI), incluimos 9 variables dicotómicas (10 al considerar la categoría base), donde cada una indicaba qué combinación original de nivel educacional poseían sus progenitores⁴² reales (sin simulaciones). Los efectos marginales de la estimación anterior, para la categoría educación superior, se presentan en la tabla A.3 en anexos.

Posteriormente se procede a predecir un nuevo nivel educacional para cada individuo, ponderando los coeficientes estimados por su variable correspondiente, de tal forma de obtener un valor predicho de la variable latente. Ésta, posteriormente, determinará el nuevo nivel educacional del individuo, dependiendo de entre qué cortes quede posicionada. Es importante mencionar que para este caso se construyen nuevas variables dicotómicas que incluyen la nueva combinación simulada/aleatorizada de niveles educacionales de los padres. Y son estas nuevas variables las que se ponderan por los coeficientes correspondientes de las dicotómicas originales⁴³ obteniéndose así el nuevo nivel educacional predicho para cada individuo para un escenario donde no hubiese habido homofilia.

Debe mencionarse que la forma en que están construidas las variables dicotómicas permite recoger un efecto no sólo de la MEI sino que también de AM, esto debido a que cada variable identifica: (i) el nivel educacional de los padres (lo que permite identificar el impacto de la MEI) y (ii) su distribución de entre ambos progenitores (lo que permitirá distinguir el efecto del AM).

Finalmente, con este último se procede a predecir nuevos ingresos laborales para cada persona, realizando⁴⁴ un MCO para los hombres y utilizando el método de Heckman⁴⁵ (1979) para las mujeres⁴⁶. Los regresores empleados son los indicados por Mincer (1974) pero en vez de usar la escolaridad se usaron variables dicotómicas de los niveles educacionales; a su vez, se usaron como

⁴² Se tomaron los mismos 4 niveles usados hasta el momento (sin educación, básica, media y superior) pero no hay 16 variables dicotómicas ya que, por ejemplo, se usa una misma variable para indicar que el padre posee básica y la madre media o para indicar que el padre posee media y la madre básica, y así con todas las demás celdas de la matriz de la tabla 10.

⁴³ Para las demás variables se usan sus observaciones originales, no son simuladas.

⁴⁴ Todas las estimaciones presentan errores estándar robustos a heterocedasticidad.

⁴⁵ Para corregir el sesgo de selección en participación en el mercado laboral, usando como variable de exclusión si la mujer estaba casada o no.

⁴⁶ La predicción fue hecha con los coeficientes obtenidos al estimar el ingreso de la ocupación principal de la metodología tradicional de la Encuesta Casen.

variables explicativas el set de 9 *dummies* relativas al tipo de clasificación *ij* del nivel educacional de los padres del individuo. Esto último se hizo por dos razones. La primera, ver el efecto del AM en los ingresos laborales de los individuos, en donde es posible que existan efectos relativos a habilidades cognitivas y no cognitivas que fuesen otorgadas por los padres, así como redes de contactos u otras vías mediante las cuales el nivel educacional de los padres afecte los ingresos laborales. La segunda razón hace referencia a poder identificar inescudadamente el efecto de la escolaridad, la cual puede estar absorbiendo el efecto del nivel educacional de los padres sobre los ingresos.

Luego se procedió a ponderar las variables de cada individuo por sus coeficientes correspondientes, pero usando los niveles educacionales simulados en vez de los observados, y con los mismos valores simulados del set de 9 variables dicotómicas usados para predecir el nivel educacional, estimando así el ingreso de la ocupación principal predicho. Las regresiones para hombres y mujeres se encuentran en la tabla A.4 de anexos.

A continuación el procedimiento, mediante el cual poder generar la nueva distribución de ingresos, se complejiza. Esto debido a que si la simulación se lleva de forma directa, tal como se indicó recién, tanto la nueva distribución de niveles educacionales, como la del ingreso de la ocupación principal, poseerán una varianza considerablemente subestimada. Gran parte de este fenómeno sucedería porque se predeciría un set limitado de variables de resultado, dado que la combinatoria de observaciones producida por las diferentes variables explicativas utilizadas, también sería limitada. Es decir, se perdería la varianza generada por las variables no observables, la cual se asumiría como idéntica entre individuos con iguales variables observables. Y dado que nuestro proceso de simulación cuenta con dos etapas (la primera que predice un nivel de escolaridad y la segunda, que a partir del último, genera ingresos), entonces en la última no sólo se generaría el inconveniente ya expuesto, sino que se exacerbaría aún más, producto que las variables explicativas de los niveles educacionales (que son un insumo para la segunda predicción) también tendrían una menor varianza, producto de su predicción en la primera etapa.

De esta forma, este proceder permitiría sólo identificar el Gini *entre* las diferentes categorías⁴⁷ observadas y no *al interior* de éstas, porque sería nulo, por construcción. Por lo tanto tendríamos lo siguiente:

$$Gini_{real} = Gini_{inter\ categ.} + Gini_{intra\ categ.}$$

Donde sólo observamos el $Gini_{inter\ categ.}$ y no el $Gini_{intra\ categ.}$. Para lidiar con esta dificultad, este trabajo emplea dos metodologías diferentes que apuntan a poder estimar el $Gini_{real}$. La primera de ellas propone aplicar el mismo procedimiento de la simulación a las variables originales, y no a las simuladas. Esto es, al predecir el nivel educacional, estimarlo con los valores ob-

⁴⁷ Que surgen a partir de las variables explicativas utilizadas.

servados (los efectivamente reportados en la encuesta) del set de 9 variables dicotómicas respectivas a las clasificaciones de emparejamiento de los padres de los individuos. Luego, con ese nivel educacional predicho, se estimarían los ingresos, para finalmente aplicar el cálculo del Gini sobre dicha distribución. En suma, sería como una especie de tratamiento placebo sobre los datos.

