ASOCIACION ARGENTINA
DE ECONOMIA POLITICA

ANALES | ASOCIACION ARGENTINA DE ECONOMIA POLITICA

XLVII Reunión Anual

Noviembre de 2012

ISSN 1852-0022
ISBN 978-987-28590-0-8

EDUCACIÓN Y DESIGUALDAD: UNA METODOLOGÍA DE DESCOMPOSICIÓN BASADA EN DOS INTERPRETACIONES DE LA ECUACIÓN DE MINCER. EVIDENCIA PARA ARGENTINA

Alejo Javier

Educación y Desigualdad: una metodología de descomposición basada en dos interpretaciones de la ecuación de Mincer. Evidencia para Argentina.

Javier Alejo

CEDLAS-UNLP

Resumen [1]

Dentro la literatura empírica que estudia la distribución salarial se encuentra regularmente evidencia de que mayores niveles de educación predicen un mayor nivel de desigualdad. Este trabajo propone una metodología para descomponer el efecto marginal de la educación sobre la desigualdad salarial utilizando encuestas de hogares. Utilizando supuestos y técnicas de estimación estándar es posible medir la importancia relativa de dos posturas adoptadas por la literatura empírica con respecto a los factores clave detrás de éste fenómeno. La aplicación de la descomposición propuesta al caso argentino muestra que la fuerza más importante detrás del efecto desigualador de la educación es la convexidad en la relación entre el salario y el nivel educativo. Una interpretación rápida de los resultados podría indicar que para aprovechar al máximo los efectos de la mejora educativa las políticas deberían poner el foco sobre la diversidad en el stock del capital humano.

Abstract

Regular evidence in empirical literature about wage distribution is that an improvement in the educational levels predicts an increase in inequality. This paper proposes a methodology based on Mincer equations estimated with household surveys to decompose the marginal effect of education on wage inequality. Using standard assumptions and methodologies, the relative importance of two explanations adopted by empirical literature on this phenomenon can be measured. The application of the proposed decomposition for the Argentinean case shows that the most important force behind the distributional effect is the convexity relationship between wage and education. A quick interpretation could indicate that policies should be focused on the diversity of human capital stock in order to take advantage of improvements in education.

Palabras clave: desigualdad salarial, retornos a la educación, quantile regression, descomposición, Argentina.

Códigos JEL: I24, I25 y D31.

[1] Este trabajo es parte del plan de tesis del Doctorado en Economía de la UNLP, financiada mediante el sistema de becas internas del CONICET. Agradezco a Walter Sosa Escudero y Leonardo Gasparini por sus valiosos comentarios. Cualquier error es de mi exclusiva responsabilidad. Sugerencias son bienvenidas a javier.alejo@depeco.econo.unlp.edu.ar.

1. Introducción

El efecto de la educación sobre la distribución del ingreso es un tema que se ha estudiado ampliamente con distintos enfoques y metodologías. Intuitivamente se tiende a pensar que una población más educada estaría asociada a una mejor distribución de los salarios. Claramente, ambos aspectos son relevantes. Por un lado la extrema desigualdad de ingresos puede ser considerada como un aspecto que es mal visto por la sociedad y en consecuencia es contraproducente para el bienestar de la misma. Por otro lado, tanto el acceso como la cantidad de educación disponible son considerados como características indispensables desde el punto de vista de la igualdad de oportunidades. Sin embargo, la literatura que basada en microdatos de encuestas de hogares generalmente encuentra una relación positiva entre el nivel educativo de los individuos y la desigualdad de salarios predicha. Es por eso que este fenómeno ha despertado el interés de varios autores de la literatura. En particular, Bourguignon et al. (2004) han llamado a este fenómeno la Paradoja del Progreso.

Éste aspecto singular aparece en distintos trabajos que utilizan métodos de estimación y bajo diferentes interpretaciones para la relación entre los salarios y la educación. Actualmente existen dos posturas que tratan de explicar por qué se observa un efecto desigualdor de la educación. La primera está basada en el hecho estilizado de la convexidad en las ecuaciones de Mincer, mientras que la segunda toma a la heterogeneidad de los retornos de la educación como principal argumentación. Cada una de ellas está asociada con una interpretación teórica acerca del funcionamiento en mercado laboral y por lo tanto puede sugerir distintas vías de acción para morigerar la desigualdad salarial. Bajo los mismos supuestos utilizados en ambas líneas de investigación, este trabajo propone una metodología para determinar cuál de ambas visiones es la más relevante dentro de un contexto de estimaciones con microdatos de corte transversal.

La aplicación al caso de Argentina cobra relevancia al considerar que es un país donde el stock de capital humano ha crecido paulatinamente en los últimos 30 años (Gasparini, 2007 y Casal et al., 2011). Además, la historia reciente muestra que la desigualdad de ingresos ha cambiado su patrón de comportamiento hacia una distribución más igualitaria. En este contexto, conocer la importancia relativa de las fuentes redistributivas de la educación es relevante como forma darle un mejor uso a los esfuerzos de la política económica para mejorar la distribución salarial.

El trabajo se ordena de la siguiente forma: en la Sección 2 se resumen los principales aspectos teóricos detrás de la relación entre los salarios y la educación así como el estado actual de la literatura empírica; la Sección 3 propone la metodología de descomposición y una estrategia de estimación mientras que en la Sección 4 se muestra la aplicación de la misma al caso de Argentina y se discuten algunos aspectos sobre los supuestos de la metodología. Por último en la Sección 6 se presentan las conclusiones y comentarios finales del trabajo.

2. Aspectos teóricos detrás de la ecuación de Mincer

Una ecuación de Mincer es una ecuación de precios hedónicos en la cual el valor que paga el mercado por un bien o factor depende de sus características observables. En el caso del mercado laboral, el salario que recibe un trabajador depende, entre otras cosas, del tiempo que haya dedicado a capacitarse para poder realizar tareas que requieran un mayor grado de complejidad (Mincer, 1974). Una buena parte de la literatura las ha utilizado como una base para el estudio del efecto distributivo de factores considerados clave en la determinación de los salarios. En el caso de la educación como determinante de los salarios, la evidencia muestra que bajo los supuestos de un equilibrio parcial de corto plazo, la misma tiene un efecto desigualador sobre la distribución de los salarios. Algunos autores han denominado a este fenómeno como la Paradoja del Progreso (Bourgignon et al., 2004). En forma casi independiente, la literatura ha dado dos interpretaciones de lo que subyace en la Paradoja del Progreso: una basada en la convexidad de la ecuación de salarios y otra en el efecto heterogéneo de la interacción de la educación con otros factores inobservables en la determinación de salarios.

2.1 *Convexidad de los salarios con respecto la educación*

La primera línea de argumentación ha enfocado el análisis empírico de la relación entre desigualdad y educación en la media condicional. Bajo un enfoque de equilibrio parcial, el argumento principal de esta rama de la literatura es que el factor desigualador es la convexidad de la ecuación de Mincer. Una relación creciente y convexa indica que la estructura de los retornos salariales de la educación aumenta con el nivel de calificación de los individuos. Es decir, se esperaría que un año más de educación beneficie más (en términos de salario) a aquellas personas más educadas, que en general son los que tienen mejores salarios. Por lo tanto, suponiendo que la estructura de salarios se mantiene constante, un incremento en la educación llevaría a un aumento en la desigualdad salarial. Además, el tamaño del impacto dependerá del grado de convexidad en la relación salario-educación. La parte (a) de la Figura 1.1 ilustra este argumento, el eje horizontal representa los años de educación y el eje vertical mide el salario en logaritmos. Un incremento de A en la educación del individuo estará asociado a un incremento salarial que dependerá de su nivel educativo inicial. Como se observa, el incremento salarial para una persona con educación baja es B mientras que para una persona más calificada el salario esperado aumenta en C. Claramente la convexidad en la ecuación de salarios hace que $C > B$ y dado que estos incrementos están expresados en logaritmos significa que el salario aumenta más que proporcionalmente con cada año adicional de educación.

Los fundamentos teóricos sobre los distintos aspectos de las ecuaciones de salarios están bien resumidos en el trabajo de Satinger (1993), entre ellas la curvatura de la ecuación de Mincer. En los modelos del mercado laboral que admiten rentas diferenciadas (*differential rents models*), los trabajadores se asignan en distintas ocupaciones de acuerdo a su nivel calificación. La misma es observable mediante alguna característica (años de educación formal, por ejemplo). Las firmas demandan trabajadores de acuerdo a los requerimientos de su capital específico. Es decir, en estos modelos el capital físico está asociado a tareas que requieren diferentes niveles capacitación y solamente los más calificados son los que pueden trabajar con un mayor tamaño de capital. Dado que las firmas también utilizan fuerza laboral, si los trabajadores más calificados son además los más productivos el resultado de ésta asignación es una relación positiva entre los salarios y el nivel educativo. La curvatura de ésta relación depende de la distribución de las características asociadas a la calificación y de los requerimientos capital (o tipo de tareas). Si el capital presenta una mayor dispersión que la distribución de los niveles de calificación disponibles, entonces hay una escasez relativa de trabajadores más calificados. Por lo tanto el mercado estará dispuesto a pagar cada vez más por ellos y esto conlleva a una relación creciente y convexa en la ecuación de salarios. [2] También existe la posibilidad de que los requerimientos de calificación del capital físico disponible estén más concentrados en relación a la dispersión del capital humano y por lo tanto, usando la misma lógica de escasez relativa, la relación se vuelve creciente pero cóncava. En consecuencia, lo relevante en estos modelos para determinar la curvatura es la discrepancia entre la disponibilidad y la necesidad del tipo de tareas en el mercado, ya que el nivel o tamaño del capital está asociado con un conjunto de tareas específicas.

