

Curvas de Salario

Análisis de Corte Transversal - CASEN (1990 - 2006)

Informe Final

Julio 2009

Felipe Lozano

felipe.lozano@lumnifinance.com

Resumen

Se presentan los modelos Minceriano Modificado y Splines Lineales y se corren dos metodologías, Errores Estándar Robustos y Corrección del Sesgo de Selección propuesto por Heckman, para estimar los salarios de la población con educación Superior en Chile con datos de CASEN 2006. Se encuentra que el grado de educación es uno de los principales determinantes de las brechas salariales en Chile, de acuerdo con otros artículos elaborados en Chile al respecto, y que las brechas de género y entre regiones comparadas con Santiago, persisten en Chile. La estimación se incluye dentro del material para la valoración de Contratos de Capital Humano de Lumni.

Posteriormente se hace uso de cada una de las versiones de CASEN para estimar el mismo modelo y ver las diferencias en el periodo 1990-2006. Se encuentra que la dinámica del retorno a la educación se puede separar en dos partes, antes de la crisis asiática donde el retorno de la educación superior aumentó para cada nueva versión, y después de la crisis asiática (98-00), cuando el retorno a la educación se estancó y perdió su ritmo de crecimiento

Tabla de contenido

Introducción y Revisión de Literatura	3
Descripción de la base de datos	5
Ingreso y Nivel de Estudios.....	6
Ingreso y oferta laboral en horas	7
Explicando la Participación	8
Estimaciones.....	10
Modelo Minceriano Modificado.....	10
Modelo de Splines Lineales	13
Corrección de Sesgo de Selección	17
Corrección de Sesgo de Selección – Modelo Minceriano Modificado	18
Corrección de Sesgo de Selección – Modelo Splines Lineales.....	20
Estimaciones Mincerianas (1990-2006)	24
Metodologías.....	24
Resultados	24
Modelo Minceriano - Pooled.....	24
Modelo Minceriano – Estimaciones Anuales	26
Modelo de Splines Lineales – Pooled	27
Modelo de Splines Lineales – Estimaciones Anuales	28
Conclusiones.....	30
Referencias	32
ANEXOS.....	33

Introducción y Revisión de Literatura

Una de las variables que más contribuye en la valuación de Contratos de Capital Humano es la proveniente de la estimación de los flujos futuros de los ingresos de los potenciales estudiantes a financiar. Durante los últimos 50 años las estimaciones mincerianas y sus modificaciones han sido la piedra angular de la literatura sobre retornos de la educación y predicciones sobre su rentabilidad futura. Las ecuaciones mincerianas, como fueron propuestas en el artículo original, Mincer (1974), han sido las más exitosas en establecer una relación entre ingreso y educación alcanzada. El logaritmo del ingreso es explicado por la cantidad de años de educación y por la experiencia, y la experiencia en forma cuadrática para poder establecer la relación convexa entre estas variables. La ecuación minceriana está expresada en la ecuación 1, donde w es el salario del individuo i ; s , es la variable de años de educación, y x , mide los años de experiencia:

$$\log[w_i(s_i, x_i)] = \alpha_0 + \rho_s s_i + \beta_0 x_i + \beta_1 x_i^2 + \varepsilon_i \quad (1)$$

La ecuación de Mincer, a través de los años, ha recibido críticas y modificaciones. Una de ellas es la que se refiere al problema de la linealidad respecto al ingreso de todos los niveles de educación (Splines Models). Los modelos de *Splines* Lineales (SL) tratan de corregir el supuesto de que un año de educación básica y un año de educación universitaria retornan una misma tasa, diferenciando entre los niveles educativos.

Otra crítica ha venido del uso de la experiencia, ya que es una variable complicada de medir. En la literatura se ha usado con frecuencia la experiencia potencial que hace a la experiencia una función de la edad y de los años de educación: $Max[age - s - 6, age - 16, 0]$ (Heckman, et Al. 2005). Al incluir esta nueva variable de la experiencia potencial, en función de los años de educación, el efecto de los años de educación sobre la ecuación minceriana se distorsiona, y el retorno a la educación no puede ser medido exclusivamente por el coeficiente ρ_s . Incluyendo la variable de experiencia potencial, el efecto de un incremento en nivel de educación resultaría en:

$$\frac{\partial \log w}{\partial s} = \rho_s + \beta_0 \frac{\partial x}{\partial s} + 2\beta_1 \frac{\partial x^2}{\partial s} \quad (2)$$

Esta especificación implicaría que las curvas del logaritmo del ingreso con respecto a la experiencia fueran paralelas a lo largo de todos los años de educación, y que el logaritmo del ingreso con respecto a la edad diverge con la edad a lo largo de los años de educación. Para evitar los problemas derivados del uso de la experiencia potencial se usa la edad como proxy de experiencia sin la transformación sugerida anteriormente.

Para romper el supuesto de elasticidad unitaria entre el ingreso y la oferta laboral, en la literatura se ha incluido el logaritmo de las horas trabajadas al mes, a las que corresponden el salario declarado (Guataqui, et Al. 2008). Otro tipo de modificaciones han sido la inclusión de variables socio - demográficas como el género o

el lugar de origen de una persona, o la inclusión del la Corrección por Riesgo de Selección, propuesta por Heckman, con características del los hogares que pueden explicar el hecho de que un individuo sea parte o no de la fuerza laboral.

Otros experimentos incluyen la diferenciación en las estimaciones por quintiles para analizar el efecto de la educación sobre los diferentes grupos sociales (Prada, 2006) o centrarse en los estudios específicos de grupos poblacionales según el nivel de estudios alcanzado (Lemieux, 2006; Guataqui et Al., 2009). En el caso de los Contratos de Capital Humano y de la educación vocacional una aproximación al grupo de profesionales y técnicos graduados de la educación superior sería lo ideal. Desafortunadamente, la fuente existente de seguimiento a los ingresos de graduados de la ES, Futuro Laboral, no está abierta al público.

Para Chile, el estudio de Sapelli, 2009, es el más actualizado que incluye estimaciones mincerianas y SL para los diferentes niveles educativos, todos estadísticamente significativos, salvo para la educación media técnica y para la educación técnica profesional, donde “los resultados no son claros”. Su estudio usa los datos de CASEN y de la EOUC para varios años disponibles. Al estimar los resultados por cohortes puede hacer la comparación los resultados de una misma generación y lo que hubiera obtenido con la estimación minceriana. Ese seguimiento de cohortes sintéticas demuestra que las estimaciones mincerianas han sido conservadoras y que están sesgadas hacia abajo en Chile durante el periodo de estudio. Su estimación sin incluir más variables hace que sea comparable con estudios internacionales, llegando a la conclusión de que en Chile las tasas de retorno para los niveles secundaria y superior son mayores que en otros países similares y que de varios países en desarrollo.

En estudios tratando de enfocarse en los profesionales, Meller y Rappoport (2004) hacen estimaciones de los déficits de profesionales y técnicos que necesita la economía chilena acorde a su nivel de ingreso per cápita. A través de un modelo de causalidad doble, en donde el ingreso causa el nivel de capital humano, y viceversa. Se usan diferentes tipos de medición para el capital humano. El estudio encuentra en Chile todavía existe un déficit de profesionales y que cuentan con una dotación relativamente baja de los mismos. Por otro lado, el estudio encuentra que existe un problema de carácter cualitativo en la ES y que los profesionales y técnicos no están del todo preparados para enfrentarse a la “sociedad del conocimiento”.

Igualmente, Mizala y Romaguera (2003) tratan de hacer seguimiento a la base de datos de CASEN para establecer el retorno que hay a la educación superior, por encima de las demás niveles educativos. Hacen una comparación entre los datos que se encuentran en CASEN 1990 y 2000, tratando de analizar las implicaciones que ha tenido en la economía chilena el cambio en la estructura educativa de la población y el incremento en participación de las mujeres, ya que en el lapso de esos 10 años la escolaridad de la sociedad como un todo cambió: en 1990, el 42% de la población económicamente activa había completado primaria contra un 6% que había finalizado secundaria; y en 2000, el 25% de la fuerza de trabajo había terminado primaria y un 30% había terminado secundaria. Por otro lado, las mujeres habían aumentado su participación en el mercado laboral especialmente en los sectores con menor grado de educación. En el estudio se incluyen las variables de género,

que muestra una diferencia significativa en el ingreso a favor de los hombres, y de urbanidad, que muestra una diferencia significativa respecto al sector rural.

El estudio de Mizala y Romaguera (2003) trata de explicar la persistencia de los diferenciales de ingreso entre los salarios de individuos de diferentes niveles educativos a pesar de los respectivos aumentos en la escolaridad de la fuerza laboral. Para resolver esa inquietud presenta dos hipótesis: la primera, como resultado del patrón que determina la intensidad transada en el comercio internacional, donde la entrada a la escena mundial de países con abundante mano de obra no calificada (China, India, Indonesia) ha deprimido los salarios de ese tipo de trabajadores en los demás países. Por el otro lado, la segunda hipótesis plantea que la tecnología es complementaria de la mano de obra calificada y que ahorra recursos de mano de obra descalificada. Esta segunda hipótesis es apoyada por el hecho de que los sectores que en el mundo han experimentado mayores crecimientos en los últimos años también son los sectores que más empleos han creado para empleados preparados.

En el presente estudio se recogen las experiencias internacionales que no han sido aplicadas en Chile y se hace especial énfasis en ellas: La corrección del sesgo de selección y los retornos a la educación por quintiles, al igual que la inclusión de otras variables dentro de la ecuación que relaciona el ingreso con la educación.

Descripción de la base de datos

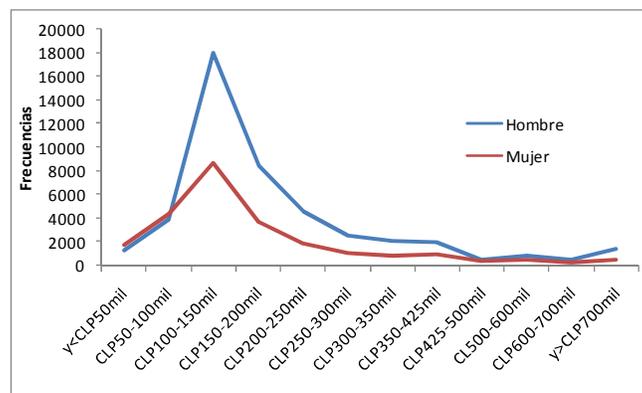
La base de Caracterización Socioeconómica CASEN tiene ediciones públicas desde 1990 hasta 2006. La edición 2006 cuenta con 268,873 observaciones por individuos, de entrevistas realizadas a 73,720 hogares. De esas observaciones 192,195 están en edad de trabajar, con edades superiores a los 12 años y 71,809 declaran ingreso salarial de forma individual. Descontando las observaciones que no responden al nivel de estudios, o cuyo nivel de estudios fue educación especial quedan 71,444 observaciones que fueron tomadas en cuenta para el presente análisis descriptivo.