La idea de lo anterior es que si se cumple el supuesto que no hubiese diferencias en el Gini intra categorías de la distribución simulada y la generada por el tratamiento placebo, entonces el Gini del escenario contrafactual de padres sin emparejamiento selectivo, vendría dado por:

$$\Delta Gini = \widehat{Gini} - \widetilde{Gini} = (\widehat{Gini}_{inter\ c.} + \widehat{Gini}_{intra\ c.}) - (\widetilde{Gini}_{inter\ c.} + \widetilde{Gini}_{intra\ c.})$$

$$Si: \widehat{Gini}_{intra\ c.} = \widetilde{Gini}_{intra\ c.}$$

$$\widehat{Gini} = \widetilde{Gini} - \Delta Gini = \widetilde{Gini} - (\widehat{Gini}_{inter\ c.} - \widetilde{Gini}_{inter\ c.})$$

Donde \widehat{Gini} corresponde al Gini de la distribución simulada a la partir del caso sin homofilia, mientras que \widetilde{Gini} es el de la distribución placebo. Por lo tanto si se cumple nuestro supuesto identificador, la diferencia entre ambos Ginis subestimados debiese equivaler al efecto real del emparejamiento aleatorio de los padres de los individuos en la desigualdad.

Paralelamente también se llevó a cabo otro procedimiento, menos restrictivo que el anterior que, por lo mismo, consideramos como el más válido. Éste consiste en hacerse cargo de la variabilidad no considerada, radicada en las variables no observables. En donde, en un escenario ideal, si conociésemos la distribución de los errores (y su varianza, ya que la media ya es conocida pues se asume que es 0), podríamos sumarle un error aleatorio, según su distribución, a cada individuo, sobre su variable predicha. Pero dado que no conocemos su distribución, ni su varianza, entonces se optó por calcular los residuos de cada individuo, y posteriormente aleatorizarlos entre éstos.

Con respecto a este último proceso, algunas aclaraciones. Fue hecho para ambas etapas, la que estimaba los niveles educacionales y la que lo hacía con el ingreso de la ocupación principal; ya que en ambas etapas se perdía variabilidad en las variables al no tomar en cuenta las no observables. Los residuos fueron calculados como la diferencia entre la variable dependiente observada y la variable predicha, calculada a partir de los valores reales de las variables explicativas⁴⁸, es decir, no el set simulado de variables dicotómicas de AM, ni del nivel educacional simulado (en el caso de la etapa de ingresos). Por otro lado, al momento de la aleatorización de los residuos entre los individuos, es razonable pensar que dicho proceso debiese hacerse con reemplazo, es decir, el residuo de un individuo pudiese tocarle nuevamente a otro, lo cual es más razonable que restringir

⁴⁸ La razón de ello es porque el vector de coeficientes estimados para las variables, que permite predecir posteriormente las variables dependientes, es calculado usando las observaciones originales (no simuladas), dado el conjunto de variables observables incluidas (y por lo tanto dado el de no observables excluidas), por ende para poder recuperar la variabilidad perdida del modelo, producto de las no observables, es más apropiado usar los residuos con las variables originales, que permitieron calcular el vector de coeficientes antes mencionado.

el set de posibles residuos a los individuos sucesivos. No obstante presentaremos los resultados de simulaciones con y sin reemplazo. También, es importante aclarar que los residuos de la ecuación de ingresos fueron aleatorizados distinguiendo por sexo y nivel educacional, ante la posibilidad de que existiese heterocedasticidad.

Para minimizar el riesgo de que en el proceso de aleatorización de los residuos tocara una asignación muy favorable o desfavorable, en términos de desigualdad y de los otros resultados relevantes, se optó por simular este proceso 500 veces, de tal forma de poder generar intervalos de confianza para cada uno de dichos resultados.

Es importante mencionar que la predicción fue hecha con respecto al ingreso de la ocupación principal, y no otro, debido a que éste era el que podía predecirse con mayor bondad de ajuste, usando en ello las variables de niveles educacionales. Esto era importante para poder reducir al mínimo el problema de identificación del efecto del escenario contrafactual. Por otra parte, considerando que el nivel educacional también pudiera tener efectos en la determinación de otros ingresos no laborales, entonces se optó por calcular el Gini solamente sobre el ingreso de la ocupación principal, pues el efecto producido por AM se podría apreciar más directamente, reduciendo otros potenciales sesgos⁴⁹.

Finalmente se debe aclarar que los resultados relativos a los ingresos de la ocupación principal predichos, presentados en la siguiente sección, fueron ajustados, de tal forma que sólo se tomó en cuenta a aquellos individuos que originalmente reportaron una cifra de sus ingresos de su ocupación principal. Es decir, no se consideró a todos quienes se les predijeron un valor para esta variable pero que originalmente no tenían.

Resultados⁵⁰

Las tablas 13 y 14 muestran los principales resultados de todas las simulaciones anteriormente descritas. En la primera se presentan los relativos al caso con imputaciones y en la segunda se exponen las estimaciones hechas sobre la submuestra que incluye sólo a los individuos que reportaron el nivel educacional de ambos de sus padres. En ambos casos, las diferencias en los resultados deben hacerse a

⁴⁹ Esto porque al estar sesgado el nivel educacional predicho, entonces cada vez que se usase una variable en estas condiciones para explicar otra, esta última también lo estaría. Por lo tanto mientras menos variables (tipos de ingresos) intentemos explicar con el nivel educacional, menos sesgaremos la variable final observada (en el caso de los ingresos, la variable observada puede ser, por ejemplo, el ingreso autónomo el cual se compone de otras variables como los ingresos laborales y otros ingresos auto generados, los cuales cada uno de ellos podría estar sesgado si es que fuesen explicados por el nivel educacional también).