Desde un punto de vista empírico, la convexidad de las ecuaciones de Mincer es un resultado frecuente en las estimaciones con datos de corte transversal. Esto tal vez ha llevado a la conjetura de que el efecto desigualador de la educación está estrechamente asociado a esa forma funcional. El procedimiento estándar para cuantificar el efecto distributivo dentro de esta serie de trabajos consiste en realizar una simulación en la cual se le asigna un año más de educación a cada uno de los individuos y se les imputa un salario en base a una ecuación de Mincer estimada previamente. Dentro de esta literatura se encuentran las investigaciones de Bourgignon, Ferreira y Lustig (2005) y Gasparini, Battistón y García Domench (2011). La misma asume exogeneidad de los regresores y que los errores de predicción tienen una distribución homogénea alrededor de la ecuación de salarios. En otras palabras, extrapola el comportamiento de la media condicional a toda la distribución condicional y por lo tanto puede que, en caso de no cumplirse este supuesto, el cálculo del cambio en la desigualdad no mida en forma adecuada el cambio redistributivo. Cuantificar el efecto de los regresores sobre toda la distribución condicional de los salarios es el eje central del argumento que se presenta a continuación.

2.2 *Heterogeneidad de los retornos a la educación*

La otra línea de investigación sobre educación y desigualdad interpreta que pueden existir distintas ecuaciones de salarios, dependiendo de ciertos factores que son inobservables en una encuesta pero relevantes en el mercado para determinar los salarios. Becker y Chiswick (1966) interpretan que, en el extremo, cada individuo tiene un retorno específico por adquirir capital humano y por lo tanto la media condicional refleja el retorno promedio de todos los individuos. Es usual interpretar que la habilidad, la inteligencia y/o el talento es/son parte de este conjunto de factores que dan heterogeneidad a los retornos a la educación entre individuos. La idea central es que si se considera a dos individuos con las mismas características observables, aún persiste una brecha en sus remuneraciones, dado que ambos pueden tener otras habilidades laborales muy diferentes. A su vez, también es posible que esa disparidad dependa de las características observables. Por ejemplo, si existe complementariedad entre la habilidad y el capital humano es probable que el retorno de un año adicional de educación sea mayor para los más hábiles, que a su vez son los mejores pagos en el mercado. Por lo tanto, la desigualdad dentro del grupo de personas más calificadas sería mayor que dentro del grupo con menor nivel educativo.

Ésta literatura utiliza los cuantiles condicionales para modelar este aspecto de la determinación de los salarios dentro del mercado laboral. Dentro de ésta

[2] Ver también los modelos de Tinberger (1951, 1956, 1970), citados por Sattinger (1993).

lógica, los cuantiles superiores representarían a los más beneficiados por un año extra de educación. Por lo tanto, la disparidad salarial aumentaría como consecuencia de una mejora educativa. La parte (b) de la Figura 1.1 muestra este razonamiento, con distintas ecuaciones de salarios definidas por las diferentes aptitudes inobservables de los individuos. La heterogeneidad en los retornos a la educación está caracterizada por la pendiente de cada una de las ecuaciones de salarios. La brecha salarial (en logaritmos) es $D - C$ para las personas con educación A, mientras que considerando otro grupo de personas más calificadas con educación B la brecha es mayor, representada por la diferencia $F - E$. Por lo tanto, en esta visión de las ecuaciones de Mincer, mayores niveles de educación también estarían asociados a una mayor desigualdad salarial. Buchinsky (1994), Martins y Pereira (2004), Staneva et al. (2010), entre otros, son sólo una parte de ésta extensa línea de investigación. En su gran mayoría, los resultados para distintos países muestran que el efecto marginal de la educación sobre los salarios es mayor para los cuantiles superiores y por lo tanto el efecto sería ampliar la brecha salarial.

Sin embargo, debe recordarse que el método de regresión por cuantiles está diseñado para caracterizar a la distribución condicional, no a la distribución agregada de salarios. Si bien estos resultados dan indicio cualitativo de lo que puede ocurrir con la distribución total de los salarios, debe advertirse que no es una forma apropiada de medir el cambio distributivo. Sobre este punto se expone en los párrafos siguientes.

2.3 *Efecto distributivo de la educación*

Nótese que utilizando un razonamiento similar al realizado por Kuznets es posible una tercera interpretación sobre el vínculo entre la desigualdad salarial y la educación. Supongamos una relación única como la que plantea el gráfico (a) de la Figura 1.1. Si se considera que inicialmente todos los individuos tienen el nivel más bajo de educación, entonces todos tienen el mismo salario y la desigualdad es nula. A medida que algunas personas comienzan educarse para aprovechar las rentas del incremento salarial la desigualdad comienza a aumentar como consecuencia de la disparidad entre los ingresos laborales de los individuos. Finalmente, una vez que todas las personas han alcanzado el nivel máximo de educación, todos tienen un mayor salario y la desigualdad desaparece. Por lo tanto, esto genera un patrón de U invertida, tal como el planteado por Kuznets (1955). Este proceso sencillo refleja el hecho de que, aún suponiendo una única curva de salarios, la forma en que están distribuidos los individuos en los niveles educativos determina la distribución total de salarios.

El interés último de los trabajos que analizan la Paradoja del Progreso es el efecto de la educación sobre la distribución de salarios de todos los individuos, sin importar si comparten o no las mismas características observables. En otras palabras, el objetivo es caracterizar el cambio de la desigualdad medido con la distribución no condicional de los salarios. En general, la forma en que la literatura ha encarado ésta tarea es mediante el uso de metodologías para caracterizar la distribución condicional (OLS, IV, *quantile regression*, etc.). Para ver los efectos sobre la distribución no condicional se recurre a simulaciones numéricas que extrapolan el comportamiento de la estructura condicional con la nueva configuración de características. En el caso de la literatura basada en la media condicional se utiliza la predicción del modelo junto con un error que representa a los factores inobservables, bajo el supuesto de homogeneidad en su distribución. Por otro lado, la literatura de cuantiles condicionales también suele recurrir a simulaciones numéricas, aunque de mayor complejidad en su aplicación. [3] Ambas literaturas han estudiado el mismo fenómeno, pero una ha ignorado la argumentación de la otra. Por lo tanto, no se ha determinado aún cuál de los dos argumentos es más relevante para explicar el efecto distributivo de la educación.

Firpo et al. (2009) construyen una metodología que permite computar directamente el efecto marginal de los regresores de una ecuación de salarios sobre cualquier funcional (indicador) de la distribución no condicional de los salarios. Además de otras bondades prácticas, ese método tiene una clara ventaja en medir el efecto sobre la desigualdad medida con indicadores complejos y muy utilizados en la literatura, tal como el índice de Gini. [4] Sin embargo, el beneficio del método en términos de obtener una medición más precisa y la generalidad en la admisión de indicadores, tiene un costo en términos de explicar cuáles son los canales por el cual la educación impacta sobre la desigualdad. En otras palabras, Firpo et al. (2009) es la metodología adecuada si el interés reside en medir correctamente el efecto sobre la desigualdad utilizando una gran variedad de indicadores (Gini, deciles, Theil, etc.). Sin embargo, no permite determinar si el efecto marginal observado proviene de una relación convexa o bien de la heterogeneidad de los mismos. El aporte de este trabajo es presentar una metodología que permite cuantificar la relevancia de las dos argumentaciones esbozadas en la literatura empírica en un contexto no condicional. El sacrificio de la descomposición propuesta es la potencial pérdida de generalidad dado que utiliza un indicador particular de la desigualdad agregada como es la varianza de los logaritmos. Sin embargo, como se verá en la sección de resultados, puede que en la aplicación utilizada ese costo no sea excesivo.

3. Metodología

En esta sección se presenta la metodología de descomposición para medir el peso que tiene las dos explicaciones alternativas de la Paradoja del Progreso, dentro de la literatura que utiliza datos de corte transversal. La misma se basa en la re-parametrización de un modelo de cuantiles condicionales propuesta por Machado y Mata (2005) y en la aplicación de la ley de varianzas iteradas. De la misma forma que gran parte de ésta literatura, se asume exogeneidad de los regresores y que el conjunto de datos a utilizar conforman una muestra aleatoria de la población, dejando para la Sección 4 la discusión estos puntos. También en línea con la literatura, se supondrá en todo el análisis que no hay efectos en el equilibrio general, es decir los cambios generados como ejercicios de estática comparada no producen cambios en la forma en que el mercado retribuye a los factores. En otras palabras, las estimaciones no incluyen los efectos indirectos de cambiar la educación sobre otros mercados. La interpretación usual es que se trata de un análisis de corto plazo. En términos analíticos, este supuesto permite utilizar la distribución condicional estimada para luego aplicarle cambios en sus regresores y obtener así el efecto sobre toda la distribución de salarios (Firpo et al., 2011).

Si bien en la literatura sobre la desigualdad del ingreso el indicador por excelencia es el índice de Gini, resulta analíticamente conveniente el uso de la varianza de los logaritmos para realizar la descomposición. En la sección de

[3] Ver por ejemplo la metodología basada en integración por Monte Carlo propuesta Machado y Mata (2005).
[4] Ver Alejo, Gabrielli y Sosa Escudero (2011).

resultados se verá que en el caso de Argentina la evolución de ambos indicadores es muy similar.

3.1 *Modelo poblacional*

Supóngase la siguiente ecuación para la relación entre el cuantil condicional de los salarios y los atributos de un individuo:

$$w = x'\alpha(\tau) \tag{3.1}$$

donde $\tau \mid x \sim U(0,1)$, w es el logaritmo del salario y x es un conjunto de atributos observables en una encuesta (educación, experiencia, género, localización geográfica, etc.). El índice τ indica el ranking que ocupa la persona en la distribución salarial, condicional en sus características observables. [5] Este ranking condicional está determinado por el conjunto de atributos que determinan el salario, pero que no son directamente observables (habilidad, inteligencia, motivación, etc.). Por lo tanto, la forma funcional de los coeficientes (cuyo argumento es τ) está determinada por la distribución de este conjunto heterogéneo de características. Los mismos pueden interpretarse como una medida indirecta del efecto que tiene la interacción de las variables observables con el conjunto agregado de inobservables.

Como se explicó en la sección anterior, la literatura sobre desigualdad salarial que comienza con Buchinsky (2005) se ha concentrado en analizar el efecto marginal de la educación sobre los coeficientes de los cuantiles condicionales, encontrando que los mismos tienen un patrón creciente en el índice τ. Es decir, si la habilidad/inteligencia/motivación de un individuo hace que sea más fácil adquirir más años de educación y ambos factores son valuados en el mercado, entonces es lógico que se observe que un año adicional de educación formal tenga un retorno mayor para aquellos individuos que se encuentran en los cuantiles superiores de la distribución condicional de salarios. Por otro lado, la línea de investigación iniciada en Bourgignon, Ferreira y Lustig (2005), modela la convexidad en la ecuación de Mincer incluyendo un polinomio de segundo grado en la educación.