Se tuvieron en cuenta 12 intervalos de ingreso mensual de la siguiente forma:

- | | |
|---|---|
| 1. Menor a CLP50,000 (Menor a USD94); | 2. Entre CLP50,000 y CLP100,000 (USD94-189); |
| 3. Entre CLP100,000 y CLP150,000 (USD189-283); | 4. Entre CLP150,000 y CLP200,000 (USD283-377); |
| 5. Entre CLP200,000 y CLP250,000 (USD377-471); | 6. Entre CLP250,000 y CLP300,000 (USD471-566); |
| 7. Entre CLP300,000 y CLP350,000 (USD566-660); | 8. Entre CLP350,000 y CLP425,000 (USD660-801); |
| 9. Entre CLP425,000 y CLP500,000 (USD801-943); | 10. Entre CLP500,000 y CLP600,000 (USD943-1,131); |
| 11. Entre CLP600,000 y CLP700,000 (USD1,131-1,320); | 12. Mayor a CLP700,000 (Más de USD1320); |

El tercer y cuarto intervalos con salarios entre CLP100,000-200,000 encierran más del 50% de los encuestados. Por otro lado, los intervalos con salarios superiores a los CLP500,000 concentran apenas al 6% de los individuos que declararon ingresos laborales. La Figura 1 presenta la frecuencia de los registros de ingreso dependiendo del género. Los hombres son mayoría en casi todos los grupos de ingreso, salvo en los dos intervalos más bajos. Igualmente, la participación relativa de las mujeres en el mercado laboral crece a medida que aumenta el intervalo de ingreso que se está observando. Más adelante se mostrará cómo la diferencia entre géneros tiende a desaparecer por el efecto de la Educación Superior en los rangos de ingresos elevados.

Figura 1 Ingreso y Género. Frecuencias



Ingreso y Nivel de Estudios

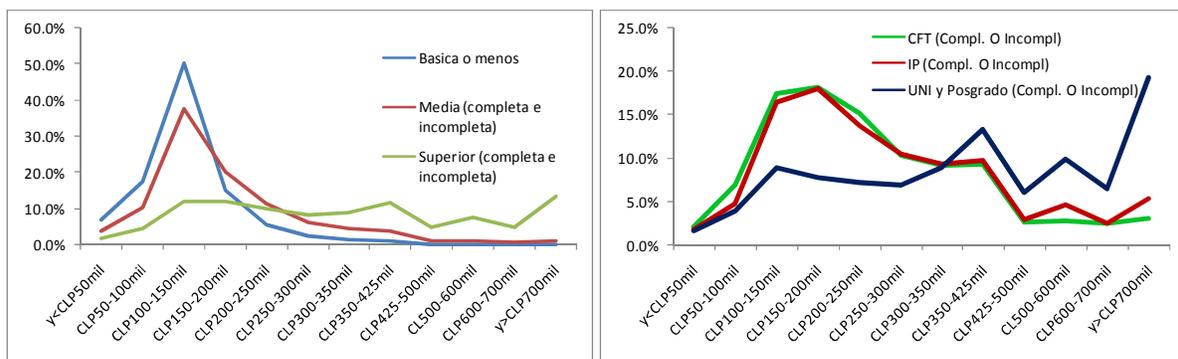
En cuanto a los niveles de estudio, observaciones cuyo máximo grado de estudio es educación básica o menos¹ concentran la mayor parte de las observaciones en los primeros intervalos de ingreso definidos: 74.3% de las observaciones se encuentran en los intervalos con salarios menores a CLP150,000 mensuales. Por su parte, estudiantes con educación media, completa e incompleta, también están concentrados en los intervalos de ingreso más bajos pero a menor escala: los intervalos de ingreso hasta CLP150,000 sólo concentran al 51.3% de las observaciones. Por último, los individuos con algún tipo de educación superior tienen una distribución casi uniforme desde el intervalo de salarios mayores a CLP100,000; los intervalos de ingreso inferiores a CLP150,000 sólo concentran el 18% del total y los intervalos con salarios mayores a CLP300,000 concentran el 51.3% de los casos. La Figura 2 resume la distribución del ingreso por los niveles de estudio definidos.

La distribución de la población en intervalos de salario también se presenta exclusivamente para los diferentes niveles de educación superior en la Figura 2. El panel izquierdo muestra una distribución cuasi-uniforme a partir del rango entre CLP100,000-150,000 mensuales, para el ingreso de la población con ES. Sin

¹ Para el grupo de Educación Básica se incluyeron las observaciones cuyo máximo nivel de estudios fuera Preparatoria en el sistema antiguo. Igualmente, para el grupo de Educación Media, se incluyeron las observaciones de Humanidades y de vocacionales Técnicas, Comerciales, Industriales y Normales del sistema antiguo.

embargo, cada nivel de educación terciaria presenta diferencias en su distribución. Los CFT e IP concentran 26% y 23%, respectivamente, en rangos con salarios mensuales por debajo de CLP150,000; en cambio, la población con educación universitaria y más apenas concentra el 14% en los primeros tres rangos. Al nivel universitario la participación de los rangos con mayores ingresos se hace cada vez más importante, de hecho, el rango que mayor cantidad de individuos agrupa es el último con sueldos sobre CLP700mil mensuales con un 19% del total de individuos con este nivel de educación. A pesar de esta diferencia con los universitarios, los estudiantes de programas no universitarios tienen una concentración en los primeros rangos de ingreso tan sólo de la mitad del nivel de exhibido por los estudiantes con algún tipo de educación Media.

Figuras 2 y 3 Distribución del Ingreso y Nivel Educativo (Todos los Niveles) / Educación Superior



Ingreso y oferta laboral en horas

Otra variable de interés es el número de horas trabajadas y medir su efecto sobre el ingreso dependiendo del nivel educativo. Se esperaría que a medida que el nivel educativo aumenta, el número de horas trabajadas por un individuo disminuyan para alcanzar un mismo nivel de ingresos. Esto es difícil de observar como tal en la práctica porque la mayoría de individuos están sujetos a algún tipo de contrato que hace rígida la oferta laboral. La Tabla 1 relaciona estas variables, individuos con salarios inferiores a CLP300,000 y con menos de 150 horas trabajadas mensualmente concentran el 62.6% de las observaciones de los individuos sin educación superior, mientras que en el grupo de los graduados universitarios representa apenas el 17.6% y en el grupo de estudiantes con postgrados representa el 12%. Por otro lado, manteniéndose en el grupo con menos de 150 horas mensuales trabajadas, salarios superiores a ese nivel cuentan con sólo el 4.7% de las observaciones para los individuos sin ES, mientras que alcanzan a concentrar el 62% y el 67% para los graduados universitarios y para los individuos con postgrados, respectivamente. Esto corresponde a la hipótesis inicial: un individuo con mayor educación recibe un mayor salario con el mismo tiempo trabajado.

Si el nivel de ingreso es considerado arbitrario, se presentan los datos agrupados en torno a un corte de CLP100,000 y se observan las mismas diferencias entre los grupos por nivel de educación. Una cosa más se

puede mencionar de la observación de la Tabla 1, es que muchas de las personas sin ES están concentradas en grupos cuya oferta laboral está por debajo de las 150 horas (67.3% del total de individuos sin ES con salarios), lo que sugiere que muchos de ellos no tienen contratos estables con al menos 160 horas mensuales y que tienen algún otro tipo de contratación.

Tabla 1 Frecuencia Relativa por nivel de ingreso y horas trabajadas

	Sin ES	UNI Completa	Postgrado
Y<300mil, h<150	62.6%	17.6%	11.9%
Y>300mil, h<150	4.7%	62.3%	67.2%
Y<100mil, h<150	14.0%	1.3%	1.2%
Y>100mil, h<150	53.3%	78.7%	77.9%

Explicando la Participación

En las encuestas de seguimiento al ingreso de la fuerza laboral, sólo se cuenta con información para aquellos individuos que están trabajando en el momento en el que la encuesta es hecha. Existe un problema de selección determinado por la decisión previa de un individuo de participar o no en el mercado laboral. Dentro de las estimaciones presentadas en el este artículo se trata de corregir ese problema de selección, a través de un modelo que explique, en la medida de lo posible, la decisión de participar de los individuos que lo hacen y de aquellos que no. La estimación del efecto de las demás variables sobre el ingreso se hace en una segunda etapa. Ese modelo se presentará con mayor extensión más adelante.

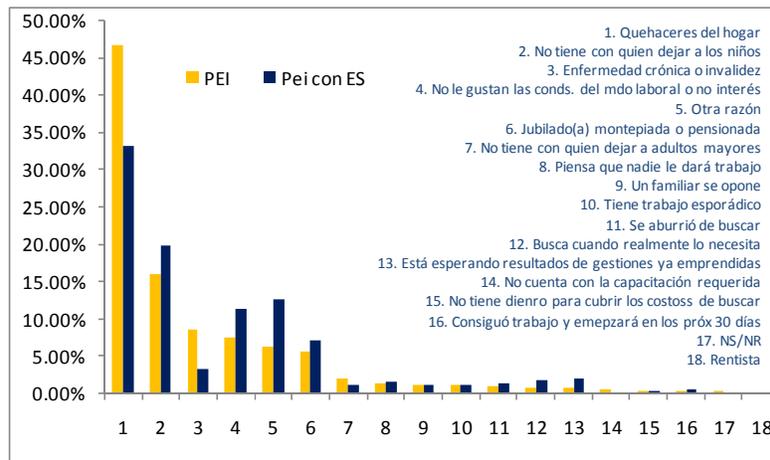
En la presente sección se analiza la pregunta de por qué una persona inactiva, no ha buscado trabajo durante las últimas cuatro semanas. La mayoría de la población económicamente inactiva se encuentra estudiando: para toda la población el porcentaje es del 42% mientras que en el grupo de inactivos con ES es del 68%. Para el siguiente análisis los estudiantes fueron excluidos.

Con una participación del 73% para hombres y de 39% para las mujeres, las razones atribuidas a la no participación deben ser analizadas por género independientemente. Las Figuras 3 y 4 presentan los principales motivos para no buscar trabajo, sin incluir a los que actualmente están estudiando para toda la población y para la población con educación superior. Los datos están ordenados por la participación de cada una de las causas en el caso de la población económicamente inactiva (PEI) total.