⁵⁰ Debemos mencionar que estos resultados representan un equilibrio parcial, por lo que eventualmente pudieran existir otras consecuencias de un escenario como el aquí descrito y que tal vez pudiesen afectar de modo indirecto a dichos resultados.

partir de las columnas comparables entre sí. Las referentes a las de suma de residuos (tanto con y sin reemplazo) pueden compararse con la primera columna, donde se muestran los datos de la muestra total y los de la submuestra, correspondientemente; mientras que las de la estimación directa deben compararse con la columna de la *falsa estimación* o placebo, como la llamamos. Cabe destacar que dado que para la simulación de suma de residuos, se realizó 500 veces el proceso de aleatorización, entonces lo que se presenta está en intervalos de confianza al 95%.

Si bien consideramos los resultados de la columna de la estimación de suma de residuos con reemplazo, de la tabla 13 (con imputaciones) como los resultados de mayor validez, se presentan también los demás a modo de mostrar la robustez de los resultados. En efecto, a modo global, podemos extraer los siguientes resultados a partir de cómo serían las distribuciones de ingresos y de nivel educacional de la sociedad chilena, si los padres de la generación actual no hubiesen presentado emparejamiento selectivo en escolaridad:

1. En todas las especificaciones cae la desigualdad, medida mediante el coeficiente de Gini, en donde observamos una caída de 5 puntos porcentuales, pasando de 0,48 a 0,43. Para poner en perspectiva esta cifra notemos que el Gini del ingreso de la ocupación principal en 1990 era de 0,53, es decir, esta disminución sería el equivalente a la reducción hecha durante casi un cuarto de siglo. En las demás simulaciones también se corrobora una caída en este índice, aunque menor.
2. Por otro lado, la desigualdad en los extremos de la distribución de ingresos, medida mediante el ratio 10/10 parece no haber disminuido, es más, aumentó levemente, pasando de 6,3 a 6,4. Esta alza es leve teniendo en cuenta que el nivel de este indicador en 1990 era de 7,7. Este resultado va en la misma dirección que en las demás simulaciones, a excepción del caso de la simulación directa para la submuestra de la tabla 14.
3. Quizás uno de los resultados más interesantes de este trabajo sea, además de la importante caída en desigualdad, que el ingreso laboral promedio (o de la ocupación principal) cae significativamente. Pasando de poco más de \$500.000 a poco más de \$425.000, monto que corresponde al ingreso promedio de la ocupación principal en el año 1998 en Chile⁵¹. Esto equivale a una caída de un 15% de los salarios. Las demás simulaciones también muestran una caída importante del ingreso promedio, en algunos casos mayor y en otras menor. Por otro lado, en todos los escenarios se aprecia una disminución de la desviación estándar de la distribución de ingresos laborales.
4. Por último podemos observar que la distribución de niveles educacionales tiende a disminuir la densidad en los extremos y abultarse más en los niveles centrales. En efecto, se aprecia un efecto positivo, al reducirse la cantidad de individuos sin escolaridad o con educación básica in-

⁵¹ Ajustado por inflación.

completa (“Sin Educación”) desde un 31% a un 26%. Sin embargo, ello viene asociado a una caída en una magnitud similar, desde 12% a 7%, en la cantidad de personas con educación superior, en donde de ese 10% (5% que se desplazó de cada una) que pertenecía a dichas categorías, pasa poco más de un 3% a educación media y un 7% a educación básica, tal que también, presumiblemente, exista un porcentaje que se haya transferido de media a básica. La tendencia de estos resultados se repite en las otras simulaciones.

TABLA 13: Resultados de las diversas metodologías de simulación, aplicadas a la muestra con imputaciones

		Valores Originales Muestra	Simulación de Iteración con Suma de Residuos con Reemplazo con Imputaciones	Simulación de Iteración con Suma de Residuos sin Reemplazo con Imputaciones	Simulación Placebo con Imputaciones	Simulación Directa con Imputaciones
Ingreso de la Ocupación Principal	Gini	0,48	[0,426; 0,428]	[0,440; 0,441]	0,24	0,22
	Ratio 10/10	6,3	[6,40; 6,44]	[6,70; 6,72]	2,68	2,80
	Promedio	501.020	[426.583; 428.284]	[452.695; 453.524]	303.264	281.209
	Desviación Estándar	775.985	[547.318; 564.487]	[622.280; 633.050]	139.639	122.893
	Mínimo	2.270	[2.972; 3.209]	[1.777; 1.881]	38.208	38.208
	Máximo	43.500.000	[22.358.910; 23.956.604]	[42.055.734; 44.714.862]	1.274.917	1.275.243
	Sin Educación (Frec.)	5.417.480	[4.498.602; 4.502.580]	[4.476.164; 4.479.683]	4.717.166	3.722.415
	Sin Educación (%)	31,49%	[26,15%; 26,17%]	[25,91%; 25,93%]	27,31%	21,55%
	Ed. Básica (Frec.)	4.198.088	[5.496.514; 5.502.014]	[5.475.268; 5.480.247]	5.360.432	6.693.344
	Ed. Básica (%)	24,40%	[31,95%; 31,98%]	[31,70%; 31,73%]	31,03%	38,75%
	Ed. Media (Frec.)	5.463.920	[5.985.929; 5.990.752]	[5.993.282; 5.997.880]	6.322.098	6.595.471
	Ed. Media (%)	31,76%	[34,79%; 34,82%]	[34,70%; 34,72%]	36,60%	38,18%
	Ed. Superior (Frec.)	2.125.105	[1.214.901; 1.217.894]	[1.320.408; 1.323.301]	873.421	261.887
	Ed. Superior (%)	12,35%	[7,06% 7,08%]	[7,64%; 7,66%]	5,06%	1,52%

Fuente: Elaboración propia a partir de la Encuesta Casen 2013.