Sea h una medida continua del capital humano (años de educación) y z el resto de los determinantes observables del salario. Si la educación al cuadrado es parte del modelo, entonces $x = [1, h, h^2, z]$ es el vector de regresores. Dado que una rama de la literatura enfoca el análisis sobre la media condicional, a continuación se hace una representación alternativa del mismo modelo siguiendo a Machado y Mata (2005). Sumando y restando $x'\beta$ en el lado derecho de (3.1) la relación entre la educación y los salarios dada por la ecuación (3.1) se puede rescribir de la siguiente manera:

$$w = x'\beta + x'\gamma(\tau) \tag{3.2}$$

donde $\tau \mid x \sim U(0,1)$, β son los parámetros de la esperanza condicional y $\gamma(\tau) \equiv \alpha(\tau) - \beta$. Es decir, los parámetros $\gamma(\tau)$ miden la diferencia entre el efecto marginal sobre la media y el τ-ésimo cuantil de la distribución condicional de salarios. En la terminología de Machado y Mata, si se agrupase a los individuos que comparten los mismos valores de x, entonces β son los parámetros asociados a la desigualdad *between* mientras que $\gamma(\tau)$ mide un componente de desigualdad *within*.

[5] El hecho de que τ sea uniforme proviene de la aplicación del método de la transformada inversa para representar variables aleatorias a través de sus cuantiles. El supuesto implícito es que $F(w|x)$ es una función continua (Devroye, 1986).

Con ésta especificación es posible capturar las dos fuentes que explican potencialmente lo que se ha denominado la Paradoja del Progreso. Para ello hay que tener presente que el objeto bajo análisis es el efecto de la educación sobre el nivel de agregado (no condicional) de desigualdad. Como primer paso, una forma conveniente de hacer esto es tomar como indicador de desigualdad I a la varianza de los logaritmos y aplicar la ley de varianzas iteradas.

$$I \equiv Var(w) = Var[E(w \mid x)] + E[Var(w \mid x)] \tag{3.3}$$

Luego, utilizando (3.2) para calcular $E(w|x)$ y $Var(w|x)$ se obtiene la siguiente descomposición de la desigualdad salarial:

$$I = \beta' V\beta + [tr(\Omega V) + E'\Omega E] \tag{3.4}$$

donde la matriz V contiene las varianzas y covarianzas de las variables incluidas en x, mientras que E es el vector que contiene las esperanzas de x. Los elementos de la matriz $\Omega \equiv Var[\gamma(\tau)]$ son parámetros que miden la distancia entre los coeficientes de los cuantiles condicionales y la media condicional. [6] Dentro de la lógica de Machado y Mata (2005), claramente se ve que la desigualdad agregada es la suma de dos componentes: uno que depende de los parámetros β (*between*) y otro de la dispersión de los parámetros $\gamma(\tau)$ (*within*).

3.2 *Descomposición del efecto marginal de la educación*

El hecho de que la expresión (3.4) incluya a E y V muestra en forma explicita que la desigualdad no solo depende de los parámetros de la distribución condicional de los salarios sino que también depende de la forma en que están distribuidos los regresores. Por lo tanto, para realizar un análisis marginal del efecto de la educación sobre la dispersión salarial es necesario definir la forma en la cual se mueve esta distribución. Siguiendo a Fortín et al. (2011) se asume una pequeña traslación horizontal (*location shift*) en la distribución de los años de educación *h*. En otras palabras, se suma un número ε a cada valor posible de *h* y se calcula el efecto de la misma sobre la desigualdad, en este caso medida por la varianza de los logaritmos.

El cambio o derivada funcional de un indicador será denotado con la letra δ. A modo de ejemplo, tómese como indicador el momento de orden *k* de la variable de educación *h*, es decir $E_k = E(h^k)$. En ese caso, la derivada funcional de E_k con respecto a una traslación horizontal en la distribución de *h* se define como:

$$\delta(E_k) \equiv \lim_{\varepsilon \to 0} \frac{E[(h+\varepsilon)^k] - E[h^k]}{\varepsilon}$$

Si la distribución del salario condicional en las x se mantiene constante (*i.e.* los parámetros β y γ no cambian), es fácil ver que el cambio en la desigualdad como consecuencia de la traslación horizontal en *h* es:

$$\delta(I) = \beta'\delta(V)\beta + tr[\Omega\delta(E)] + 2E'\Omega\delta(E) \tag{3.5}$$

donde $\delta(E)$ es un vector que contiene la derivada de los primeros momentos de x ante una traslación horizontal en *h*, mientras que $\delta(V)$ una matriz que contiene el

[6] Ver la sección A.1 del Apéndice para los detalles en la obtención de la ecuación (3.4).

cambio en las varianzas y covarianzas de x originadas por la misma traslación en h (ver la sección A.2 del Apéndice).

De la ecuación (3.5) puede verse que el cambio en la desigualdad se compone de dos factores: un efecto que depende de la media condicional (efecto *between*) y otro que depende de los cuantiles (efecto *within*).

$$EF_{betw} = \beta'\delta(V)\beta$$

$$EF_{with} = tr[\Omega\delta(V)] + 2E'\Omega\delta(E)$$

de forma tal que $\delta(I) = E_{betw} + E_{with}$.

En otras palabras, por un lado se tiene que la educación influye sobre la desigualdad al modificar la brecha salarial promedio entre los distintos grupos educativos (medido por la media condicional). Por otro lado, hay un efecto adicional que considera la relación entre el nivel de capital humano y la desigualdad al interior de cada grupo, que depende de la varianza condicional a computada a través de los cuantiles.

Considérese un caso sencillo para entender cómo se relaciona la convexidad en la relación entre *w* y *h* con el efecto marginal de la educación. Sin pérdida de generalidad, supóngase que el caso en donde x solamente incluye a las variables educativas. En ése caso, la ecuación (3.2) queda:

$$w = \beta_0 + \beta_1 h + \beta_2 h^2 + \gamma_0(\tau) + \gamma_1(\tau)h + \gamma_2(\tau)h^2$$

Utilizando los parámetros de este modelo para calcular los efectos *between* y *within* se obtiene la siguiente expresión:

$$EF_{betw} = 4(\beta_1 V_{11} + \beta_2 V_{12})\beta_2$$

$$EF_{with} = 2[\Omega_{01} + 2\Omega_{02}E_1 + 3\Omega_{12}E_2 + \Omega_{11}E_1 + 2\Omega_{22}E_3]$$

Claramente, la convexidad juega un rol crucial en el efecto distributivo proveniente del salario promedio condicional (*between*), pero no necesariamente sobre los cuantiles condicionales (*within*). Es decir, pueden existir modelos que son lineales en la educación ($\beta_2 = 0$) y en los cuales una traslación horizontal de *h* lleva a un cambio en la desigualdad que no pasan por el efecto de la media condicional. Tres casos extremos ilustran este punto:

Caso 1 (*modelo lineal homocedastico*): $\beta_2 = 0$ y Ω tiene elementos nulos a excepción de $\Omega_{00} > 0$ (varianza de la ordenada al origen del cuantil condicional). En este caso, cualquier traslación horizontal en *h* no perturba a I, es decir:

$$EF_{betw} = EF_{with} = 0$$

Caso 2 (*modelo cuadrático homocedástico*): $\beta_2 > 0$ y Ω tiene elementos nulos a excepción de $\Omega_{00} > 0$. En este caso, todo el efecto distributivo se encuentra exclusivamente en la convexidad de la media condicional.

$$EF_{betw} = 4(\beta_1 V_{11} + \beta_2 V_{12})\beta_2 > 0$$

$$EF_{with} = 0$$

Caso 3 (*modelo lineal heterocedástico*). $\beta_2 = 0$ y Ω tiene elementos nulos fuera de la diagonal principal. [7] En este caso, todo el efecto distributivo de h es explicado por la heterocedasticidad, tanto en su efecto lineal (Ω_{11}) como cuadrático (Ω_{22}).

$$EF_{betw} = 0$$
$$EF_{with} = 2\Omega_{11}E_1 + 4\Omega_{22}E_3 > 0$$

El aporte de ésta descomposición del efecto marginal de la educación sobre la desigualdad es que permite cuantificar el peso relativo de las dos explicaciones esbozadas en la literatura empírica en forma separada. Por un lado, el efecto *between* está asociado a la rama de la literatura se ha enfocado en la convexidad de la media condicional como explicación de la paradoja del progreso (Bourgignon et al., 2005; Gasparini et al., 2011, etc.), mientras que el efecto *within* ofrece una buena medida para cuantificar el rol que juega la heterogeneidad no observada en el cambio de la desigualdad, tal como se analiza en la literatura de regresión por cuantiles (Buchinsky, 1994; Martins y Pereira, 2004, etc.). En consecuencia, es factible ver cuál de las dos fuentes desigualadoras tiene mayor peso sobre el cambio total en la varianza ante una traslación horizontal de los años educación.

3.3 *Estimación*

Sea (w_i, x_i) con $i = 1, \ldots, n$ una muestra de n asalariados, donde x_i incluye a h_i (la educación del individuo i) y su cuadrado, junto con un conjunto de otras características observables z_i. Suponiendo regresores exógenos es posible obtener estimadores consistentes de cada uno de los elementos de la descomposición (3.5).

En primer lugar, los parámetros β de la media condicional pueden ser estimados por el método de mínimos cuadrados ordinarios (OLS),

$$\hat{\beta} = \arg\min_{b} \Sigma_{i=1}^{n}(w_i - x_i'b)^2$$

mientras que el método de regresión por cuantiles (QR) condicionales (Koenker y Basset, 2005) puede utilizarse para estimar los parámetros $\alpha(\tau)$.

$$\hat{\alpha}(\tau) = \arg\min_{a} \Sigma_{i=1}^{n}\rho_\tau(w_i - x_i'a) \qquad \text{, donde } \rho_\tau(u) = \begin{cases} (\tau - 1) & si \quad u < 0 \\ \tau & si \quad u \geq 0 \end{cases}$$

Bajo el supuesto de exogeneidad en los regresores, ambos métodos son consistentes y con distribución asintótica normal.