Viendo las principales causas de no participación en el caso de las mujeres, los quehaceres del hogar y el cuidado de los niños retienen al 63% de participar en el mercado laboral, aunque esta proporción cae hasta el 52% en el caso de mujeres con ES. Estas causas no se encuentran entre las primeras que retienen a los hombres de entrar al mercado laboral: 3% para toda la población y apenas 1% para aquellos con ES. Motivos de salud e incapacidad son la siguiente razón para las mujeres y la primera causa de inactividad entre los hombres,

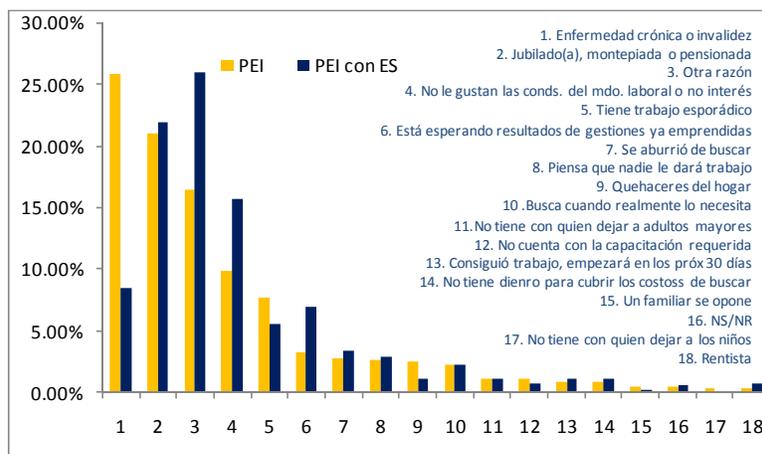
con 9% y 26%, respectivamente, en el caso de toda la PEI. En el caso de la PEI con ES los motivos de salud representaron el 3% y el 8%, para mujeres y hombres. Es de resaltar la falta de peso de esta razón en la población con ES, indicando algún tipo de relación entre el nivel de educación y el riesgo de sufrir enfermedades que generen incapacidad.

Figura 3 ¿Por qué no buscan trabajo? – Mujeres (Rango de Edades)



En el caso de la población con ES los motivos tienen diferentes pesos a los de toda la población. En el caso de los hombres no queda claro cuál es el principal motivo, dado que el 26% atribuyen la inactividad en la búsqueda de empleo a “otra razón”, seguido por los jubilados antes de los 65 años con el 22% de los hombres. Las mujeres con ES también tienen esa diferencia después de tener en cuenta el cuidado de hogar e hijos.

Figura 4 Por qué no buscan trabajo – Hombres



Estimaciones

Para las estimaciones una variable de años de estudio fue creada tomando las variables de nivel de estudios y de curso ofrecidas en CASEN 2006. No se tuvieron en cuenta observaciones que tuvieran un grado de instituciones de Educación Especial. Igualmente, en las regresiones se tomaron en cuenta sólo observaciones que declararan algún tipo de salario laboral, con excepción de los modelos en los que se corrige el sesgo de selección con un modelo de participación. En este último caso se definió la variable participación, que incluyó aquellas observaciones entre 12 y 65 años, que declararon tener algún tipo de ocupación o que estaban activamente buscando empleo. Para los estudiantes con hasta educación básica terminada la variable de años de educación tomó el valor de curso, 0 a 6 en el sistema antiguo y 0 a 8 en el sistema actual. A partir de ahí, para la construcción de la variable años de educación al último curso se le sumaban los años adicionales del nivel inmediatamente anterior, así una persona que terminó sus estudios en 3 años de media básica completó 11 años de estudio. Posteriormente, se hizo lo mismo con los datos de educación terciaria, al curso declarado por el encuestado se le sumaron 12 años de educación básica y media.

Otras transformaciones sobre los datos incluyeron la obtención del logaritmo del ingreso individual y el logaritmo de las horas trabajadas mensualmente. Aquellas observaciones que no declararon a qué número de horas correspondía su salario y declararon que correspondía a un número de días o a un número de meses se les calculó el salario devengado por hora bajo un esquema de 8 horas diarias 20 días al mes. Se construyó una variable dicotómica para el género donde la variable de referencia son los hombres y otra para las regiones donde la referencia es Santiago.

Para el modelo de SL se crearon variables dicotómicas en formas de premios para observaciones que hubieran terminado algún nivel de estudios: Primaria, Secundaria, CFT, IP, Universitaria, acorde con el último año de grado de educación básica y media, y de que se declarara que el nivel de estudios superiores había sido concluido. También se construyó una variable para medir el cambio en el coeficiente que acompaña los años de educación una vez el nivel educativo es terminado, siguiendo a Prada (2006), quien hace un análisis de Splines para Colombia. La transformación es explicada con mayor detalle en la sección de Splines Lineales.

Modelo Minceriano Modificado

La primera ecuación estimada se presenta en la ecuación 2, donde s es la variable de años de educación y en la Tabla 2 se presentan los principales resultados de la estimación con mínimos cuadrados ordinarios MCO, y errores estándar robustos (RSE). Se usaron RSE porque las pruebas de hipótesis rechazaron las hipótesis de normalidad y de homocedasticidad. En el Anexo A.1 en la Tabla A.1 se presenta el resultado de las pruebas de Shapiro-Wilk de normalidad y Cameron & Trivedi de homocedasticidad y en la Figura A.1 se presenta la distribución Kernel empírica de los errores. Se hicieron también, pruebas de multicolinealidad y a

juzgar por el factor de inflación de varianza (vif), sólo se presenta un vif elevado sobre el nivel recomendado (vif=10) en el caso de la edad y su transformación al cuadrado, que por definición, una es la transformación de la otra variable y debe existir un grado elevado de correlación.

$$\ln y = \alpha + \rho_s s + \beta_1 edad + \beta_2 edad^2 + \beta_3 \ln\left(\frac{horas}{mes}\right) + \beta_4 sexo + \beta_5 regiones + \varepsilon \quad (2)$$

Tabla 2 Estimación Minceriana Modificada (MM) - RSE

Source	SS	df	MS			
Model	14,234.01	6.00	2,372.34	Number of obs	70,328	
Residual	20,389.67	70,321.00	0.29	F(6, 70321)	5,016.46	
Total	34,623.68	70,327.00	0.49	Prob > F	0.000	
				R-squared	0.4111	
				Adj R-squared	0.4111	
				Root MSE	0.5385	

In_y	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
s	0.0960	0.0007	144.24	0.0000	0.0947	0.0973
edad	0.0338	0.0011	30.39	0.0000	0.0316	0.0360
edad2	-0.0002	0.0000	-18.01	0.0000	-0.0003	-0.0002
ln_hr	0.4848	0.0075	64.28	0.0000	0.4701	0.4996
sexo	-0.2558	0.0045	-56.49	0.0000	-0.2647	-0.2469
regiones	-0.1574	0.0049	-31.82	0.0000	-0.1671	-0.1477
alfa	7.7899	0.0445	175.22	0.0000	7.7028	7.8770

El modelo es significativo de forma conjunta a juzgar tanto por la prueba F como por el R-cuadrado del orden del 41%. Todas las variables incluidas tienen un impacto significativo sobre el ingreso y tienen los signos esperados. Un año más de estudios se traduce en un aumento en el ingreso de un 0.07% y cada año de experiencia medido con la edad se traduce en un aumento del salario de 0.02%. La tasa de crecimiento del salario no es constante en el tiempo y su decrecimiento es medido con el coeficiente de la edad al cuadrado. Ser mujer o vivir en una región diferente a Santiago, son factores que afectan negativamente los salarios.

La Figura 5 presenta estimaciones determinísticas de la evolución del ingreso esperado según la edad para hombres y mujeres con diferentes años de educación para casos de la región metropolitana. Por otro lado la Figura 6, muestra las estimaciones separadas por género junto a sus respectivos intervalos de confianza², sólo para hombres y mujeres con 16 o 20 años de educación. Más adelante, la Figura 7 muestra la estimación determinística de la evolución del ingreso esperado acorde al nivel de educación de un individuo de 30 años, y la Figura 8 separa los dos géneros correspondientes e incluye sus intervalos de confianza.

² Aunque se rechazó la hipótesis de normalidad, los intervalos de confianza fueron construidos suponiendo $\widehat{Y}_t \sim t(N - k - 1)$. Así los intervalos de confianza están definidos por: $\widehat{\ln Y}_t \pm t_{N-k-1}^{0.05} \times \left(\frac{S_{mse}}{\sqrt{n}}\right)$; donde $\widehat{\ln Y}_t$, es el valor medio estimado; t_{N-k-1}^α , es el valor de la distribución t con $N-k-1$ grados de libertad a un nivel del α en prueba de dos colas; S_{mse} es el error estándar medio de la regresión, y n es el número de observaciones para los cuales se está estimando la media, en el presente estudio en todos los casos $n = 10$.

Figura 5 Ingreso Estimado y Edad – (MM)

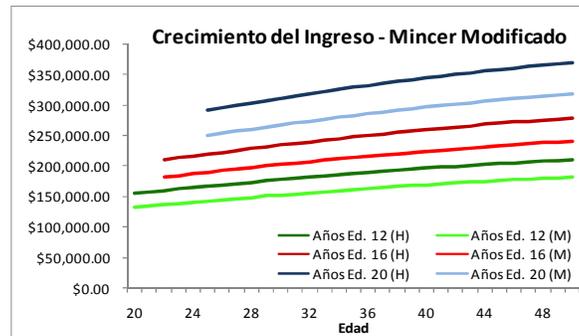
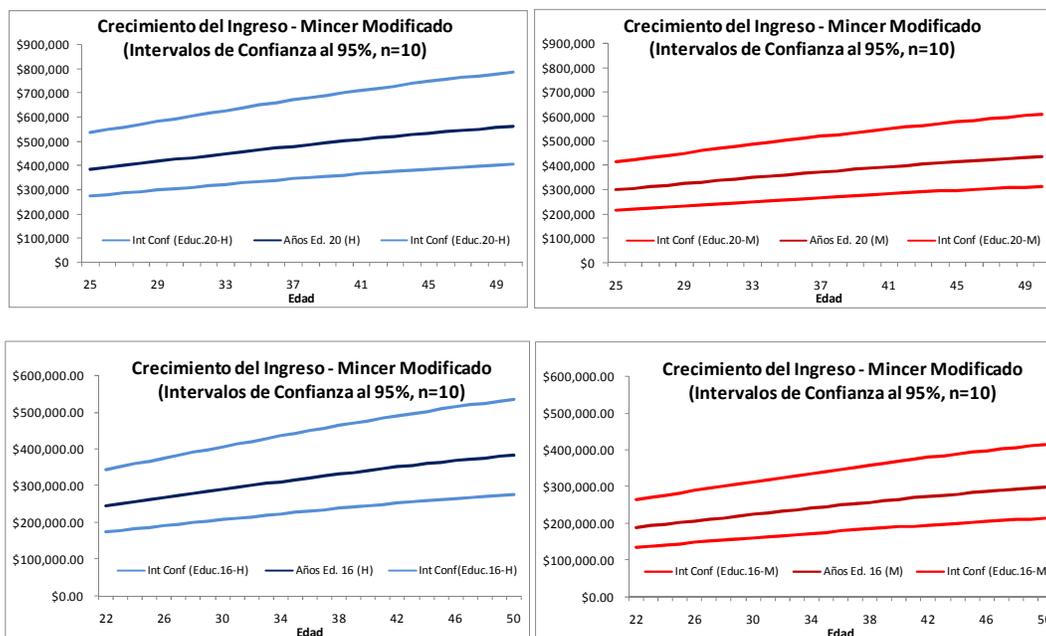


Figura 6 Ingreso Estimado y Edad (Intervalos de Confianza) – (MM)



Las Figuras 5 y 6 con pendientes más pronunciada en los primeros años de inserción al mercado laboral, y con una concavidad medida con el coeficiente del cuadrado de la edad desacelerando el crecimiento del ingreso, se comportan como es esperado. Las estimaciones mincerianas de este estilo concuerdan con Sapelli 2009 en que son conservadoras. Los coeficientes difícilmente logran explicar salarios por encima de CLP350,000 donde se encuentra sólo el 11% de las observaciones. Este esquema promedia hacia abajo los salarios de la población con educación superior.