TABLA 14: Resultados de las diversas metodologías de simulación, aplicadas a la muestra sin imputaciones

		Muestra Restringida	Simulación de Iteración con Suma de Residuos con Reemplazo (M.R.)	Simulación de Iteración con Suma de Residuos sin Reemplazo (M.R.)	Simulación Placebo (Muestra Restringida)	Simulación Directa (Muestra Restringida)
Ingreso de la Ocupación Principal	Gini	0,51	[0,462; 0,463]	[0,463; 0,464]	0,27	0,26
	Ratio 10/10	7,3	[7,56; 7,60]	[7,596; 7,627]	3,04	2,96
	Promedio	592.710	[494.775; 496.548]	[506.019; 507.159]	333.165	311.297
	Desv. Estándar	935.981	[699.797; 719.797]	[745.513; 765.538]	172.926	170.571
	Mínimo	2.270	[3.319; 3.494]	[3.199; 3.312]	41.140	41.140
	Máximo	43.500.000	[32.557.526; 34.969.131]	[44.925.607; 47.577.842]	1.400.202	1.648.179
	Sin Educación (Frec.)	839.650	[612.414; 614.338]	[613.969; 615.881]	264.036	127.471
	Sin Educación (%)	15,21%	[11,10%; 11,13%]	[11,09%; 11,12%]	4,77%	2,30%
	Ed. Básica (Frec.)	1.395.572	[1.702.050; 1.705.366]	[1.704.889; 1.708.055]	2.015.886	1.903.782
	Ed. Básica (%)	25,29%	[30,84%; 30,90%]	[30,78%; 30,84%]	36,40%	34,38%
	Ed. Media (Frec.)	2.093.977	[2.382.725; 2.386.088]	[2.377.964; 2.381.521]	2.728.559	3.341.099
	Ed. Media (%)	37,94%	[43,17%; 43,24%]	[42,94%; 43,00%]	49,27%	60,33%
	Ed. Superior (Frec.)	1.189.654	[816.183; 818.542]	[835.854; 838.134]	529.653	165.782
	Ed. Superior (%)	21,56%	[14,79%; 14,83%]	[15,09%; 15,13%]	9,56%	2,99%

Fuente: Elaboración propia en base a Encuesta Casen 2013.

Análisis

Podemos rechazar la hipótesis que planteaba que el nivel educacional de los padres aporta de manera creciente pero a tasas *marcadamente* decrecientes en la propensión de escolarización de sus hijos, puesto que se observa una caída importante en la cantidad de individuos con educación superior. Ello quiere decir que, en muchos casos, no basta con tener un solo padre con alta educación para que el hijo también la tenga, sino que se necesita del aporte que hacen ambos progenitores en su nivel educacional. Es decir, en el agregado no mejora la cantidad de personas con educación superior, lo cual hubiese ocurrido si bastase con sólo el aporte del padre más educado, ya que al disminuir el AM habría más hijos con al menos un padre educado, y dado que el aporte educacional del segundo progenitor no hubiese aportado marginalmente mucho, entonces hubieran incrementado el porcentaje de individuos con mayores niveles educacionales. Nótese que, tal vez, podría pensarse que el aumento hubiese sido absorbido por gente con educación media, pero el porcentaje de incremento de este nivel fue menor que la reducción de quienes tenían educación superior.

En otro ámbito, notamos que las simulaciones hechas de manera directa, comparadas con la *falsa simulación*, muestran diferencias más acotadas de las predichas mediante la metodología de sumar los residuos. Ello puede deberse a que no se haya cumplido el supuesto identificador: ($\widehat{Gini}_{intra\ c.} = \widehat{Gini}_{intra\ c.}$). Por lo tanto es probable que estemos observando sólo la diferencia entre: ($\widehat{Gini}_{inter\ c.} - \widehat{Gini}_{inter\ c.}$), lo que explicaría que éstas sean más acotadas que las presentadas por el otro procedimiento. De todas formas, constituyen un activo valioso para verificar la robustez de nuestras estimaciones.

En relación a los 4 puntos analizados en el apartado anterior podemos inferir que:

1. Si bien es una cuantía considerable la caída en 5 puntos del Gini que se presenta, esto podría ser mucho mayor para las próximas generaciones, en el caso que las parejas de hoy no presentasen emparejamiento selectivo. Esto ocurriría debido a que en el análisis presentado, casi 8 de cada 10 personas poseían originalmente a ambos padres sin educación o con educación básica. De esta forma, el impacto de aleatorizar a los padres no sería muy alto, ya que, tal como lo muestra la distribución *benchmark*, en dicho escenario seguirían quedando poco más de 7 de cada 10 individuos con tanto padre o madre con educación básica o sin educación. Esto quiere decir que la redistribución de padres educados no tendría un efecto tan significativo en las variables de resultado, ya que era un porcentaje bajo los que habían cursado educación media o superior. Por lo tanto, si se analiza con esta perspectiva, entonces efectivamente puede decirse que los resultados aquí expuestos son de gran impacto, ya que bastó redistribuir la poca gente educada que había para ya haber generado impactos importantes.

Si también se observa la tabla A.4 (en Anexos), puede concluirse que los ingresos laborales también se ven afectados de forma directa por la *redistribución* de padres, en algunos casos positivamente y en otros de manera negativa. Por ejemplo, vemos que si comparamos dos parejas, una con ambos progenitores con educación básica y otra con ambos sin educación, entonces la hija de los primeros tendría en promedio 9% más de salarios que la hija de los segundos. Lo cual es menor que la suma de cuánto tendría la hija de dos parejas que se componen de un padre con educación básica y el otro sin educación. En tal caso, entre ambas tendrían 12% más que la hija de parejas con ambos padres sin educación. No obstante, esto contrasta con un caso análogo pero para un par de parejas con educación superior y la otra sin educación, en cuyo caso el primer porcentaje (caso con AM) ascendería a 51% y el segundo (en el caso sin AM) llegaría sólo a 26%.