Para calcular la descomposición (3.5) es necesario contar con una estimación de Ω, la matriz de varianzas y covarianzas de $\gamma(\tau) = \alpha(\tau) - \beta$, donde τ sigue una distribución uniforme entre 0 y 1. Si $\Omega < \infty$, una opción factible es estimar $\alpha(\tau)$ para una grilla de cuantiles $\tau_1, \ldots, \tau_M$ y luego calcular:

$$\hat{\Omega} = (M-1)^{-1}\Sigma_{m=1}^{M}[\hat{\alpha}(\tau_m) - \hat{\beta}]\cdot[\hat{\alpha}(\tau_m) - \hat{\beta}]'$$

[7] Este es sólo un caso sencillo de heterocedasticidad, el modelo admite otras variantes según las combinaciones de los parámetros de la matriz Ω.

Si el número de cuantiles M es grande se esperaría que este estimador sea una buena aproximación a Ω. [8] Obviamente, dada la consistencia de los métodos OLS y QR, la aproximación de Ω será cada vez más adecuada si el tamaño de la muestra n es grande.

Finalmente, es necesario estimar el cambio en E y V como consecuencia de la traslación horizontal en los años de educación. Siendo Q el número de regresores incluidos en z, la sección A.4 del Apéndice muestra que la forma analítica para $\delta(E)$ y $\delta(V)$ es:

$$\delta(E)=\begin{bmatrix}0\\1\\2E_1\\0_{1\times Q}\end{bmatrix} \quad \text{y} \quad \delta(V)=2\begin{bmatrix}0 & 0 & 0 & 0_{1\times Q}\\ 0 & 0 & V_{11} & 0_{1\times Q}\\ 0 & V_{11} & 2V_{12} & M_{1z}\\ 0_{Q\times 1} & 0_{Q\times 1} & M_{z1} & 0_{Q\times Q}\end{bmatrix} \tag{3.6}$$

con $E_1 = E(h)$, $V_{12} = COV(h, h^2)$, $M_{1z} = COV(h,z)$ (vector de 1xQ) y donde 0_{AxB} indica una matriz nula de dimensión AxB. [9] Todos estos componentes son momentos poblacionales y por lo tanto pueden ser estimados consistentemente utilizando sus análogos muestrales:

$$\hat{E}=(n-1)^{-1}\Sigma_{i=1}^{n}x_i$$
$$\hat{V}=(n-1)^{-1}\Sigma_{i=1}^{n}(x_i-\hat{E})(x_i-\hat{E})'$$

Luego, la estimación de la descomposición en cambio de la desigualdad consiste en reemplazar cada elemento estimado en (3.5), es decir:

$$\hat{\delta}(I)=\hat{\beta}'\delta(\hat{V})\hat{\beta}+tr[\hat{\Omega}\delta(\hat{E})]+2\hat{E}'\hat{\Omega}\delta(\hat{E}) \tag{3.7}$$

donde $\delta(\hat{E})$ y $\delta(\hat{V})$ son construidas según (3.6).

El primer sumando de (3.7) es la estimación del efecto *between* producto de la convexidad en los retornos a la educación, mientras que el resto corresponde al efecto *within* como consecuencia de la heterogeneidad en los retornos al capital humano. Por lo tanto, es posible separar y cuantificar la relevancia de ambas explicaciones del efecto desigualador de la educación.

En cuanto a las propiedades asintóticas de ésta estimación, todos los componentes surgen de métodos consistentes y por lo tanto, por el teorema de mapeo continuo, la estimación del cambio en la desigualdad también es consistente. Por otro lado, el uso de varios métodos de estimación para calcular el cambio en la desigualdad implica una cierta complejidad analítica para calcular la distribución asintótica de (3.7) que excede a este trabajo. Es por ello que se optó por el método de bootstrap para realizar la inferencia estadística de los valores estimados.

El comportamiento de la estimación con muestras finitas se evaluó con un experimento de Monte Carlo. [10] Con un tamaño de muestra de 500 observaciones y distintas configuraciones de convexidad y heterocedasticidad, los resultados muestran que el error cuadrático medio del efecto *between* está explicado casi en su

---

[8] Portnoy (1991) muestra que la cantidad de cuantiles (estimados por QR) que son numéricamente diferentes es $O(n\log(n))$.

[9] Lógicamente, M_{z1} es la transpuesta de M_{1z}, por simetría de las covarianzas.

[10] Ver la sección A.4 del Apéndice para más detalles sobre el diseño experimental.

totalidad por la varianza muestral. Es decir, en promedio la estimación es cercana a su valor poblacional. Por otro lado, la estimación del efecto *within* presenta un mayor componente de sesgo, siendo aproximadamente el 25% del error cuadrático medio (ver Tabla 3.1). Adicionalmente, se compara el estimador (3.6) con la estimación del efecto marginal de h sobre la desigualdad salarial que surge de hacer una simulación simple como se hace usualmente en la literatura: moviendo la distribución de h a la derecha (*location shift*) y re-computando los salarios con los parámetros de una Mincer estimada por OLS a la que se le agrega el error de predicción de cada observación. [11] Como era de esperar, bajo homocedasticidad un ejercicio de simulación simple que contemple la convexidad en la media condicional funcionará bien en promedio; mientras que en presencia de heterogeneidad no observada el estimador de cuantiles es superior, dado que se reduce notablemente el sesgo (ver Tabla 3.2). El Gráfico 3.1 ilustra este punto: los ejes representan la diferencia entre el cambio en la desigualdad estimado por algún método y el cambio poblacional (es decir, el sesgo). El eje horizontal corresponde a la estimación de la ecuación (3.7) mientras que el vertical a la simulación numérica. Los puntos sobre el plano son los resultados de cada uno de los 500 experimentos. Como se observa, en ausencia de heterogeneidad (columna izquierda) no hay mayor problema en ambos métodos dado que el centro de la nube de puntos es el origen del plano, es decir en promedio ambos tienen sesgo cercano a cero. Sin embargo, cuando se considera una población con retornos diferenciales (columna derecha), la nube de puntos desciende, mostrando que en promedio la simulación numérica estará sesgada (subestimando el cambio en la desigualdad) mientras que las estimaciones por cuantiles siguen teniendo como centro al cero.

4. Resultados

Para la implementación de la metodología descomposición al caso de Argentina se utilizó la Encuesta Permanente de Hogares (EPH) elaborada por el Instituto Nacional de Estadísticas y Censos (INDEC). A modo de explorar tres periodos diferentes la descomposición se realiza para tres años relativamente distantes: 1992, 1998 y 2008. Los datos corresponden a las encuestas recolectadas durante el segundo semestre. Los primeros dos años pertenecen a la versión puntual de la EPH mientras que el último corresponde a la metodología continua de relevamiento. Dado que la cobertura se ha ido ampliando en el tiempo, a modo de mantener comparable las estimaciones se restingue las observaciones que pertenecen a los aglomerados presentes en la EPH de 1992. [12] La muestra utilizada son hombres entre 16 y 65 años de edad. Tanto la muestra como los años elegidos son exactamente los mismos que los utilizados en Alejo, Gabrielli y Sosa Escudero (2011).

A modo ilustrativo, la Tabla 4.1 presenta las estimaciones de los parámetros de las variables de educación de las ecuaciones de Mincer para la media condicional y algunos cuantiles. [13] Las estimaciones incluyen además otros regresores que son los usuales en la literatura. Como se observa, dado el resto de las variables el término cuadrático es estadísticamente significativo en los tres años considerados y en distintos puntos de la distribución condicional. Esto significa que, dado el resto de

[11] Para hacer comparable el ejercicio numérico con la noción de derivada, la traslación considerada es de h a $h' = h + 0.01$.

[12] Los aglomerados considerados son Gran La Plata, Gran Santa Fe, Gran Paraná, Comodoro Rivadavia - Rada Tilly, Gran Córdoba, Neuquén - Plottier Santiago del Estero - La Banda, Jujuy - Palpalá, Río Gallegos, Salta, San Luis - El Chorrillo, Gran San Juan, Santa Rosa - Toay, Ushuaia - Río Grande, Ciudad de Buenos Aires, Partidos del GBA.

[13] Como se verá en los párrafos siguientes, para la descomposición se utiliza un número mucho más amplio de cuantiles.

las variables constantes, la relación entre los salarios y el nivel educativo es marcadamente convexa. Desde el punto de vista de los modelos de rentas diferenciadas, esto sería un indicio de que la disparidad en los requerimientos del capital físico es mayor que la que puede ofrecer los individuos en los distintos niveles educativos. Por otro lado, la diferencia en el valor de los parámetros al considerar los distintos cuantiles condicionales muestra que también hay evidencia de un patrón heterogéneo en los retornos a la educación. Por ejemplo, al considerar el comportamiento del coeficiente del término cuadrático entre deciles se ve que la diferencia es alta en 1992 (el último decil es 3 veces el primero), leve en 1998 (1.2 veces) y practicante nula en 2008. El Gráfico 4.1 ilustra este punto mostrando las distintas ecuaciones de Mincer para un caso promedio. Como se observa, la convexidad pareciera ser más marcada en 1998, mientras que el patrón de heterogeneidad en los retornos ha ido desapareciendo a lo largo de la última década y media.

Como fuese discutido en la Sección 2, tanto la convexidad como la heterogeneidad son dos factores clave detrás del efecto marginal de la educación sobre la desigualdad salarial. La idea de ésta investigación es tratar de medir la importancia relativa de cada uno de ellos. La razón por la cual no basta con las estimaciones del párrafo anterior es que a partir de las mismas se necesita computar el cambio en la distribución no condicional, utilizando la distribución de las distintas características observables de los individuos. Para ello se utilizará la metodología de descomposición propuesta en la Sección 3.

La Tabla 4.2 resume los principales resultados del trabajo. En el primer bloque se muestra la evolución de la desigualdad salarial para los hombres de 16 a 65 años de edad, medida tanto con el coeficiente de Gini como por la varianza de los logaritmos. Como se observa, ambos indicadores muestran una evolución semejante: aumento de la desigualdad hacia fines de los 90 y una mejora distributiva cerrando la primera década de los 2000. Por lo tanto, esto puede ser un punto a favor de la descomposición propuesta dado que no se estaría perdiendo demasiada generalidad al analizar el comportamiento distributivo mediante la varianza de los logaritmos.