Las Figuras 7 y 8 también muestran las formas esperadas de una función exponencial positiva. Igualmente, el modelo MM explica salarios sólo hasta CLP500,000, en el caso extremo de que un hombre tenga un nivel de educación superior de 22 años o más.

Figura 7 Ingreso Estimado y Años de Educación – (MM)

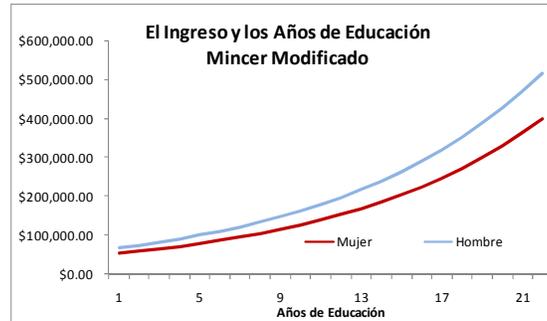
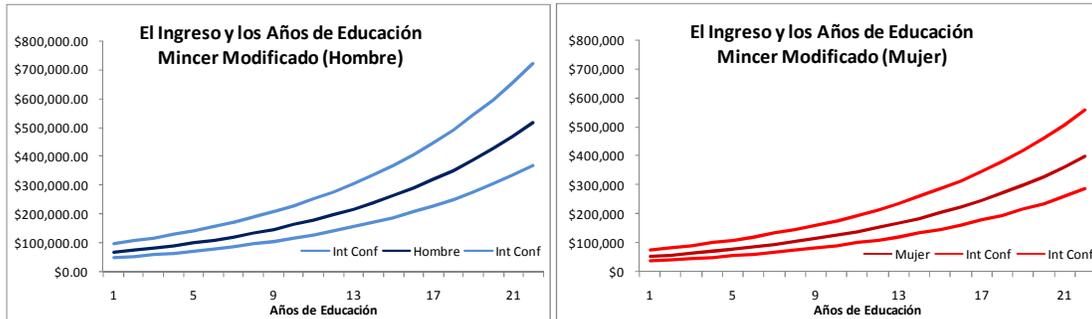


Figura 8 Ingreso Estimado y Años de Educación (Intervalos de Confianza) – (MM)



Modelo de Splines Lineales

Para la estimación de un modelo que incluyera los efectos de cada nivel educativo se incluyeron variables que afectaran el coeficiente de años de educación reconociendo que un año de educación en un nivel tiene un retorno diferente al del nivel precedente, y se definieron premios al ingreso por nivel concluido. Los premios se definieron de la siguiente forma: *primaria*, *secundaria*, *CFT*, *IP* y *UNI*; variables dicotómicas que toman el valor de uno cuando es terminado el nivel de educación respectivo. Por otro lado, se crearon variables para estimar el cambio en el coeficiente de los años de educación al terminar un nivel educativo, siguiendo a Prada (2006) y teniendo en cuenta que el promedio de duración de un programa universitario en Chile es de 5 años. Las variables creadas fueron:

$$\phi_{prim} = primaria \times (s - 8)$$

$$\phi_{secun} = secundaria \times (s - 12) \tag{3}$$

$$\phi_{univ} = UNI \times (s - 17)$$

Es importante observar que una persona que concluye su educación básica no activa inmediatamente el ϕ_{prim} ; este coeficiente sólo se activa en el momento en el que se cursa un año de secundaria o $s > 8$. De esa forma, una persona que terminó la universidad y tiene 20 años de educación va a encontrar una pendiente a sus años de educación definida por: $(\rho + \phi_{prim} + \phi_{sec} + \phi_{univ})$. Como no hay registros organizados de personas que sigan sobre el nivel formal de educación formal después de terminar un CFT o IP y que puedan terminar una licenciatura, no se pueden estimar cambios en el coeficiente de años de educación para esos niveles, y se asume que una persona que obtiene un título no universitario no sigue estudiando o que su coeficiente de años de educación no se ve afectado hasta que obtenga un título universitario.

La Ecuación 4 presenta el modelo a estimar, donde se incluyen las mismas variables del modelo MM más las variables definidas anteriormente. En la Tabla 3 se presentan los principales resultados de las estimaciones con los coeficientes de MCO y RSE. En el Anexo A.2 se presentan las pruebas de hipótesis y la distribución empírica de los errores para el caso de MCO. En la Figura 4 se presentan las estimaciones de curvas de ingreso según el nivel educativo terminado desde educación media para la región metropolitana.

$$\ln y = \alpha + \rho_s s + \beta_1 edad + \beta_2 edad^2 + \beta_3 \ln\left(\frac{horas}{mes}\right) + \beta_4 sexo + \beta_5 regiones + \beta_6 \phi_{prim} + \beta_7 \phi_{secund} + \beta_8 \phi_{univ} + \beta_9 primaria + \beta_{10} secundaria + \beta_{11} CFT + \beta_{12} IP + \beta_{13} UNI + \varepsilon \quad (4)$$

Nuevamente, el modelo en su conjunto es significativo y la representación de los datos se ajusta mejor evaluada por su R cuadrado, alcanzando un nivel de 0.46, no sólo por el número de variables incluidas como lo indica el R cuadrado ajustado por el número de variables. Todas las variables incluidas son significativas y con los signos esperados. Cabe resaltar que el coeficiente de años de educación se reduce considerablemente, comparado con la regresión anterior, esto por la presencia de los premios de cada nivel, que absorben parte del efecto de la educación sobre el ingreso. Los resultados se presentan en la Tabla 3.

Interesados particularmente en los efectos de la educación superior, también cabe mencionar el premio para cada nivel es estadísticamente diferente de los premios para los demás niveles, a juzgar por la falta de intersecciones en los intervalos de confianza entre CFT, IP y Universidades. El hecho de haber terminado la Universidad permite que los salarios crezcan un 60% y que crezcan a una tasa m[ás acelerada. Los otros niveles de educación superior tienen premios de: 20% para IPs y 10% para CFTs. La tasa de cambio de los salarios en esta especificación es mayor comparada con el modelo anterior, pero la velocidad a la que cae la tasa de cambio es mayor; estos dos últimos resultados afectan la concavidad de las curvas de ingreso.

Los dos paneles de la Figura 9 muestran cómo el efecto del género pesa sobre los diferentes niveles de educación a través de las estimaciones determinísticas. En esas gráficas se encuentran la relación de crecimiento del salario junto a la proxy de la experiencia, que en este estudio es la edad, para los diferentes niveles de educación del sistema chileno. Se mantuvo el máximo nivel del eje vertical para comparar los resultados entre géneros. Este modelo logra explicar salarios de hasta CLP900,000 para hombres con 20 años de educación (MA) y de hasta CLP600,000 para mujeres con el mismo nivel educativo. Las gráficas incluidas en

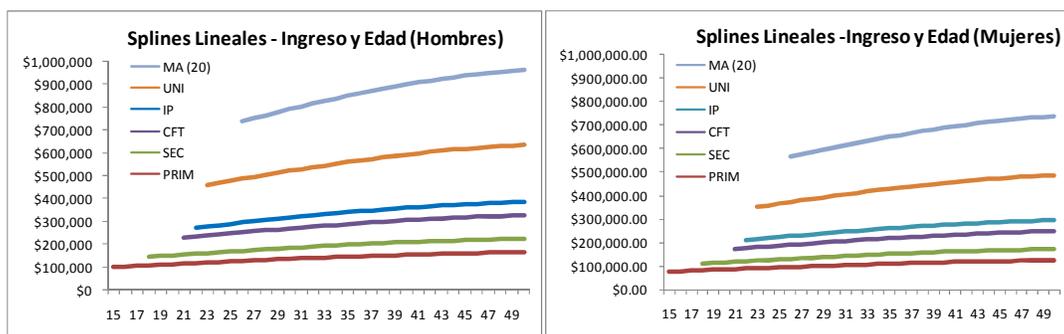
la Figura 10 muestran las curvas, con sus respectivos intervalos de confianza para los niveles de educación superior del sistema chileno. En el caso del nivel post-universitario se tomaron 20 años de estudio correspondientes a 3 años de maestría a partir de la graduación universitaria de 17 años.

Tabla 3 Estimación Splines Lineales (SL) - RSE

Source	SS	df	MS			
Model	16,065.40	14.00	1,147.53	Number of obs	70,328	
Residual	18,558.28	70,313.00	0.26	F(14, 70313)	2,797.47	
Total	34,623.68	70,327.00	0.49	Prob > F	0.000	
				R-squared	0.4640	
				Adj R-squared	0.4640	
				Root MSE	0.5138	

In_y	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
s	0.0297	0.0022	13.71	0.0000	0.0254	0.0339
edad	0.0334	0.0011	30.48	0.0000	0.0312	0.0355
edad2	-0.0003	0.0000	-21.13	0.0000	-0.0003	-0.0003
ln_hr	0.4916	0.0074	66.06	0.0000	0.4770	0.5061
sexo	-0.2680	0.0043	-61.85	0.0000	-0.2764	-0.2595
regiones	-0.1609	0.0047	-34.36	0.0000	-0.1700	-0.1517
Ø prim	0.0260	0.0033	7.7900	0.0000	0.0195	0.0326
Ø secund	0.0327	0.0044	7.4000	0.0000	0.0240	0.0413
Ø univ	0.0552	0.0140	3.9500	0.0000	0.0278	0.0825
primaria	0.0658	0.0093	7.1000	0.0000	0.0476	0.0839
secundaria	0.0914	0.0096	9.4700	0.0000	0.0725	0.1103
CFT	0.1041	0.0179	5.8200	0.0000	0.0690	0.1391
IP	0.1873	0.0162	11.5700	0.0000	0.1556	0.2190
UNI	0.5768	0.0187	30.9100	0.0000	0.5402	0.6133
alfa	8.2836	0.0442	187.3600	0.0000	8.1970	8.3703

Figura 9 Ingreso Estimado y Edad – (SL)



La Figura 11 muestra la relación que existe entre el ingreso y los años de educación para individuos de 30 años. Allí se puede apreciar el efecto de los cursos terminados y de las nuevas pendientes una vez se comienza un nivel educativo diferente. La mayor inclinación y el mayor salto se obtiene después de concluir los estudios universitarios (17 años) y continuar con estudios de postgrado. El salto por la educación primaria es

muy pequeño e imperceptible en la gráfica (8 años). El de la educación media se percibe pero no es pequeño si se compara con el de la educación superior. Los premios al ingreso por otros niveles de educación superior no se incluyen en la gráfica. La Figura 12 presenta las estimaciones por género y sus intervalos de confianza.