2. Es interesante notar que la caída en desigualdad que se observa en el Gini, aparentemente proviene de una contracción de la distribución de ingresos de los tramos medios, sin embargo la desigualdad de los deciles extremos parece mantenerse similar. Si bien esto puede parecer extraño, al observar que ahora menos de un 10% de la población posee educa-

ción superior (lo que implicaría que el decil 10 ahora ganase menos proporcionalmente que el decil 1), también debe tenerse en cuenta que en nuestro proceso de simulación también se incluye el nivel de clasificación de los padres *ij* como regresores de los ingresos. Por lo tanto es posible que el hecho de que ahora más gente del decil 1 posea padres con mayores niveles educacionales les permita, si bien no necesaria tener más escolaridad, pero sí mayores ingresos.

3. Tal como se vio en los gráficos 5 y 7, los individuos con educación superior ganan salarios considerablemente mayores que el resto de los niveles educacionales. De esta forma, la caída en el porcentaje de personas con dicho nivel implica un impacto negativo mucho mayor en términos de los salarios promedio, que el impacto positivo producido por la menor proporción de gente sin educación.
4. Si bien no es trivial la comparación de si la nueva distribución de niveles educacionales es mejor o peor que la original, sí se puede decir que es positiva la disminución de gente sin educación, pero cayendo también el porcentaje de individuos de mayor cualificación. Por otra parte, este escenario es más virtuoso para otorgar oportunidades de forma más equitativa en la sociedad, lo cual a largo plazo podría ser más beneficioso, ya que, de existir emparejamiento selectivo, éste, hipotéticamente, ocurriría bajo un *piso* educacional mayor.

De todas formas, es interesante observar detalladamente la tabla A.3 (en Anexos). En ella se muestran los efectos marginales de las variables explicativas del Probit Ordenado para estimar el nivel educacional, para la categoría de educación superior. Como vimos, si bien la proporción de personas con este nivel cae en la simulación, podemos observar que en un escenario con bajo AM (en contraste a otro con AM elevado), mediante la mayoría de las distintas clasificaciones *ij*, mejora la propensión de tener educación superior, sólo que, seguramente, el aumento de esta variable no es suficiente para superar el corte que clasificaría al individuo como alguien con este nivel. A ello debe sumarse que la excepción de este análisis son las clasificaciones de ambos padres con educación superior y ambos sin educación, comparado con una situación de tener parejas de padres uno con educación superior y otro sin educación.

Para ilustrar con claridad lo anterior revisemos un ejemplo: observemos los efectos marginales de la tabla A.3 y tomemos a un individuo con ambos padres con el máximo nivel educacional y a otro con ambos con el mínimo. Veríamos que el primero tendría 67% más de probabilidades de tener educación superior que el segundo. Si suponemos, como lo hicimos en nuestra simulación, que los mismos individuos tienen un padre o madre con educación superior y su otro progenitor sin educación, en este caso ambos tendrían 28% más de probabilidades de tener educación supe-

rior que alguien con sus dos padres sin educación. Es decir, en el agregado, el haber *distribuido* a los padres de estos dos individuos, termina generando una propensión de un 56%⁵¹ a que alguien tenga educación superior. Lo que contrasta con el caso con AM en el cual la probabilidad de que alguien tuviese educación superior sería 11% mayor. Ahora, si realizamos el mismo ejercicio con la mayoría de las demás clasificaciones *ij* veremos que ahí el caso es distinto, y sí mejora la propensión de tener educación superior en el agregado para el escenario sin AM. Sin embargo, y como ya ilustramos anteriormente, la proporción de padres sin educación era tan elevada que al final ello terminó “nivelando hacia abajo”.

Implicancias para políticas públicas de un alto AM y baja MEI

Tal como señala Torche (2010), un elevado nivel de segregación se traduce en elevados índices de homofilia, ya que la primera reduce las probabilidades de encontrar una pareja de distinto nivel socioeconómico. Este trabajo por su parte nos permite pensar en una hipótesis complementaria a la anterior. Si a dicho contexto se le suma una baja movilidad educacional intergeneracional, entonces los hijos que surjan de tales familias también heredarán el nivel educacional de sus padres. Si vivían en un entorno segregado, entonces desde el establecimiento educacional al que asistan hasta en el barrio donde vivan frecuentarán otros hijos con un nivel socioeconómico similar. Esto sesgará considerablemente sus probabilidades de emparejarse con alguien de otro nivel educacional, reforzando nuevamente la segregación y el ciclo aquí descrito.

El análisis descrito es especialmente trascendente al ver la elevada magnitud y significancia estadística que tienen el set de variables de clasificaciones de los niveles educacionales de los padres que se presentan en la tabla A.3⁵² (en Anexos).

Lo anterior ilustra cómo, bajo un contexto de poca MEI, el AM es tanto consecuencia y causa de la segregación social. Por lo tanto, asumiendo constante la movilidad educacional, una manera de terminar con este círculo vicioso que menoscaba la meritocracia, es realizando políticas públicas que apunten a reducir la segregación. Por ejemplo, para el contexto chileno, medidas como el fin al financiamiento compartido en los colegios subvencionados o políticas urbanísticas⁵³ que inhiban la segregación podrían permitir romper con este ciclo.