El segundo segmento muestra un ejercicio de simulación numérica en línea con la literatura basada en la media condicional. El mismo consiste en utilizar la ecuación de Mincer estimada por OLS junto con sus errores de predicción para re-computar una nueva distribución de salario luego de una pequeña traslación horizontal en los años de educación. Es decir, si un individuo i tiene un nivel de educación de h_i, el ejercicio consiste en asignarle $h_i+\varepsilon$ años de educación y construirle un salario contrafactual (en logaritmos) w^s_i a través de la media condicional. Luego, la variación marginal en la desigualdad como consecuencia del cambio en h se calcula como $[V(w^s) - V(w)]/\varepsilon$, donde ε es un valor pequeño.[14] El siguiente bloque en la Tabla 4.2 es la estimación del cambio en la varianza utilizando la metodología de Fortín et al (2009) a través de regresiones RIF. Ambas metodologías muestran que el efecto de la educación sobre la distribución de salarios es desigualador y que tuvo su máxima expresión a fines de los 90. La diferencia entre una y otra estimación reside en que la estimación RIF es una aproximación lineal del efecto marginal que contempla todos los canales condicionales de la relación el salario y la educación, mientras que en la simulación solo usa información de la media condicional.

Para saber qué parte del efecto desigualador corresponde a la convexidad o a la heterogeneidad en los retornos a la educación se utiliza la metodología de descomposición. La misma se muestra en el cuarto segmento de la Tabla 4.2. Para las regresiones por cuantiles se utilizó una grilla amplia de valores para $\tau = \{0.005,$

[14] El valor utilizado para éste trabajo es $\varepsilon = 0.01$.

$0.01, \ldots, 0.99, 0.995\}$, implicando un total de $M = 199$ estimaciones. [15] Como se observa, el efecto total se aproxima bastante a la estimación realizada por el método RIF. El aporte de la descomposición es que permite ver de una forma explicita el aporte relativo de las explicaciones basadas tanto en la media condicional como con los cuantiles condicionales. Como se observa en el la última parte del tabulado, a comienzos de los 90 ambos argumentos tenían la misma relevancia como fuente desigualadote de ingresos. Sin embargo, la explicación basada en la convexidad de los retornos esperados se vuelve más relevante en 1998 (poco más del 70%) y pareciera ir creciendo hacia fines de los 2000. Este cambio en la relevancia de las dos fuerzas desigualadoras podría dar lugar a algunas interpretaciones de cómo ha cambiado el funcionamiento del mercado laboral y poner el foco de atención en la más relevante a la hora de diseñar políticas redistributivas basadas en mejoras educativas.

Es importante cuestionar si estos resultados son demasiado sensibles a los supuestos sobre los que se basan los estimadores: exogeneidad de los regresores y muestra aletoria. En el primer caso, las consecuencias de la endogeneidad sobre los retornos a la educación han sido arduamente estudiadas para estimaciones de la media condicional, siendo Card (2001) la principal referencia. Esto afectaría a las estimaciones de la convexidad y por lo tanto al efecto denominado *between*. Por otro lado, el estudio de las consecuencias de la endogeneidad sobre los cuantiles condicionales es relativamente más nuevo e incluso hay pocos trabajos que exploren sus consecuencias en las estimaciones de *cross-setcion*. Algunos ejemplos son Arias et al. (2001) y Staneva et al. (2010). Todas estas metodologías necesitan del uso de instrumentos para corregir el potencial sesgo por endogeneidad, variables que no se encuentran disponibles en las EPHs utilizadas en este trabajo. [16] En general, los resultados de esas aplicaciones indican que es la media condicional la que se ve mayormente afectada por la corrección de variables instrumentales, no así el caso de los cuantiles. Si bien son alternativas consistentes, en general estos métodos son bastante ineficientes y por lo tanto no presentan diferencias estadísticamente significativas con las estimaciones estándar. En el caso Arias et al., (2001), la consideración de la endogeneidad en la educación solo produce un cambio en el nivel de los coeficientes, pero no cambia el patrón de los cuantiles condicionales. Por lo tanto se esperaría que la corrección por variables instrumentales, en caso de que se contara con ellas, no debería afectar demasiado el componente *within* de la descomposición.

Finalmente, el sesgo por selección en la muestra es otro aspecto usualmente atribuido a la estimación de ecuaciones de Mincer, especialmente para el grupo de mujeres dado que tiene una baja participación laboral. Si bien en este trabajo se usa una muestra de hombres en edades activas, cuestionar la aleatoriedad en la muestra es una pregunta válida. Nuevamente, los métodos para corregir la estimación de la media condicional son los más difundidos en comparación a la literatura de cuantiles. La Tabla 4.3 muestra las estimaciones con y sin corrección por sesgo de selección de los parámetros asociados a los años de educación, con la misma muestra de hombres utilizada para la descomposición. Tanto para la media condicional como para los cuantiles se utilizó los métodos en dos etapas actualmente disponibles en la literatura: Heckman (1979) para la media y Buchinsky (2001) para los cuantiles. Los errores estándar de los estimadores, así como el de la diferencia entre ellos fueron calculados por un bootstrap de 500 replicas. Como se observa, la diferencia en la estimación de la media condicional como consecuencia de la corrección por sesgo de selección no difiere sistemáticamente de la estimación

[15] Este valor es similar a los considerados por Melly (2005) en ejercicios realizados para generar distribuciones contrafactuales con regresiones por cuantiles.

[16] En general, las variables instrumentales utilizadas son distancia al centro de estudios (universidad, colegio, etc.) o con bases de datos particulares, tal como las encuestas a gemelos.

de OLS. La evidencia es menos clara en el caso de los cuantiles condicionales, con algunas diferencias sistemáticas en 1992 y 2008, pero en general se observa que no hay mayor discrepancia con la estimación estándar. Vale aclarar, que al igual de lo ocurre con corrección por variables instrumentales, las estimaciones que incluyen una segunda etapa son estadísticamente más ineficientes. Por lo tanto, realizar las correcciones por sesgo de selección puede cambiar algunos valores en la descomposición, pero se esperaría que los cambios no sean demasiado drásticos.

5. Conclusiones

Este trabajo presentó una metodología de descomposición para el efecto marginal de la educación sobre la desigualdad salarial. El procedimiento propuesto permite cuantificar la relevancia de dos argumentaciones empíricas esbozadas en la literatura para explicar lo que se ha denominado la Paradoja del Progreso. La primera de ellas está basada en la convexidad de la ecuación de Mincer interpretada como una media condicional, mientras que la otra pone el foco de atención en el efecto heterogéneo de los retornos a la educación. El método utiliza las mismas técnicas de estimación empleadas en ambas literaturas (OLS y QR) junto con la noción de derivada funcional (recientemente propuesta por Firpo et al. 2009). Los supuestos utilizados son estándar dentro de la literatura que analiza la desigualdad salarial con microdatos.

La implementación de la descomposición al caso Argentino muestra que, bajo los supuestos del corto plazo, a principios de los ‘90 tanto la convexidad como la heterogeneidad tenían el mismo peso en el efecto desigualador de la educación sobre los salarios. Sin embargo, el primer efecto ha ganando terreno hacia fines de la primera década de los 2000. Ésta evidencia sobre la mayor importancia del efecto distributivo a través de la media condicional en relación lo que ocurre en otros puntos de la distribución condicional mostraría que el mercado ha ido perdiendo interés en pagar en forma diferencial los atributos no observables. Todo este proceso coincide con un período marcado por una clara mejora en la distribución de los ingresos. Por otro lado, la curvatura de la ecuación Mincer aparece como el factor más influyente sobre la distribución de salarios, hecho relacionado con el efecto de la escasez relativa entre los requerimientos del capital y la oferta de trabajadores calificados.

Desde la óptica de las teorías mencionadas, esto significa que existe una escasez relativa en la diversidad de la oferta educativa que no llega a cubrir las necesidades requeridas para el uso del stock de capital existente. Esto pareciera ser clave en el caso argentino, donde se ha registrado un avance educativo importante en términos de años de educación formal. Una mejora en el *matching* entre el nivel educativo de los trabajadores y el stock de capital específico que morigere ésta escasez relativa de reduciría el retorno adicional que presentan aquellos ocupados con un nivel de calificación superior. Finalmente, se debe advertir al lector que los resultados encontrados no son más que una guía a tener en cuenta, construida sobre una serie de supuestos de equilibrio parcial utilizados ampliamente en la literatura. El uso de los mismos surge de una necesidad de dar alguna respuesta al problema distributivo, en ausencia de técnicas y datos más adecuados que permitan levantar esos supuestos, quedando relegado para investigaciones futuras.

Referencias

Alejo, J., Gabrielli, F. y Sosa Escudero, W. (2011). “The Distributive Effects of Education: An Unconditional Quantile Regression Approach”. En Anales de la

AAEP, ISSN: 1852-0022, ISBN: 978-987-99570-9-7. (también como Documento de Trabajo CEDLAS N°125).

Arias, O., Hallock K. y Sosa Escudero, W. (2001). "Individual heterogeneity in the returns to schooling: instrumental variables quantile regression using twins data". Empirical Economics, 2001, Volume 26, Number 1, Pages 7-40.

Aysit, T. y Bircan, F. (2010). "Wage Inequality and Returns to Education in Turkey: A Quantile Regression Analysis" IZA Discussion Paper No. 5417.

Becker, G. y Chiswick, B. (1966) “Education and the Distribution of Earnings”. The American Economic Review. Vol. 56, No. 1/2, pp. 358-369. American Economic Association.

Bourguignon, F., N. Lustig, y F. Ferreira (2004): The Microeconomics of Income Distribution Dynamics. Oxford University Press, Washington.

Buchinsky, M. (1994). "Changes in the U.S. Wage Structure 1963-1987: Application of Quantile Regression," Econometrica, 62(2), 405-458.

Buchinsky, M. (2001). "Quantile regression with sample selection: Estimating women’s return to education in the U.S." Empirical Economics, 26: 87-113.