Figura 10 Ingreso Estimado y Edad (Intervalos de Confianza) – (SL)

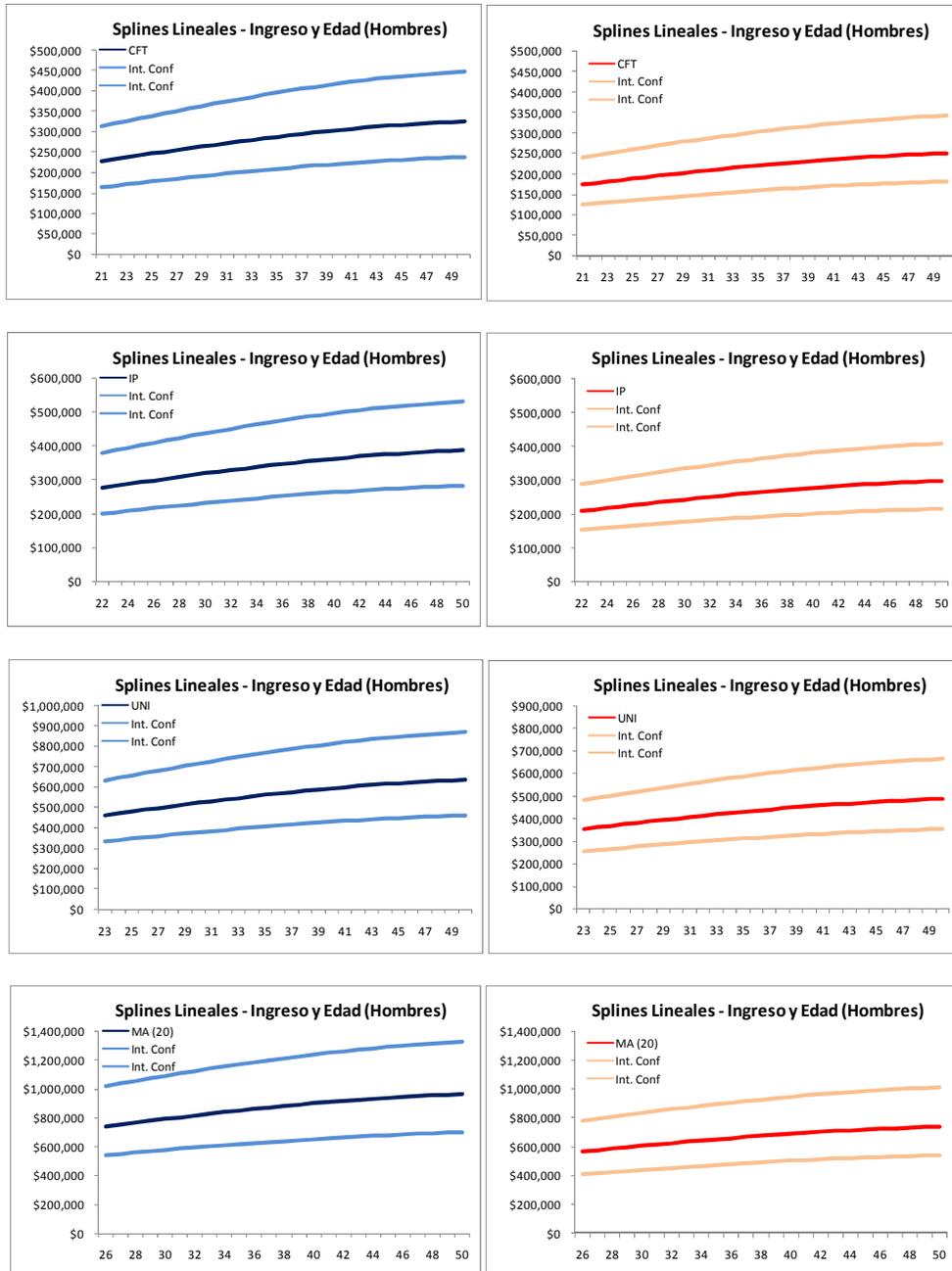


Figura 11 Ingreso Estimado y Años de Educación – (SL)

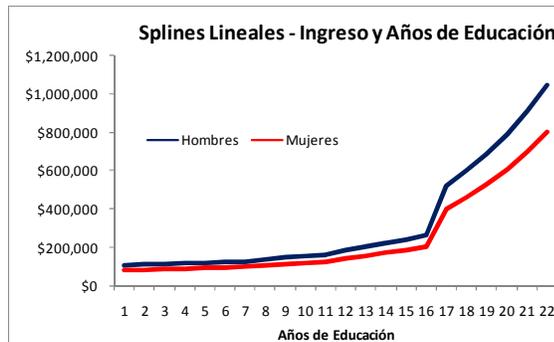
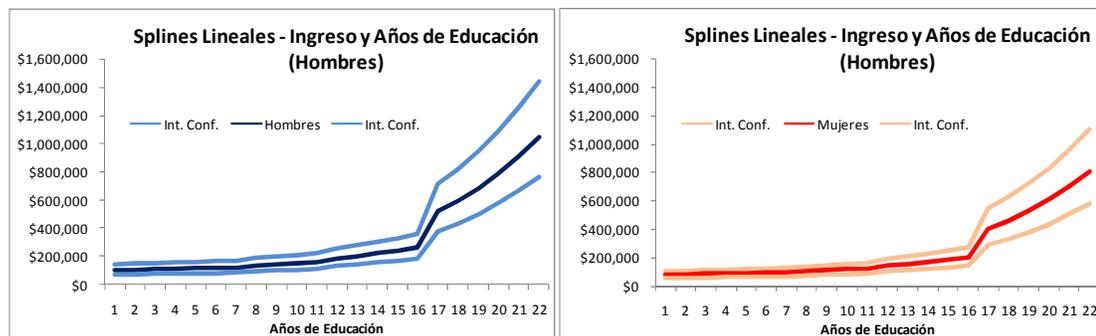


Figura 12 Ingreso Estimado y Años de Educación (Intervalos de Confianza) –(SL)



Corrección de Sesgo de Selección

Cuando se quiere estimar el salario promedio de toda la población, el procedimiento de MCO sería ideal en el caso de que toda la Población en Edad de Trabajar (PET) devengara un salario observable. Sin embargo, al recolectar información, existe un problema de Sesgo de Selección cuando sólo se observan salarios para las personas que están trabajando (Wooldridge, 2001; Heckman, 1979). De hecho, los individuos que se encuentran trabajando, son personas que han recibido una oferta de salario sobre su salario de reserva. Lo que se observa no es simplemente el salario, sino el salario condicionado a que el individuo decidió aceptar un salario. Con esa información, sólo se puede aspirar a estimar $E(w^o|x_1, Participacion = 1)$, a diferencia de $E(w^o|x_1)$ que es lo que se está tratando de estimar, donde x_1 es un vector de variables determinantes del salario y *Participación* es una variable dicotómica que toma el valor de 1 cuando la observación está participando y cero en el caso contrario.

Si una muestra aleatoria incluye elementos que permitan describir los determinantes de la decisión de participar en el mercado laboral, el sesgo puede ser corregido. Heckman (1979), plantea el problema del sesgo de selección como un problema de variables omitidas, porque al añadir la condición de selección en la estimación derivada de los datos –en este caso la participación en el mercado laboral– aparece un nuevo término que no es incluido en el modelo básico de MCO. La Ecuación 5 presenta esta condición:

$$E(w^o | x_1, Participacion = 1) = x_1\beta + \gamma_1\lambda(x\delta) + \varepsilon \quad (5)$$

La estimación en dos etapas sugerida por Heckman (1979), produce estimadores consistentes y asintóticamente normales para grandes muestras. Este procedimiento consiste en estimar los coeficientes del modelo probit $P(\text{Participacion} = 1 | x_i) = \Phi(x_i x \delta)$, para obtener el inverso de la razón de Mills estimado, el ratio entre la función de densidad de probabilidad y la función acumulada de probabilidad $\hat{\lambda}_i \equiv \lambda(x\delta) = \left[\frac{\phi(x\delta)}{\Phi(x\delta)} \right]^{-1}$. El siguiente paso es obtener los coeficientes β y γ_1 de una regresión de MCO. El grado de significancia de γ_1 constituye la prueba de hipótesis sobre la necesidad de corregir el sesgo de selección.

En el modelo de participación se incluyeron las variables socio demográficas determinadas por la relación o parentesco con el jefe del hogar (referencia jefe de hogar), el estado civil (referencia casado), el número de personas en el hogar y el nivel de educación del jefe de hogar. Esta última variable se construyó a partir de la información para los jefes de hogar, cuando no hubo información disponible se utilizó la información de la Esposa del Jefe, otros casos sin información no fueron considerados.

Corrección de Sesgo de Selección – Modelo Minceriano Modificado

Tabla 4 Estimaciones Modelo Heckman Probit – (MM)

Modelo Probit Participacion (MM)						
Participacion	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
Mujer	-0.8326	0.0096	-87.03	0.0000	-0.8513	-0.8138
Regiones	-0.1969	0.0100	-19.76	0.0000	-0.2164	-0.1773
Esposo(a)/Pareja	-1.0647	0.0138	-77.09	0.0000	-1.0918	-1.0376
Hijo(a) de ambos	-1.2167	0.0200	-60.79	0.0000	-1.2559	-1.1775
Hijo(a) jefe	-0.9591	0.0220	-43.56	0.0000	-1.0022	-0.9159
Hijo(a) pareja	-1.0932	0.0338	-32.32	0.0000	-1.1595	-1.0269
Padre o Madre	-0.4174	0.0722	-5.78	0.0000	-0.5590	-0.2759
Suegro(a)	-0.7401	0.0899	-8.23	0.0000	-0.9162	-0.5639
Yerno o nuera	-0.8788	0.0324	-27.08	0.0000	-0.9423	-0.8152
Nieto(a)	-1.8489	0.0404	-45.73	0.0000	-1.9281	-1.7696
Hermano(a)	-0.2921	0.0361	-8.09	0.0000	-0.3629	-0.2214
Cuñado(a)	-0.4299	0.0573	-7.50	0.0000	-0.5422	-0.3176
Otro familiar	-0.7794	0.0354	-22.00	0.0000	-0.8488	-0.7100
No familiar	-0.3043	0.0460	-6.61	0.0000	-0.3945	-0.2141
Serv. Doméstico	2.7873	0.2479	11.24	0.0000	2.3013	3.2732
Union Libre	0.2495	0.0132	18.86	0.0000	0.2236	0.2755
Anulado(a)	-0.1593	0.1121	-1.42	0.1550	-0.3790	0.0604
Separado(a)	0.1618	0.0237	6.83	0.0000	0.1154	0.2081
Divorciado(a)	0.0520	0.1365	0.38	0.7030	-0.2155	0.3196
Viudo(a)	-0.8575	0.0300	-28.58	0.0000	-0.9164	-0.7987
Soltero(a)	-0.3376	0.0178	-18.93	0.0000	-0.3725	-0.3026
NS/NR	-0.2670	0.2885	-0.93	0.3550	-0.8324	0.2985
Numper	0.0235	0.0024	9.62	0.0000	0.0187	0.0283
educ_jefe	-0.0590	0.0013	-44.95	0.0000	-0.0616	-0.0565
s	0.1490	0.0015	101.23	0.0000	0.1461	0.1519
edad	0.0130	0.0004	29.50	0.0000	0.0122	0.0139
Constante	-0.0422	0.0316	-1.34	0.1820	-0.1042	0.0198