⁵¹ Cifra que proviene de sumar toscamente el 28% de ambos individuos. Evidentemente en una economía con dos individuos como la descrita en el ejemplo dicha probabilidad sería menor aun, ya que sería la unión de ocurrencia de dos eventos, por lo que habría que todavía restarle la intersección de ocurrencia de ambos.

⁵² Si bien en ésta se exhiben sólo los efectos marginales para la categoría de educación superior, los resultados de las demás también son altamente significativos y de gran magnitud.

⁵³ En especial en ciudades fuera de la Región Metropolitana, donde la segregación aún no está totalmente arraigada y queda más espacio de acción.

Sin embargo, esta investigación también ha mostrado evidencia en favor del emparejamiento selectivo (y con ello de la segregación), al menos para el contexto aquí analizado (con una generación de padres con altos porcentajes de gente sin educación y bajo un escenario de altos retornos a la educación superior), en donde se aprecian mayores ingresos promedio y una mayor cantidad de personas con educación superior.

Es importante señalar que si el premio a la educación superior fuese menor que el que posee Chile en la actualidad, entonces es posible que la disminución que se produciría en la gente con este nivel tendría un efecto menor en los salarios. O que si dado los efectos marginales de la tabla A.3, hubiese menos padres sin educación (como eventualmente ocurrirá para la próxima generación), entonces incluso podría aumentar la cantidad de individuos con educación superior, como se revisó en el apartado anterior.

Otra conclusión relevante a tener en cuenta es la referente a la heterogeneidad de efectos que existe entre las distintas clasificaciones de niveles educacionales de los padres sobre los ingresos y la propensión a escolarizarse de los individuos. En el primer caso, dicho efecto puede provenir del capital cultural que proveen los padres; mientras que en el segundo, puede, por ejemplo, tener relación con las redes de contactos que éstos otorgan. Lo trascendental es que en algunos casos se ve que existe algo similar a un *efecto par*, en donde el padre con mayor educación logra aportar más favorablemente a la propensión a escolarizarse y a los salarios de su hijo, de lo que podría perjudicar su pareja de menor educación. Por otra parte, en otras de las clasificaciones *ij* podría haber algo similar a un *efecto "manzana podrida"* en donde ocurriría lo contrario que en el caso anterior.

En suma, se espera que esta sección de este trabajo sirva como insumo para la toma de decisiones políticas y la discusión de algunos de los costos y beneficios asociados a medidas que afecten directa o indirectamente el AM.

Conclusiones

La primera parte de este trabajo se centró en presentar cuál había sido el panorama de la desigualdad de Chile entre los años 1990 y 2013. Ésta se caracterizó, durante este período, por haber disminuido significativamente, en donde el Gini medido para el ingreso monetario mostró una caída desde 0,56 a 0,52, lo cual es positivo al contrastarlo con la evolución mostrada por otros países desarrollados que retrocedieron en sus índices de desigualdad. No obstante, debe tenerse en cuenta que Chile aún es una economía emergente y se encuentra en una etapa de desarrollo distinta. Además sus cifras siguen siendo muy altas, lo que permite especular acerca de que esfuerzos en el margen pudieran tener mayores rendimientos en el fin de reducir la desigualdad.

Ya sea el índice de Gini o las razones de deciles, todas muestran que nuestro país se caracterizó por un alza en la desigualdad de ingresos hacia fines de los años 90 e inicios de los 2000, que luego cayó a niveles más bajos de los observados en 1990 para la década del 2010. Dicha evolución se condice (es similar) con las tendencias mostradas por otras variables, relevantes a la hora de analizar la desigualdad, como los retornos de la educación, participación de los ingresos laborales en los ingresos totales del hogar y correlación de la escolaridad entre parejas. Ambos factores, sumados a una caída en la varianza de la escolaridad, una reducción del AM para los niveles más educados (que fueron los que más crecieron durante el período), y efectos períodos que mostraron una evolución positiva en pos de disminuir la desigualdad, fueron elementos relevantes a la hora de explicar su caída durante el período 1990-2013.

Por otra parte, también se analizó que a pesar de la considerable expansión de la matrícula en educación superior (que supondría una expansión en la oferta laboral que redujera los salarios de la mano de obra calificada), aparentemente también existió una expansión de la demanda, lo cual explicaría por qué no se redujeron tanto más los salarios (y con ello el retorno de la educación superior). De igual forma, pudimos observar que proporcionalmente hubo una mayor expansión de la matrícula de educación superior técnica, la cual efectivamente sí vio reducidos en un mayor porcentaje sus retornos; en comparación con la universitaria.

También es interesante mencionar que, controlando por el promedio y la dispersión de la escolaridad, no se aprecian efectos cohortes significativas, ni tampoco tendencias significativas de efectos edad. Sólo los efectos período parecen ser robustos ante todas las especificaciones.

Para explicar por qué se han mantenido tan elevadas las cifras de desigualdad, aludimos a que el premio a la educación superior (en especial el de la educación universitaria) presenta sólo leves caídas durante el período observado. Pero además presentamos un análisis que hace referencia a la existencia de factores estructurales de la sociedad chilena que limitan la distribución de oportunidades a los grupos de bajos ingresos. Ante ello surge como explicación plausible el hecho de que nuestro país posee elevados índices de homofilia o emparejamiento selectivo en escolaridad, lo que provoca dos efectos negativos en la desigualdad: uno directo, el cual implica que se formen familias de altos ingresos y otras de bajos ingresos y muy pocas de ingresos medios; y otro efecto más de largo plazo que se refiere a que una elevada homofilia, en conjunto con baja MEI, lo que puede provocar efectos perjudiciales para la distribución de ingresos en términos de igualdad, como fue presentado en la simulación del escenario contrafactual. Sin embargo, se constata, en este caso, que disminuirían en 5% las personas con educación superior, producto que el tener al menos un padre de baja escolaridad parece ser más perjudicial que el tener al menos uno con alto nivel educacional. Ello, además, repercute en una disminución de un 15% de los salarios. Aunque se constata una caída en el Gini equivalente a lo que se ha avanzado en el último cuarto de siglo, además de un escenario con menos gente sin educación y con más igualdad de oportunidades.