Casal, M., Morales, M. y Paz Terán, C. (2011). “Educational Inequality in Argentina: 1970-2010”. Anales de la AAEP, Mar del Plata 2011.

Card, D. (2001). "Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems," Econometrica, 69, 1127-1160.

Devroye, L. (1986) “Non-Uniform Randor Variable Generation”. Springer-Verlag, New Inc. Capítulo 2.

Falaris, E. (2008) "A Quantile Regression Analysis of Wages in Panama". Review of Development Economics, 12(3), 498-514.

Fersterer J. y Winter-Ebmer, R. (2003). "Are Austrian Returns to Education Falling over Time?", Labour Economics 10(1): 73-89.

Firpo, S., Fortin, N. y Lemieux, T. (2009) "Unconditional Quantile Regressions," Econometrica, 77(3), 953-973. (2011) Handbook of Labor Economicschap. Decomposition Method in Economics. Elsevier, en prensa.

Fiszbein, A., Giovagnoli, P. y Patrino, H (2007) "Estimating the Returns to Education in Argentina using Quantile Regression Analysis: 1992-2002". Económica, Vol. LIII, Nro. 1-2.

Galiani, S. y Titiunik, R. (2005). "Changes in the Panamian wage structure: a quantile regression análisis". Económica, Vol. LI, Nro. 1-2.

Gasparini, L. (2007): Monitoring the Socio-Economic Conditions in Argentina 1992-2006. World Bank and CEDLAS Working Paper.

Gasparini, L., Battiston, D y García Domench, C. (2011). “Could an Increase in Education Raise Income Inequality? Evidence for Latin America”. Anales de la AAEP, Mar del Plata 2011.

Gonzalez, X. y Miles, D. (2001). "Wage Inequality in a Developing Country: Decrease in Minimum Wage or Increase in Education Returns", Empirical Economics 26(1): 135-148.

Jacob A. Mincer, 1974. "Schooling, Experience, and Earnings". National Bureau of Economic Research, Inc. NBER Books Series.

Heckman (1979). Sample selection as a specification error, Econometrica, 47: 153-161, 1979.

Koenker, R. (2005): Quantile Regression. Cambridge University Press, Cambridge. 15.

Kuznets S. (1955) "Economic growth and income inequality". American Economic Review 45, N° 1.

Martins, P.S. y Pereira, P.T. (2004). "Does Education Reduce Wage Inequality? Quantile Regressions Evidence from 16 Countries", Labour Economics 11(3): 355-371.

Mata, J., y J. Machado (2005): "Counterfactual decomposition of changes in wage distributions using quantile regression," Journal of Applied Econometrics, 20(445-465).

Melly, B. (2005): "Decomposition of Differences in Distribution Using Quanitle Regressions," Labour Economics, 12, 577-90.

Portnoy, S. (1991): “Asymtotic Behavior of the Number of Regression Quantile Breakpoints”, SIAM Journal of Scientific and Statitistical Computing, 12, 867-883.

Sattinger, M. (1993) "Assignment Models of the Distribution of Earnings". Journal of Economic Literature. Vol. 31, No. 2, pp. 831-880. American Economic Association.

Staneva, A., Arabsheibani, R. & Murphy, P. (2010). "Returns to Education in Four Transition Countries: Quantile Regression Approach," IZA Discussion Papers 5210, Institute for the Study of Labor (IZA).

Tinbergen, J. "The Impact of Education on Income Distribution," Rev. Income Wealth, Sept. 1972, 18(3), pp. 255-65.

Wambugu, A. (2002). "Real Wages and Returns to Human Capital in Kenya Manufacturing Firms", Göteborg University Working Papers in Economics No. 75.

Tablas

Tabla 3.1: Sesgo, Varianza en relación al Error Cuadrático Medio (ECM).

Efecto Between					Efecto Within				
		$\beta_2 = 0$	$\beta_2 = 1$	$\beta_2 = 2$			$\beta_2 = 0$	$\beta_2 = 1$	$\beta_2 = 2$
Sesgo	$\gamma_1 = 0$	0.00%	0.01%	0.00%	Sesgo	$\gamma_1 = 0$	26.6%	29.2%	28.0%
	$\gamma_1 = 1$	0.01%	0.01%	0.01%		$\gamma_1 = 1$	13.2%	15.0%	13.5%
	$\gamma_1 = 2$	0.28%	0.02%	0.00%		$\gamma_1 = 2$	34.4%	35.0%	34.0%
Varianza	$\gamma_1 = 0$	100.0%	100.0%	100.0%	Varianza	$\gamma_1 = 0$	73.4%	70.8%	72.0%
	$\gamma_1 = 1$	100.0%	100.0%	100.0%		$\gamma_1 = 1$	86.8%	85.0%	86.5%
	$\gamma_1 = 2$	99.7%	100.0%	100.0%		$\gamma_1 = 2$	65.6%	65.0%	66.0%
ECM		100%	100%	100%			100%	100%	100%

Nota: basado en 5000 experimentos de Monte Carlo con muestras de 500 observaciones.

Tabla 3.2: Desempeño muestral. Sesgo y Varianza en relación al Error Cuadrático Medio (ECM).

Estimación por cuantiles					Simulación numérica				
		$\beta_2 = 0$	$\beta_2 = 1$	$\beta_2 = 2$			$\beta_2 = 0$	$\beta_2 = 1$	$\beta_2 = 2$
Sesgo	$\gamma_1 = 0$	13.3%	7.3%	1.7%	Sesgo	$\gamma_1 = 0$	0.00%	0.27%	0.61%
	$\gamma_1 = 1$	7.8%	5.4%	2.0%		$\gamma_1 = 1$	83.4%	69.1%	36.1%
	$\gamma_1 = 2$	27.1%	19.6%	10.0%		$\gamma_1 = 2$	90.9%	84.6%	65.7%
Varianza	$\gamma_1 = 0$	86.7%	92.7%	98.3%	Varianza	$\gamma_1 = 0$	100.0%	99.7%	99.4%
	$\gamma_1 = 1$	92.2%	94.6%	98.0%		$\gamma_1 = 1$	16.6%	30.9%	63.9%
	$\gamma_1 = 2$	72.9%	80.4%	90.0%		$\gamma_1 = 2$	9.1%	15.4%	34.3%
ECM		100%	100%	100%	ECM		100%	100%	100%

Nota: basado en 5000 experimentos de Monte Carlo con muestras de 500 observaciones.

Tabla 4.1: Relación entre el salario y el nivel educativo. Argentina 1992 – 2088.
Hombres entre 16 y 65 años de edad

	OLS	QR(0.10)	QR(0.25)	QR(0.50)	QR(0.75)	QR(0.90)
				1992 (n = 12196)		
Años de educación	0,007	0,023	0,006	-0,002	0,001	-0,005
	(1,00)	(1,09)	(0,38)	(0,18)	(0,03)	(0,20)
(Años de educación)^2	0,004	0,002	0,003	0,005	0,005	0,006
	(12.73)**	(2.24)*	(4.38)**	(7.76)**	(6.41)**	(4.96)**
				1998 (n = 11228)		
Años de educación	-0,02	-0,006	-0,029	-0,033	-0,024	0,003
	(2.76)**	(0,22)	(1,78)	(3.23)**	(2.46)*	(0,20)
(Años de educación)^2	0,007	0,005	0,006	0,007	0,007	0,006
	(19.58)**	(4.22)**	(8.49)**	(14.84)**	(15.77)**	(7.78)**
				2008 (n = 14580)		
Años de educación	-0,016	-0,022	-0,022	-0,031	-0,023	-0,018
	(2.17)*	(1,13)	(1,59)	(2.34)*	(1,92)	(1,20)
(Años de educación)^2	0,005	0,005	0,005	0,006	0,005	0,005
	(15.09)**	(5.85)**	(8.29)**	(9.56)**	(10.02)**	(7.72)**

Fuente: estimaciones propias en base a la EPH.
Nota: otros regresores incluidos son experiencia potencial (y su cuadrado), estado marital y controles por región geográfica. Los errores estándar se muestren entre paréntesis, * indica significatividad al 10%, ** al 5% y *** al 1%.

Tabla 4.2: Efecto marginal de la educación sobre la desigualdad.
Argentina 1992 – 2008. Hombres de 16 a 65 años de edad

	1992	1998	2008
1. Desigualdad			
Gini	40.5	44.0	39.8
Varianza de los logaritmos	42.5	53.8	48.5
2. Simulación numérica (*location shift*)	1.82**	3.69**	1.71**
	(0.32)	(0.36)	(0.24)
3. Estimación RIF (Firpo et al. 2009)	4.65**	6.36**	2.03**
	(0.51)	(0.55)	(0.36)
4. Descomposición por cuantiles			
Efecto Between	1.81**	3.69**	1.71**
	(0.32)	(0.36)	(0.24)
Efecto Within	1.88**	1.47**	0.560
	(0.31)	(0.29)	(0.29)
Cambio total	3.69**	5.16**	2.27**
	(0.48)	(0.48)	(0.35)
5. Variaciones relativas (en %)			
Efecto Between (convexidad)	49.2%	71.5%	75.3%
Efecto Within (heterogeneidad)	50.8%	28.5%	24.7%
Cambio total	100%	100%	100%

Fuente: estimaciones propias en base a la EPH.
Nota: errores estándar entre paréntesis, * indica significatividad al 10%, ** al 5%.