La primera etapa de la estimación, el modelo de participación, se presenta en la Tabla 4. En este modelo se encuentra que el género determina de manera significativa la decisión de entrar en el mercado laboral, tal como se esperaba por la descripción anterior de la participación. Igualmente, habitar en un lugar diferente a Santiago disminuye la probabilidad de participar en el mercado laboral. En cuanto al tipo de parentesco o relación existente con el jefe de hogar, todas las variables dicotómicas son significativas y disminuyen la probabilidad de participar, con la excepción de Servicio Doméstico Puertas Adentro, que per se constituye un oficio. Para el estado civil se encuentra que las parejas en Unión Libre están más dispuestas a trabajar, al igual que los separados. Es probable que lo anterior esté relacionado con la independencia del hogar que se adquiere al ir a un hogar con otra persona, o con la independencia que se quiere mantener en el caso de los separados. El contraste se hace con los solteros, pues este estado disminuye la probabilidad de participar. Muchos solteros son estudiantes menores de 30 años que se encuentran estudiando o bajo el techo familiar. La viudez disminuye la probabilidad de participar, pareciendo ser una proxy de retiro pues la mayoría de los viudos son mayores de 55 años. Estar divorciado o tener el matrimonio anulado no ejercen un efecto estadísticamente significativo sobre la participación.

Tabla 5 Estimaciones Modelo Heckman (MM) – (MCO)

Heckman selection model = twostep estimates (regression model w. sample selection)			Number of obs	
i.PCO1	1_1-14	Jefe de hogar omitted	138,700	
i.ECIVIL	L_1-9	Casado omitted	75,379	
			63,321	
			Sigma	0.54136
			Wald chi2(9)	43,734.09
			Prob > chi2	0.0000

In_y	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
s	0.0984	0.0007	141.71	0.0000	0.0971	0.0998
edad	0.0357	0.0012	29.14	0.0000	0.0333	0.0382
edad2	-0.0003	0.0000	-17.92	0.0000	-0.0003	-0.0002
ln_hr	0.4912	0.0050	98.51	0.0000	0.4815	0.5010
sexo	-0.2549	0.0070	-36.21	0.0000	-0.2687	-0.2411
regiones	-0.1538	0.0052	-29.62	0.0000	-0.1640	-0.1436
alfa	7.7024	0.0380	202.45	0.0000	7.6278	7.7769
mills						
lambda	-0.0050	0.0082	-0.6100	0.5440	-0.0210	0.0111

El número de personas en el hogar también aumenta la probabilidad de participar en el mercado laboral, al obligar al hogar a conseguir mayores recursos para su manutención. La educación del jefe del hogar disminuye la probabilidad de buscar empleo o estar empleado: el jefe de hogar con mayor nivel de educación va a buscar que otros miembros de su hogar persigan iguales o más altos niveles de educación y no que salgan a buscar empleo. El nivel educativo de la persona afecta positivamente la participación, reforzando la idea de que la educación es una inversión que se quiere recuperar en el mercado laboral. Finalmente, la edad también aumenta la probabilidad de participación.

En la segunda parte de la estimación, el modelo de ingreso que sigue la anterior especificación del modelo MM, se encuentran los signos esperados en los coeficientes y con magnitudes muy similares al MM de RSE. Se encuentra que el modelo es significativo en su conjunto por el test de Wald. Por otro lado se encuentra que el coeficiente γ que acompaña la razón de inversa de Mills no es estadísticamente significativo. Este resultado constituye una prueba de hipótesis sobre la necesidad de corregir el sesgo de selección, indicaría no ser necesario. Sin embargo, las diferencias entre los géneros presentados en la sección de descripción de la base de datos en cuanto a participación, y en los modelos de esta sección en cuanto a salarios, hacen pensar que se requiere un análisis más detallado. En el Anexo A.3 se presentan los resultados de correr el modelo separando por géneros y se encuentran varias diferencias, entre ellas que la corrección de sesgo es necesaria.

Corrección de Sesgo de Selección – Modelo Splines Lineales

Tabla 6 Heckman Probit – Splines Lineales

Modelo Probit Participacion (SL)						
Particip	Coef.	Std. Err.	z	P> z	[95% Conf	
Mujer	-0.8663	0.0097	-89.11	0.0000	-0.8853	-0.8472
regiones	-0.1982	0.0101	-19.65	0.0000	-0.2180	-0.1785
Esposo(a)/Pareja	-1.0646	0.0139	-76.49	0.0000	-1.0919	-1.0373
Hijo(a) de ambos	-1.2525	0.0202	-61.87	0.0000	-1.2922	-1.2128
Hijo(a) jefe	-0.9901	0.0223	-44.44	0.0000	-1.0338	-0.9465
Hijo(a) pareja	-1.1117	0.0342	-32.53	0.0000	-1.1787	-1.0447
Padre o Madre	-0.4215	0.0720	-5.85	0.0000	-0.5627	-0.2803
Suegro(a)	-0.7690	0.0895	-8.59	0.0000	-0.9445	-0.5935
Yerno o nuera	-0.9295	0.0328	-28.37	0.0000	-0.9937	-0.8653
Nieto(a)	-1.8654	0.0409	-45.65	0.0000	-1.9455	-1.7853
Hermano(a)	-0.3132	0.0362	-8.65	0.0000	-0.3842	-0.2423
Cuñado(a)	-0.4667	0.0577	-8.09	0.0000	-0.5797	-0.3537
Otro familiar	-0.8162	0.0358	-22.80	0.0000	-0.8864	-0.7461
No familiar	-0.3594	0.0465	-7.73	0.0000	-0.4505	-0.2683
Serv. Doméstico	2.7015	0.2430	11.12	0.0000	2.2253	3.1777
Union Libre	0.2496	0.0133	18.75	0.0000	0.2235	0.2757
Anulado(a)	-0.1589	0.1142	-1.39	0.1640	-0.3827	0.0649
Separado(a)	0.1883	0.0238	7.91	0.0000	0.1416	0.2349
Divorciado(a)	0.0643	0.1378	0.47	0.6410	-0.2057	0.3344
Viudo(a)	-0.8570	0.0300	-28.61	0.0000	-0.9157	-0.7983
Soltero(a)	-0.3301	0.0180	-18.37	0.0000	-0.3653	-0.2949
NS/NR	-0.2339	0.2867	-0.82	0.4150	-0.7958	0.3280
Numper	0.0286	0.0025	11.62	0.0000	0.0238	0.0335
educ_jefe	-0.0583	0.0013	-43.59	0.0000	-0.0609	-0.0557
s	0.0442	0.0029	14.99	0.0000	0.0384	0.0500
primaria	0.1960	0.0156	12.54	0.0000	0.1653	0.2266
secundaria	0.5815	0.0136	42.67	0.0000	0.5547	0.6082
terciaria	0.8636	0.0216	40.05	0.0000	0.8214	0.9059
edad	0.0077	0.0005	16.94	0.0000	0.0068	0.0086
constante	0.7115	0.0350	20.34	0.0000	0.6429	0.7800

El modelo SL se corrió igualmente con la metodología de corrección de sesgo de selección sugerida por Heckman. En las Tablas 6 y 7 se presentan los resultados de la estimación de incorporar la Ecuación 4 (SL) en la Ecuación 5 (Corrección de Sesgo). La Tabla 6 presenta el primer paso para el modelo Probit de selección, donde se incluyeron variables de término de niveles educativos para ver cómo afecta el culminar un nivel a la decisión de entrar al mercado laboral.

Tabla 7 Heckman MCO – Splines Lineales

Heckman selection model = twostep estimates
(regression model w. sample selection)

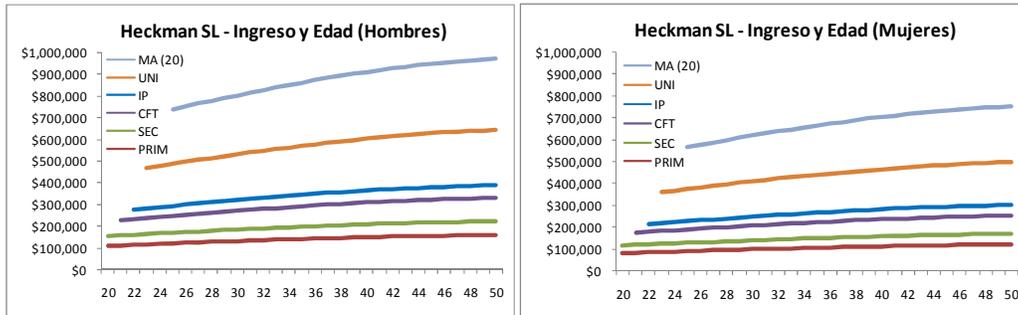
Number of obs	138,700
Censored obs	75,379
Uncensored obs	63,321
Sigma	0.51642
Wald chi2(19)	48,303.03
Prob > chi2	0.0000

ln_y	Coef.	Std. Err.	z	P> z	[95% Conf	
s	0.0301	0.0023	12.87	0.0000	0.0255	0.0347
edad	0.0336	0.0012	28.71	0.0000	0.0313	0.0359
edad2	-0.0003	0.0000	-20.99	0.0000	-0.0003	-0.0003
ln_hr	0.4965	0.0048	104.23	0.0000	0.4872	0.5059
sexo	-0.2465	0.0067	-37.01	0.0000	-0.2595	-0.2334
regiones	-0.1538	0.0050	-31.04	0.0000	-0.1636	-0.1441
prim	0.0277	0.0036	7.58	0.0000	0.0205	0.0348
secund	0.0328	0.0042	7.89	0.0000	0.0246	0.0409
univ	0.0485	0.0109	4.43	0.0000	0.0271	0.0700
primaria	0.0625	0.0104	5.98	0.0000	0.0420	0.0830
secundaria	0.0831	0.0106	7.82	0.0000	0.0622	0.1039
CFT	0.0942	0.0170	5.56	0.0000	0.0610	0.1275
IP	0.1770	0.0148	11.92	0.0000	0.1479	0.2061
UNI	0.5735	0.0161	35.63	0.0000	0.5419	0.6050
alfa	8.2823	0.0361	229.22	0.0000	8.2115	8.3531
mills						
lambda	-0.0412	0.0077	-5.36	0.0000	-0.0562	-0.0261

Los resultados de la Tabla 6 son similares a los encontrados en el modelo de participación del MM, en cuanto a dirección y magnitud, de hecho las características que parecen no influir en el hecho de participar. En esta especificación al incluir las variables de término de nivel educativo se encuentra que el terminar un nivel, la probabilidad de participar en el mercado laboral aumenta, y entre más alto sea el nivel, su finalización hace que la posibilidad de encontrarse en el mercado laboral sea aún más alta. Terminar un grado de educación terciaria hace que la probabilidad de participar de una persona aumente en 0.86 desviaciones estándar. En el caso del coeficiente de los años de educación del individuo el coeficiente es menor que en el anterior modelo porque parte de su poder explicativo ahora está en las variables dicotómicas de finalizar un nivel académico.