Si no nos hacemos cargo de estos factores estructurales que limitan considerablemente el actuar de otras políticas públicas que pudiesen atacar a la desigualdad de ingresos, entonces todo indica que sólo algunos grupos privilegiados de la población contarán con los medios para poder surgir y la alta educación se concentrará sólo en ciertos grupos, generando que la solución al problema sea simplemente exclusiva para algunos.

Dos desafíos: desigualdad y escolaridad. Todo este trabajo viene a mostrar el importante impacto que posee esta última para explicar la primera, y cómo resulta ser la clave esencial para poder construir una sociedad más justa, más inclusiva y con más oportunidades. Siendo éste el diálogo que surge entre ambos.

Anexos

TABLA A.1: Distribución porcentual por nivel educacional de la educación de los padres de los encuestados

	Educación del Padre				
Educación de la Madre	Sin Educación	Básica	Media	Superior	Total
Sin Educación	44,22%	6,95%	2,20%	0,28%	53,65%
Básica	5,02%	16,89%	4,10%	0,85%	26,85%
Media	1,37%	2,70%	10,35%	1,83%	16,24%
Superior	0,09%	0,31%	0,78%	2,07%	3,25%
Total	50,70%	26,85%	17,43%	5,03%	100,00%

TABLA A.2: *Assortative Mating* de los padres de los encuestados

	Educación del Padre			
Educación de la Madre	Sin Educación	Básica	Media	Superior
Sin Educación	1,63	0,48	0,24	0,10
Básica	0,37	2,34	0,88	0,63
Media	0,17	0,62	3,65	2,24
Superior	0,05	0,35	1,38	12,68

TABLA A.3: Efectos marginales de Probit Ordenado del nivel educacional con variables dicotómicas⁵⁴ que indican MEI y AM, para el nivel de educación superior

Variab les	(1)	(2)
Padre o Madre sin educ. y Padre o Madre con básica		0,0808
P. o M. sin educ. y P. o M. con media		0,1345
P. y M. con básica		0,1437
P. y M. con media		0,2298
P. o M. sin educ. y P. o M. con superior		0,2755
P. o M. con básica y P. o M. con media		0,3113
P. o M. con básica y P. o M. con superior		0,5446
P. o M. con media y P. o M. con superior		0,5834
P. y M. con superior		0,6705
Sexo (Hombre = 1)	0,0015	0,0021
Edad	0,0175	0,0099
Edad*Edad	-0,0002	-0,0001
Tiene pareja (Sí = 1)	-0,0106	-0,0094
Zona (Urbano = 1)	0,0585	0,0284
Número de personas en el hogar	-0,0070	-0,0034
R2	0,17	0,28

Nota: Todas las variables son significativas al 0,1%.

⁵⁴ La categoría base es madre y padre sin educación.

TABLA A.4: Estimación del ingreso de la ocupación principal para hombres (MCO) y mujeres (método de Heckman)

VARIABLES	Hombres	Mujeres (ecuación salarios)	Mujeres (ec. participación)
Educación Básica	0.186***	0.132***	-0.0448
Educación Media	0.449***	0.421***	0.298***
Educación Superior	1.177***	1.187***	0.934***
Experiencia	0.0322***	0.0312***	0.0817***
Experiencia*Experiencia	-0.000450***	-0.000510***	-0.00149***
Tiene Pareja (1 = Sí)	-	-	0.0791***
Padre o Madre sin educ. y Padre o Madre con básica	0.0385*	0.0636***	0.0256
P. o M. sin educ. y P. o M. con media	0.0726***	0.0940***	0.0867**
P. y M. con básica	0.0845***	0.0859**	0.0858**
P. o M. sin educ. y P. o M. con superior	0.106***	0.131***	0.105***
P. o M. con básica y P. o M. con media	0.183***	0.236***	-0.202
P. y M. con media	0.0644	0.228***	-0.00564
P. o M. con básica y P. o M. con superior	0.206***	0.158***	-0.0132
P. o M. con media y P. o M. con superior	0.264***	0.371***	0.225***
P. y M. con superior	0.412***	0.505***	-0.344***
Constante	11.89***	11.41***	-0.960***
N	50,723	91,154	
R2	0.31	-	-
*** p<0.01, ** p<0.05, * p<0.1			