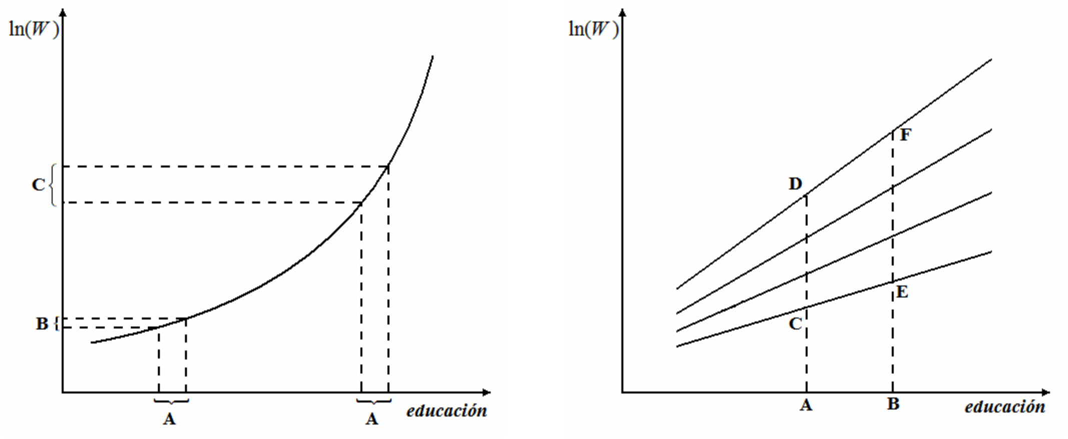
Tabla 4.3: Correcciones por selección muestral en las ecuaciones de Mincer

	1992		1998		2008	
	Educación	(Educación)^2	Educación	(Educación)^2	Educación	(Educación)^2
Media Condicional						
OLS	0.699	0.42***	-1.976	0.65***	0.342	0.39***
Heckman (1979)	-1.484	0.56***	0.159	0.46*	6.569	0.071
Diferencia	-2.183	0.14	2.135	-0.193	6.227	-0.323
Cuantil Condicional 0.10						
Koenker (1978)	2.280	0.21**	-0.564	0.49***	1.025	0.34***
Buchinsky (2001)	1.577	0.097	-93.77	5.364	8.58*	-0.0483
Diferencia	-0.704	-0.115	-93.21	4.871	7.55*	-0.39**
Cuantil Condicional 0.25						
Koenker (1978)	0.624	0.34***	-2.89*	0.64***	0.0352	0.40***
Buchinsky (2001)	1.738	-0.0903	-38.17	2.477	8.24**	-0.00363
Diferencia	1.114	-0.429	-35.28	1.842	8.207	-0.403
Cuantil Condicional 0.50						
Koenker (1978)	-0.223	0.45***	-3.34***	0.72***	-1.031	0.46***
Buchinsky (2001)	5.085	-0.494	-46.01	2.963	5.048	0.167
Diferencia	5.31**	-0.95**	-42.67	2.247	6.08*	-0.29*
Cuantil Condicional 0.75						
Koenker (1978)	0.056	0.50***	-2.40**	0.73***	0.711	0.41***
Buchinsky (2001)	2.752	-0.088	-50.56	3.28	7.10**	0.119
Diferencia	2.696	-0.587	-48.16	2.549	6.39**	-0.29**
Cuantil Condicional 0.90						
Koenker (1978)	-0.49	0.59***	0.347	0.63***	1.677	0.37***
Buchinsky (2001)	7.568	-0.631	-72.11	4.475	6.029	0.171
Diferencia	8.06**	-1.22*	-72.46	3.842	4.35	-0.204

Fuente: estimaciones propias en base a la EPH.
Notas: Otros regresores incluidos son experiencia potencial (y su cuadrado), estado marital y controles por región geográfica. Para mayor eficiencia computacional al implementar el ejercicio de bootstrap (500 replicas), la primera etapa usa un estimador lineal de probabilidad (OLS) en lugar de Klein and Spady (1993). * indica significatividad al 10%, ** al 5% y *** al 1%. Todos los coeficientes están multiplicados por 100.

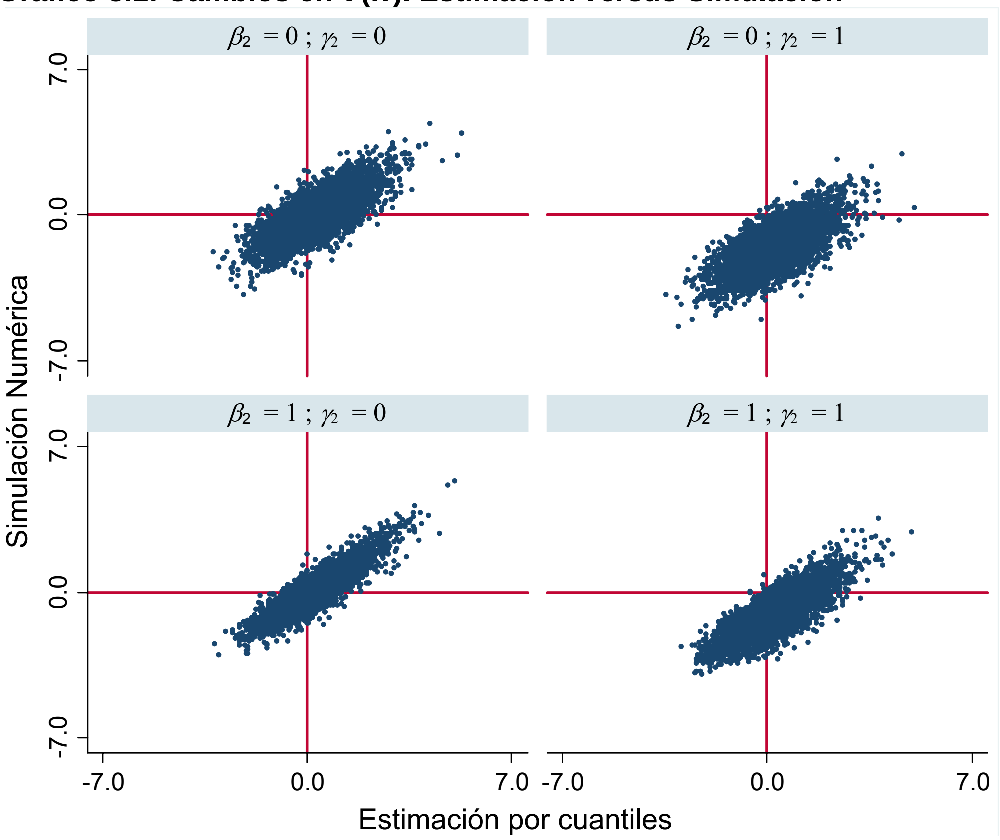
Figuras y Gráficos

Figura 1.1: Ecuaciones de Mincer – Efectos distributivos de la educación.



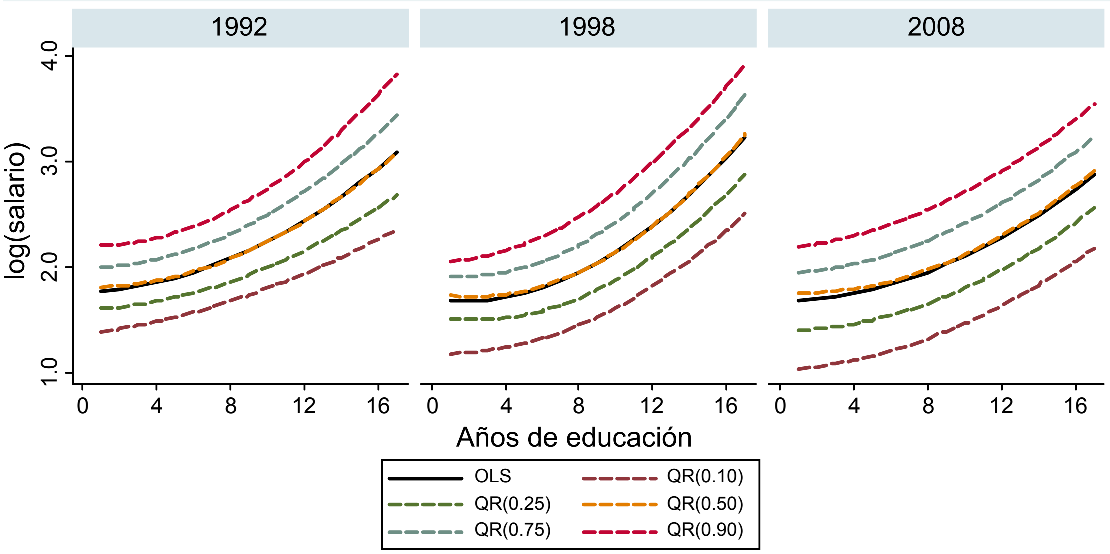
(*a*) *Convexidad* (*b*) *Heterogeneidad*

Gráfico 3.1: Cambios en *V*(*w*). Estimación *versus* Simulación



Nota: 5000 experimentos de Monte Carlo con tamaño de muestra de 500 observaciones.

Gráfico 4.1: Convexidad y Heterogeneidad de los retornos a la educación.
Argentina 1992 – 2008. Hombres entre 18 y 65 años de edad



Fuente: estimaciones propias en base a la EPH.
Nota: el resto de las regresores incluidos en la ecuación de Mincer se evalúan en sus medias muestrales.

Apéndice

A.1 Ecuación (3.4)

Partiendo de la ecuación (3.1), calcular la esperanza *w* condicional en x:

$$w = x'\alpha(\tau)$$

Luego,

$$E(w \mid x) = x' E[\alpha(\tau) \mid x] = x' \underbrace{E[\alpha(\tau)]}_{\beta} = x'\beta \qquad (a.1)$$

donde se ha definido a β como la esperanza del vector de variables aleatorias $\alpha(\tau)$, dado que $\tau \mid x$ es una variable aleatoria $U(0,1)$. Es decir, los parámetros β de la esperanza condicional no son otra cosa que el promedio de los parámetros de los cuantiles condicionales.

Por otro lado, utilizando la ecuación (3.2) calcular la varianza condicional:

$$w = x'\beta + x'\gamma(\tau)$$

Luego,

$$\begin{aligned} V(w \mid x) &= V[x'\beta \mid x] + V[x'\gamma(\tau) \mid x] = V[x'\gamma(\tau) \mid x] \\ &= x'V[\gamma(\tau)]x = x'\Omega x \end{aligned} \qquad (a.2)$$

donde se han aplicado propiedades de la varianza de producto vectoriales y se ha definido a la matriz Ω como $V[\gamma(\tau)]$. Nótese que $E[\gamma(\tau)] = E[\alpha(\tau) - \beta] = 0$ y por lo tanto $V[\gamma(\tau)] = E[\gamma(\tau)\gamma(\tau)'] = \Omega$. Es decir, en cierto sentido la matriz Ω mide la distancia de los parámetros entre la media y los cuantiles de la distribución de w condicional en x.