En este caso la corrección del sesgo de selección sí es necesaria a juzgar por la significancia del coeficiente de lambda rechazando la hipótesis de que sea igual a cero. Sobre el error estándar de las estimaciones el nivel reportado de sigma (pseudo MSE) para el modelo de Heckman es levemente mayor, que el MSE en el caso las MCO – RSE, lo que se traduce en intervalos de confianza más amplios para los coeficientes y para los pronósticos.

Figura 13 Ingreso Estimado y Edad – Heckman (SL)



En la Figura 13 se presentan las estimaciones de ingreso por edad para diferentes niveles de educación y la Figura 14 presenta algunos de los intervalos de confianza por niveles educativos. La Figura 15 presenta las estimaciones del ingreso por años de educación y en la Figura 16 se encuentran sus intervalos de confianza. Al igual que en la sección anterior, los años de educación superior se refieren a años en el sistema universitario y post universitario.

Figura 14 Ingreso Estimado y Edad (Intervalos de Confianza) – Heckman (SL)

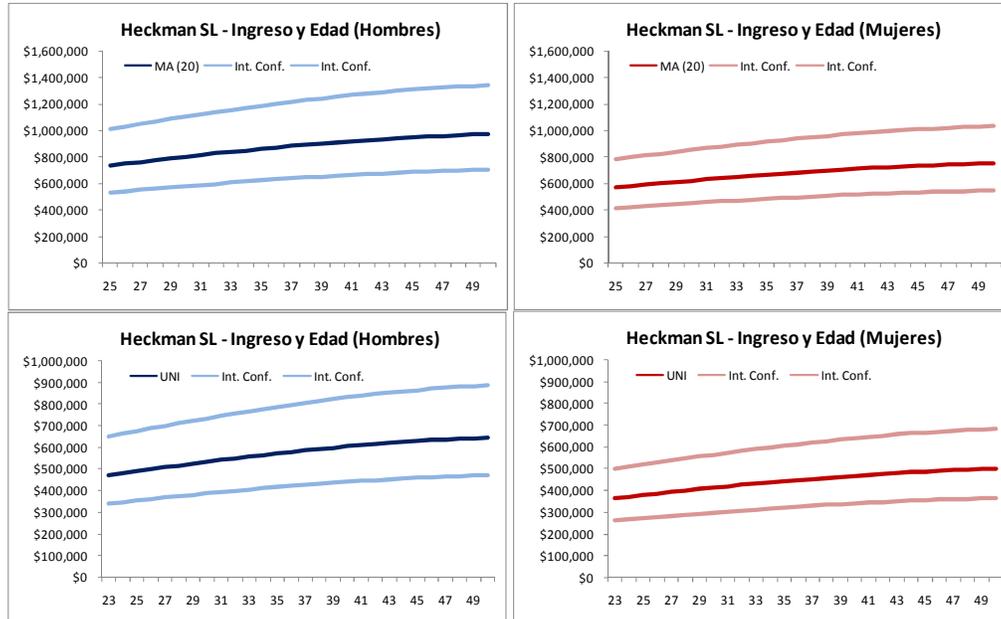


Figura 14 (Continuación)

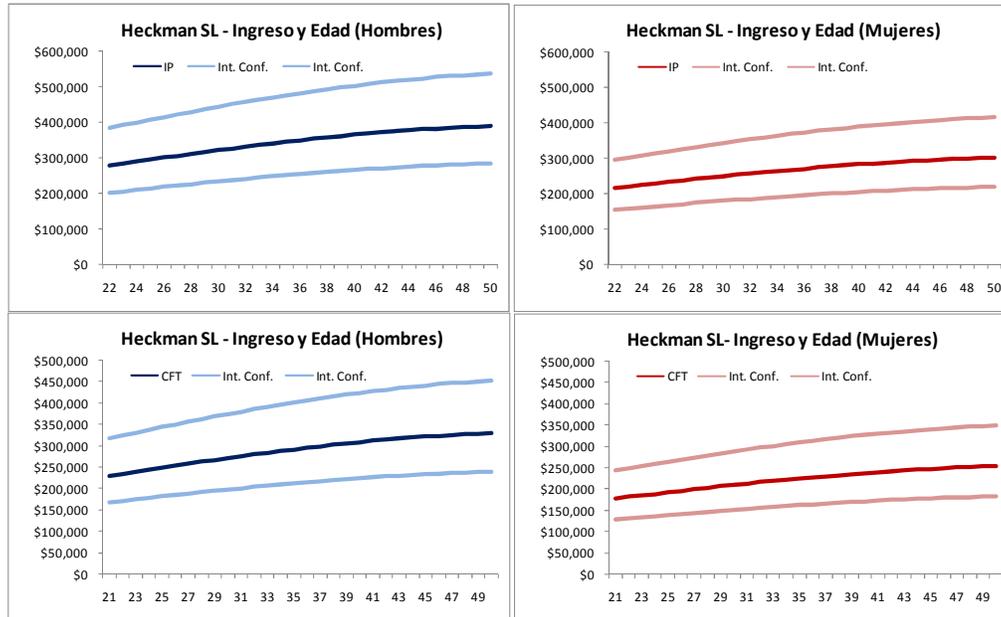


Figura 15 Ingreso Estimado y Años de Educación – Heckman (SL)

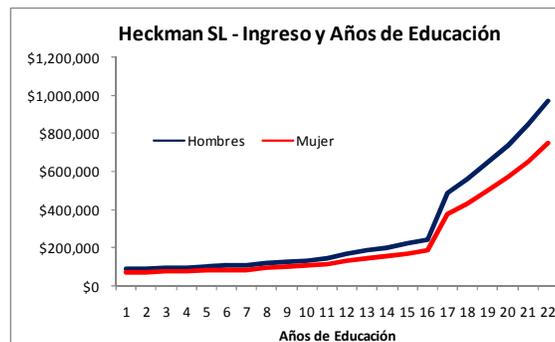
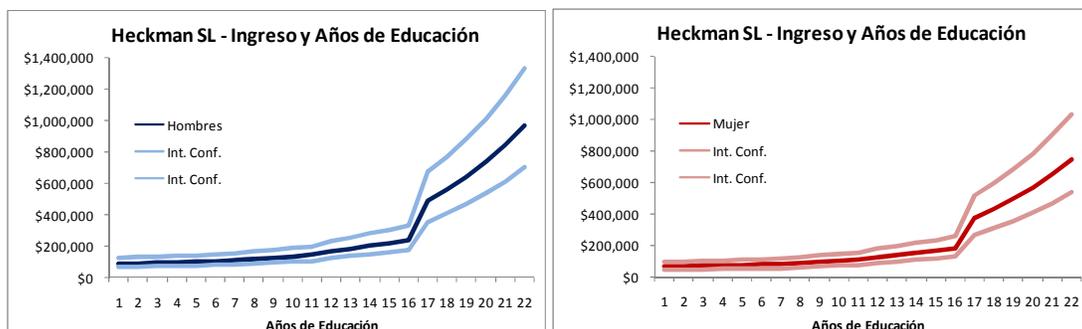


Figura 16 Ingreso Estimado y Años de Educación (Intervalos de Confianza) – Heckman (SL)



Para el modelo de Splines Lineales también se separaron las estimaciones por género y se presentan en el Anexo A.4 los resultados para los coeficientes que diferencia el modelo de SL con en el MM. Para los demás coeficientes los resultados fueron similares al ejercicio del MM con las estimaciones separadas por géneros.

Estimaciones Mincerianas (1990-2006)

Para las estimaciones se tomaron en cuenta las observaciones de la CASEN desde 1990 hasta 2006 que hubieran respondido a la preguntas de hasta qué curso y hasta en qué nivel habían realizado estudios, diferentes de Educación Especial. Igualmente, en las regresiones se tomaron en cuenta sólo observaciones que declararan algún tipo de salario laboral y que declararan estar empleados por el sector público o privado. Se utilizó la variable construida por CASEN de años de escolaridad, y en este caso se ignoraron las diferencias que existen entre los sistemas antiguos y contemporáneos de educación básica y media por simplicidad de cálculo.

Metodologías

Para acercarse al problema del tiempo, primero se hizo una estimación *pooled* que agrupa a todas las observaciones y que sólo se enfoca en las grandes diferencias entre un año y otro. Posteriormente se estimaron los modelos individuales para cada año. En ambos casos se construyó un individuo representativo, para diferentes características educativas, que permitieran hacer un seguimiento a los pronósticos de las funciones mincerianas a lo largo del tiempo para hacer comparaciones.

Resultados

A continuación se presentan estimaciones para las diferentes versiones de CASEN con el objetivo de mostrar el impacto que ha tenido el tiempo sobre ellas. Sólo se utilizó la metodología de MCO – RSE debido a que a lo largo de la serie se hicieron cambios en las mediciones de las características socio demográficas, y por simplificar el procedimiento se obvió la estimación de un modelo más complejo de participación laboral y de corrección de riesgo de selección. En esta sección se presentan los resultados de dos tipos de estimaciones de las series temporales, la primera agrupa todos los datos (*pooled*) e incluye una variable dicotómica para cada año. En el segundo las se estimaron los modelos separadamente para cada año. Las estimaciones se hicieron bajo una transformación en donde $IPC_{2000} = 1$, usando datos del IPC proveído por el INE.

Modelo Minceriano - Pooled

Los resultados para el MM con RSE se presentan en la Tabla 8. El modelo resulta significativo a juzgar por la prueba F, y se puede rechazar la hipótesis de que todos los coeficientes son conjuntamente iguales a

cero. El valor del R cuadrado no es muy elevado, de 0.37, pero se mantiene en el nivel de este tipo de estimaciones para otros estudios.