Referencias

- Becker, G. 1973. "A Theory of Marriage: Part I". *Journal of Political Economy*, 813-846.
- Beyer, H. 2000. "Educación y desigualdad de ingresos: una nueva mirada". *Estudios Públicos*.
- Beyer, H. 2011. "¿Que veinte años no es nada...? Una mirada a la desigualdad de ingresos a partir de las Encuestas Casen". *Estudios Públicos*.
- Breen, R., & Andersen, S. 2012. "Educational Assortative Mating and Income Inequality in Denmark". *Demography*: 867-887.
- Buss, D. 1985. "Human Mate Selection: Opposites Are Sometimes Said to Attract, but in fact We Are Likely to Marry Someone Who Is Similar to Us in Almost Every Variable". *American Scientist*: 47-51.
- Contreras, D., Rodríguez, J., & Urzúa, S. 2013. *The origins of inequality in Chile*.
- Deaton, A. 1997. "The Analysis of Household Surveys". *Johns Hopkins University Press*.
- Eika, L., Mogstad, M., & Zafar, B. 2014. "Educational Assortative Mating and Household Income Inequality". *Federal Reserve Bank of New York*.
- Fernández, R., Guner, N., & Knowels, J. 2005. "Love And Money: A Theoretical and Empirical Analysis of Household Sorting And Inequality". *The Quarterly Journal of Economics*.
- Ferrer, A., & W, C. 2001. *Sheepskin Effects and the Returns to Education*.
- Gaviria, A. 2007. "Social Mobility and Preferences for Redistribution in Latin America". *Economía*: 55-96.
- González, R., & Mackenna, B. 2015. "Cuando el éxito personal no se traduce en movilidad social: el caso de los "triunfadores frustrados" en Chile".
- Greenwood, J., Guner, N., Kocharkov, G., & Santos, C. 2014. Marry Your Like: Assortative Mating and Income Inequality. *American Economic Review* 104 (5): 348-53.
- Heckman, J. (1979). "Sample Selection Bias as a Specification Error". *Econometrica* 47 (1): 153-161.
- Hungerford, T., & Solon, G. 1996. "Sheepskin Effects in the Returns to Education". *Review of Economics and Statistics*.
- Jaeger, D., & Page, M. 1996. "Degrees Matter: New Evidence on Sheepskin Effects in the Returns to Education". *Review of Economics and Statistics*: 733-740.
- Mason, K., Winsborough, H., Mason, W., & Poole, W. 1973. "Some Methodological Issues in Cohort Analysis of Archival Data". *American Sociological Review*: 242-258.
- Mincer, J. (1974). "Schooling, Experience and Earnings". *National Bureau of Economic Research*.
- Ministerio de Desarrollo Social. 2013. *Informe de desarrollo social*. Santiago.
- Nuñez, J., & Miranda, L. 2007. "Recent findings on Intergenerational Income and Educational Mobility in Chile". *Serie Documentos de Trabajo, Departamento de Economía Universidad de Chile*.
- OECD. 2011. *An Overview of Growing Income Inequalities in OECD Countries: Main Findings*.

- Sapelli, C. 2003. "Ecuaciones de Mincer y las Tasas de Retorno a la Educación en Chile: 1990-1998". Instituto De Economía Pontificia Universidad Católica De Chile, Documento de Trabajo.
- Sapelli, C. 2011. "A cohort analysis of the income distribution in Chile". *Estudios de Economía*: 223-242.
- Torche, F. 2005a. "Unequal but Fluid: Social Mobility in Chile in Comparative Perspective". *American Sociological Review*: 422-450.
- Torche, F. 2010. "Educational Assortative Mating and Economic Inequality: A Comparative Analysis of Three Latin American Countries". *Demography*, 47(2): 481-502.
- Valenzuela, J., & Duryea, S. 2011. "Examinando la prominente posición de Chile a nivel mundial en cuanto a desigualdad de ingresos: comparaciones regionales". *Estudios de Economía*: 259-293.
- Warren, B. 1996. "A Multiple Variable Approach to the Assortative Mating Phenomenon". *Eugenics Quarterly*: 285-290.
- Yang, Y., Fu, W., & Land, K. 2004. "A Methodological Comparison of Age-Period-Cohort Models: The Intrinsic Estimator and Conventional Generalized Linear Models". *Sociological Methodology*: 75-110.
- Yang, Y., Schulhofer-Wohl, S., Fu, W., & Land, K. 2008. "The Intrinsic Estimator for Age-Period-Cohort Analysis: What It Is and How To Use It". *American Journal of Sociology*: 1697-1736.

Serie Debates de Política Pública

Cambios en la participación electoral tras la inscripción automática y el voto voluntario
LORETO COX A. Y RICARDO GONZÁLEZ T.
Número 14, Marzo de 2016

Las desigualdades en la atención médica en los últimos 20 años
CAROLINA VELASCO O. Y JOSEFA HENRÍQUEZ (autores)
Número 13, Noviembre de 2015

Enfoques complementarios para la evaluación social de proyectos
CLAUDIO A. AGOSTINI Y SLAVEN RAZMILIC (autores)
Número 12, Octubre de 2015

Reforma de la política: Una mirada sistémica
ISABEL ANINAT S., LUCAS SIERRA I. Y RICARDO GONZÁLEZ T. (autores)
Número 11, Septiembre de 2015

Comunidades locales y proyectos de inversión: Hacia la construcción de consensos
ANDRÉS HERNANDO Y SLAVEN RAZMILIC (autores)
Número 10 julio 2015

La prohibición a los fines de lucro y propuestas de gobierno para las universidades chilenas
IGNACIO VALENZUELA NIETO
Número 9 junio 2015

Fortalecimiento de la carrera docente
SYLVIA EYZAGUIRRE Y FERNANDO OCHOA
Número 8 abril 2015

Fortalecimiento de la función fiscalizadora del Servel
ISABEL ANINAT S. Y RICARDO GONZÁLEZ T.
Número 7 abril 2015

Subsidio al arriendo: primeros resultados y pasos a seguir
SLAVEN RAZMILIC
Número 6 marzo 2015

El sector energético en Chile y la Agenda de Energía 2014: Algunos elementos para la discusión
ANDRÉS HERNANDO
Número 5 diciembre 2014

Impuesto territorial y financiamiento municipal
SLAVEN RAZMILIC
Número 4 noviembre 2014

Desafíos y algunos lineamientos para el sistema de seguros de salud en Chile
CAROLINA VELASCO O.
Número 3 octubre 2014

Financiamiento permanente no electoral de los partidos políticos
ISABEL ANINAT S. Y RICARDO GONZÁLEZ T.
Número 2 septiembre 2014

Un sistema electoral mixto para el presidencialismo chileno
LUCAS SIERRA I. Y RICARDO GONZÁLEZ T.
Número 1 agosto 2014

CENTRO DE ESTUDIOS PÚBLICOS

www.cepchile.cl