Utilizando la ley de varianzas iteradas, mostrada en (3.3), junto con (*a*.1) y (*a*.2) se obtiene que:

$$V(w) = V[x'\beta] + E[x'\Omega x]$$

Luego, usando propiedades de la varianza para el producto de vectores y propiedades de la esperanza para formas cuadráticas:

$$V(w) = \beta' V(x)\beta + tr[\Omega V(x)] + E(x)'\Omega E(x)$$

que es la ecuación (3.4) con la notación $I = V(w)$, $V = V(x)$ y $E = E(x)$.

A.2 Ecuación (3.5)

Es usual en esta literatura hacer el supuesto de que la distribución de w condicional en las x no se ve afectada por los cambios en la distribución de las x. La consecuencia del mismo sobre el modelo utilizado en este trabajo es que los parámetros β y Ω no cambian ante una traslación horizontal en cualquiera de los regresores incluidos en x. Este supuesto hace explicito el hecho de que se trata un análisis de equilibrio parcial, en el sentido de que un pequeño cambio en la educación (medido por *h*) no cambia los retornos a la educación. Para obtener la derivada de la desigualdad I con respecto a una traslación horizontal en h se diferencia la ecuación (3.4). Por ejemplo, derivar la expresión $\beta' V\beta$ (primer sumando) implica realizar el siguiente ejercicio:

$$\begin{aligned}
\delta[\beta' V(x)\beta] &= \lim_{\varepsilon\to 0}\frac{\beta' V(x+\varepsilon)\beta - \beta' V(x)\beta}{\varepsilon} \\
&= \lim_{\varepsilon\to 0}\frac{\beta'[V(x+\varepsilon) - V(x)]\beta}{\varepsilon} = \beta' \underbrace{\lim_{\varepsilon\to 0}\frac{[V(x+\varepsilon) - V(x)]}{\varepsilon}}_{\equiv\delta(V)}\beta \\
&= \beta'\delta(V)\beta
\end{aligned}$$

donde se emplea la propiedad de que el límite pasa a través de las transformaciones continuas.

Utilizando el razonamiento anterior aplicado al resto de los sumandos de la ecuación (3.4) se obtiene que:

$$\delta[tr(\Omega V)] = tr[\Omega\delta(V)] \qquad \text{y} \qquad \delta(E'\Omega E) = 2E'\Omega\delta(E)$$

Sumando estos tres componentes se obtiene la expresión (3.5).

A.3 Expresiones (3.6)

Para obtener las expresiones (3.6) es conveniente analizar cada elementos de E y V. La matriz V contiene todas las varianzas y covarianzas de las variables incluidas en x, mientras que E es un vector que contiene la esperanza x. Formalmente,

$$E \equiv \begin{bmatrix} E_0 \\ E_1 \\ E_2 \\ E_z \end{bmatrix} \quad \text{y} \quad V = \begin{bmatrix} V_{00} & V_{01} & V_{02} & M_{0z} \\ V_{10} & V_{11} & V_{12} & M_{1z} \\ V_{20} & V_{21} & V_{22} & M_{2z} \\ M_{z0} & M_{z1} & M_{z2} & M_{zz} \end{bmatrix}$$

donde la notación de los elementos incluye los siguientes escalares,

$$E_k = E(h^k) \qquad , k = 0,\ 1,\ 2$$
$$V_{jk} = COV(h^j, h^k) = E_{j+k} - E_j E_k = V_{kj} \qquad , k = 0,\ 1,\ 2 \quad j = 0,\ 1,\ 2$$

junto con los vectores de dimensión (Qx1),

$$E_z = E(z)$$
$$M_{kz} = COV(h^k, z) = M'_{kz} \qquad , k = 0,\ 1,\ 2$$

y la matriz de dimensión (QxQ).

$$M_{zz} = V(z)$$

Nótese que cuando $k = 0$ el vector M_{0z} es un vector nulo, dado que se trata de las covarianzas entre $h^0 = 1$ y cada uno de los regresores.

Las expresiones $\delta(E)$ y $\delta(V)$ se construyen con las derivadas funcionales de cada uno los elementos de E y V, respectivamente. Es decir:

$$\delta(E) \equiv \begin{bmatrix} \delta(E_0) \\ \delta(E_1) \\ \delta(E_2) \\ \delta(E_z) \end{bmatrix} \quad \text{y} \quad \delta(V) \equiv \begin{bmatrix} \delta(V_{00}) & \delta(V_{00}) & \delta(V_{00}) & \delta(M_{0z}) \\ \delta(V_{10}) & \delta(V_{11}) & \delta(V_{12}) & \delta(M_{1z}) \\ \delta(V_{20}) & \delta(V_{21}) & \delta(V_{22}) & \delta(M_{2z}) \\ \delta(M_{z0}) & \delta(M_{z1}) & \delta(M_{z2}) & \delta(M'_{zz}) \end{bmatrix}$$

Supongamos una traslación horizontal en la distribución de los años de educación h, es decir sumemos un ε a cada valor posible de h y considérese las siguientes derivadas funcionales sobre los distintos elementos de E y V.

(*i*) *Momentos de h de orden k*.

$$\delta(E_k) \equiv \lim_{\varepsilon \to 0} \frac{E[(h+\varepsilon)^k] - E[h^k]}{\varepsilon} = E\left[\lim_{\varepsilon \to 0} \frac{(h+\varepsilon)^k - h^k}{\varepsilon}\right] = E\left[kh^{k-1}\right] = kE_{k-1} \quad \text{(a.3)}$$

Por otro lado, $E_z = E(z)$ no cambia ante una traslación horizontal en ε, es decir $\delta(E_z) = 0$. Sustituyendo en $\delta(E)$ se obtiene la primera parte de (3.6).

(*ii*) *Varianzas y covarianzas entre h^j y h^k*

El origen de este segmento de la matriz V es la inclusión de h y h^2 como parte de los regresores x. Por definición, $V_{jk} = E_{k+j} - E_k E_j$. Luego, derivando ésta expresión y utilizando el resultado (*a*.3):

$$\begin{aligned}\delta(V_{jk}) &= \delta(E_{k+j} - E_k E_j) = \delta(E_{k+j}) - \delta(E_k)E_j - E_k\delta(E_j) \\ &= (k+j)E_{k+j-1} - kE_{k-1}E_j - jE_k E_{j-1} \\ &= kV_{j(k-1)} + jV_{(j-1)k} \end{aligned} \tag{a.4}$$

para $j = 0, 1, 2$ y $k = 0, 1, 2$.

(*iii*) *Covarianzas entre h^k y los regresores z*

Nuevamente, la inclusión en x de h y h^2 junto con otros regresores z da origen a este segmento de la matriz V. Notar que para el caso de $k = 0$, el vector M_{z0} no cambia ante una traslación horizontal en h, por lo tanto $\delta(M_{0z})$ es un vector de ceros de dimensión Q. Es decir,

$$\delta(M_{0z}) = 0_{1\times Q} \tag{a.5}$$

Para los casos con $k > 1$, se debe mirar cada elemento por separado. Sea z_q un regresor perteneciente al vector z (de dimensión Q). El elemento q del vector M_{kz} es $COV(h^k, z_q) = E(h^k z_q) - E(h^k)E(z_q)$, por lo tanto:

$$\begin{aligned}\delta[COV(h^k, z_q)] &= \delta[E(h^k z_q) - E(h^k)E(z_q)] = \delta[E(h^k z_q)] - \delta[E(h^k)]E(z_q) \\ &= kE(h^{k-1}z_q) - kE(h^{k-1})E(z_q) = kCOV(h^{k-1}, z_q)\end{aligned}$$

para $k = 1, 2$ y $q = 1,\ldots, Q$, donde se ha hecho uso de (*a*.3) junto con el hecho de que $E(z_q)$ no cambia ante una traslación de h. Además,

$$\delta[E(h^k z_q)] = E\left[\lim_{\varepsilon\to 0}\frac{(h+\varepsilon)^k - h^k}{\varepsilon}\cdot z_q\right] = E\left[kh^{k-1}z_q\right] = kE\left[h^{k-1}z_q\right]$$

De aquí surge que el vector M_{kz} cambia de la siguiente manera:

$$\delta(M_{kz}) = kM_{(k-1)z}\text{, con } k = 1,2. \tag{a.6}$$

Notar que para el caso de $k = 1$, el cambio en M_{1z} es un vector nulo de dimensión Q, ya que M_{0z} es un vector de ceros.

(*iv*) *Varianzas y covarianzas de z*

Dado que estos momentos no dependen de la distribución de h, la matriz M_{zz} no cambia ante una traslación en h. En consecuencia $\delta(M_{zz})$ es una matriz nula de dimensión QxQ.

$$\delta(M_{zz}) = 0_{Q\times Q} \qquad (a.7)$$

Sustituyendo los resultados (a.4) hasta (a.7) en $\delta(V)$ se obtiene la segunda parte de la expresión (3.6).

A.4 Experimento de Monte Carlo

El diseño del experimento de Monte Carlo consta de considerar el siguiente proceso generador de datos:

$$w_i = \beta_0 + \beta_1 h_i + \beta_2 h_i^2 + \beta_z z_i + (\gamma_0 + \gamma_1 h_i + \gamma_2 h_i^2 + \gamma_z z_i)u_i$$

donde los regresores $(h_i, z_i) \sim N_2(\mu_h = 0, \mu_z = 0, \sigma_h^2 = 1, \sigma_z^2 = 4, \rho_{hz} = 0.5)$ y u_i es una variable aleatoria $N(0,1)$. Notar que en este modelo, la heterocedasticidad se incorpora por el hecho de que u_i multiplica una función de h y z, que representa a las variables inobservables en una encuesta. Por lo tanto la dispersión de los mismos puede ser modificada por los regresores, dependiendo de los γ parámetros elegidos.

El experimento consiste en generar 5000 muestras de tamaño $n = 500$, con los siguientes valores $\beta_0 = \beta_1 = \gamma_0 = \gamma_z = 1$. Por simplicidad, el experimento asume que no hay convexidad en los cuantiles, es decir $\gamma_2 = 0$. Para el resto de los parámetros (β_2 y γ_1) se usan valores alternativos (ver en Tablas y Gráficos) que ilustran el efecto de la convexidad sobre la media condicional (β_2) y la heterocedasticidad en los errores (γ_1).