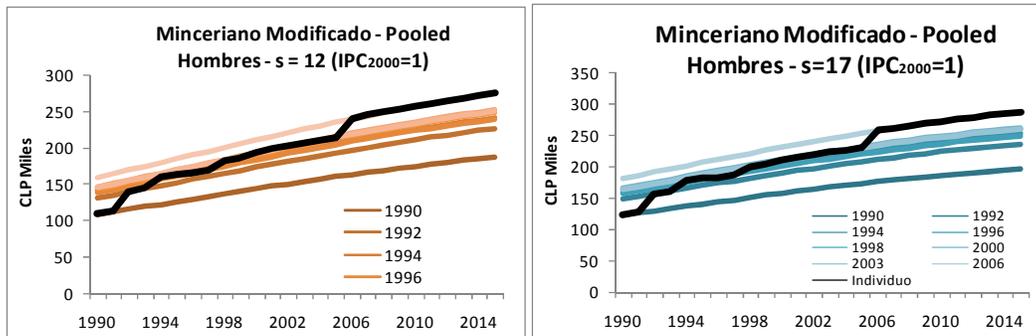
Todos los coeficientes del MM resultaron significativos al 1% y con los signos esperados. El año de 1990 se utilizó como base. Cada coeficiente para cada año tiene un resultado positivo, debido al crecimiento económico de la década de los 90, que hace que este año tenga el salario más bajo como base. Sin embargo, es notable el impacto de la crisis asiática sobre los salarios de los trabajadores. Al realizar la prueba de Wald para evaluar la hipótesis nula de que los coeficientes de 1998 y 2003 son iguales, ésta no puede ser rechazada. Haciendo la misma prueba para el coeficiente del año 2000 con cada uno e los dos años en cuestión, la hipótesis nula podía ser rechazada. Lo anterior se puede implicaría que los salarios reales de las personas con empleadas siguieron creciendo después de haber empezado la crisis, en el periodo 98-00, lo que implicaría una rigidez de los salarios reales. Y es en el periodo 00-03 que los efectos de la crisis son absorbidos por los salarios, seguramente por el ingreso de nuevas generaciones a la fuerza laboral. A este nivel no se puede interpretar si la rigidez en los salarios contribuyó en la duración de la crisis y en la habilidad de adaptación del sistema productivo chileno, que es algo que en otros estudios se argumenta (Gatica y Romaguera 2005).

Tabla 8 Estimaciones CASEN (90 – 06) – MM Pooled

ln y real ('00=100)	Coef.	Rob. Std. Err.	t	P> t	[95% Conf. Interval]	
s	0.0989	0.0003	370.2300	0.0000	0.0984	0.0994
edad	0.0491	0.0005	90.9600	0.000	0.0481	0.0502
edad2	-0.0004	0.0000	-59.0900	0.000	-0.0004	-0.0004
ln_hr	0.3382	0.0039	86.8000	0.000	0.3305	0.3458
sexo	-0.2494	0.0021	-117.8400	0.000	-0.2536	-0.2453
regiones	-0.1758	0.0020	-86.6100	0.000	-0.1797	-0.1718
1992	0.1845	0.0048	38.0500	0.000	0.1750	0.1940
1994	0.2561	0.0047	54.1800	0.000	0.2469	0.2654
1996	0.2393	0.0049	49.1800	0.000	0.2298	0.2489
1998	0.2833	0.0050	57.0700	0.000	0.2736	0.2930
2000	0.2949	0.0049	60.6900	0.000	0.2854	0.3045
2003	0.2876	0.0048	59.9300	0.000	0.2782	0.2970
2006	0.3835	0.0048	79.5900	0.000	0.3741	0.3930
alfa	7.8048	0.0232	336.4800	0.000	7.7594	7.8503

La Figura 17 muestra los pronósticos para cada uno de los años de CASEN, y simula un individuo que ha sido evaluado en cada momento del tiempo. En el caso de la educación secundaria, este individuo es un hombre que ingresa al mercado laboral a los 21 años con 12 años de escolaridad en 1990. En el caso de la educación universitaria, el individuo ingresa al mercado laboral a los 25 años con 17 años de escolaridad. Las demás características se mantienen como en las secciones anteriores. Las curvas son iguales, debido a que el efecto del tiempo es fijo independiente del nivel educativo o de cualquier otra característica. En las gráficas, según las estimaciones, se puede apreciar como desde 1990 hasta 1998 hubo un aumento continuo en el salario real, mientras que a partir de ese último año el salario real se estancó y que sólo creció debido a que la experiencia del individuo hipotético aumentó. Finalmente a partir del 2003, los salarios volvieron a crecer.

Figura 17 Ingreso Estimado y Edad/Tiempo – MM Pooled CASEN (90 – 06)



Modelo Minceriano – Estimaciones Anuales

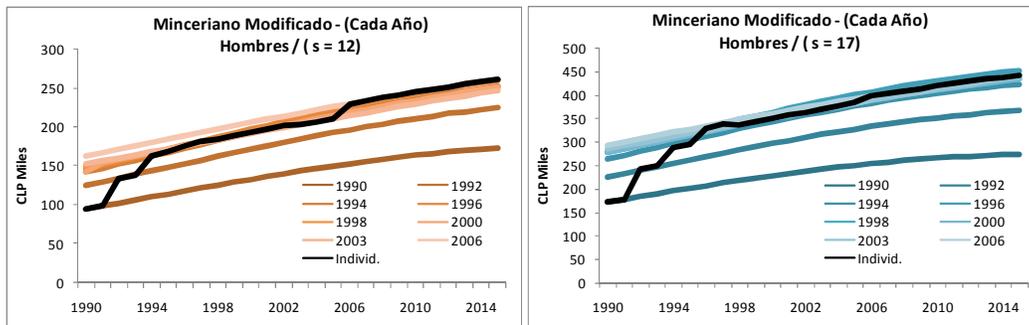
Tabla 9 Estimaciones CASEN (90 – 06) – MM Modelos Anuales

(lnY)2000=1	1990	1992	1994	1996	1998	2000	2003	2006
S	0.0887	0.0919	0.0962	0.1034	0.1024	0.0987	0.1068	0.0977
Rob. Std. Err.	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
P> t	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
edad	0.0668	0.0589	0.0555	0.0558	0.0427	0.0427	0.0349	0.0386
	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
edad2	-0.00064	-0.00053	-0.00049	-0.00049	-0.00033	-0.00032	-0.00023	-0.00030
	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
ln_hr	0.2605	0.2955	0.2634	0.1727	0.3212	0.4538	0.5539	0.3733
	0.03	0.01	0.01	0.01	0.01	0.01	0.01	0.01
	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
sexo	-0.2857	-0.2712	-0.2715	-0.2681	-0.2537	-0.2093	-0.1726	-0.2602
	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
regiones	-0.0427	-0.1557	-0.2153	-0.2463	-0.1811	-0.1479	-0.1812	-0.1303
	0.01	0.01	0.00	0.01	0.01	0.01	0.01	0.01
	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
alfa	7.926	8.091	8.396	8.788	8.266	7.570	7.145	8.209
	0.17	0.07	0.06	0.05	0.06	0.06	0.05	0.06
	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Al correr un modelo para cada año se permite que los coeficientes varíen en cada una de las fotografías de los cortes transversales. La Tabla 9 presenta los coeficientes del MM para cada año. Los modelos individuales fueron significativos bajo las pruebas F de significancia conjunta. Los R cuadrado se ubicaron entre 0.30 para 1990 y 0.46 para 2003. El coeficiente para los años de educación mantuvo una tendencia creciente hasta 1996, a partir de ese año ha tenido un comportamiento relativamente volátil. La brecha de géneros también se fue cerrando a lo largo de la década pasada y hasta 2003. Sin embargo, en la medición de la última versión disponible de CASEN 2006 se volvió a abrir y tanto los datos sin deflactar (de las secciones anteriores) como los datos deflactados tienen un registro superior a 0.25 para el 2006.

El crecimiento del salario también se ha desacelerado. De hecho las personas que ingresaron al mercado a comienzos de los 90s han sido la generación que más rápido ha visto crecer su ingreso real en Chile, seguramente en toda su historia, y pues sobre los datos, en comparación con los 17 años de análisis.

Figura 18 Ingreso Estimado y Edad/Tiempo – MM Anuales CASEN (90 – 06)



La Figura 18 presenta los pronósticos para individuos definidos como en la parte anterior de esta sección con secundaria y educación universitaria completa. En este caso el efecto de la crisis de 1998 sobre los salarios es más notorio y los pronósticos, en 2003, caen bajo el nivel precedente de 2000. El efecto es más notorio en el caso del individuo con Educación Superior. En los pronósticos del MM para el año 2006, no se ha recuperado por completo al nivel de lo que se hubiera pronosticado para 1996 antes de que la crisis estallara.

Modelo de Splines Lineales – Pooled

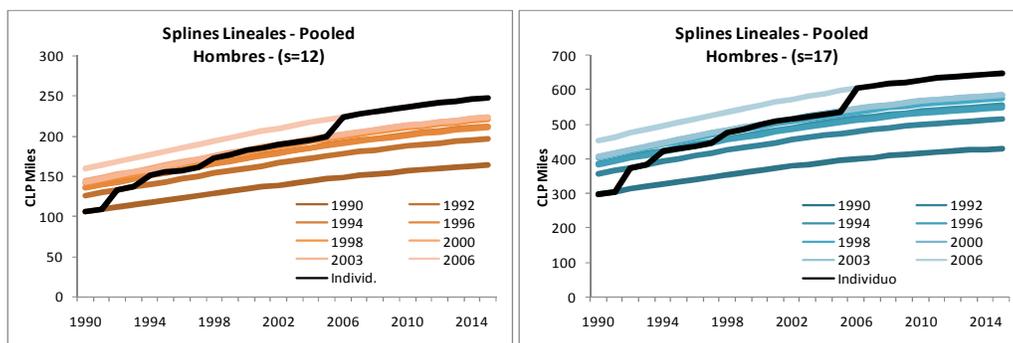
Se hicieron las mismas estimaciones para el caso del modelo de Splines lineales. Primero, en la Tabla 10 se presentan las estimaciones del modelo Pooled. Nuevamente, se encuentra un modelo altamente significativo por la prueba de significancia F. En esta agrupación el R cuadrado es más elevado de un nivel del 0.44, alto para este tipo de estimaciones. Todas las variables resultaron significativas y con los símbolos esperados. Llama la atención que el premio recibido se percibe por la educación superior no universitaria tenga un impacto menor que el de terminar la educación secundaria. Lo anterior no implica que los graduados de la educación terciaria ganen menos que los de la educación secundaria; los coeficientes no son excluyentes, son aditivos por lo que todavía se justificaría este tipo de inversión.

Bajo esta metodología, además de los cambios involucrados con las mediciones de nuevas variables educativas, hay una diferencia con el modelo anterior en los coeficientes de las variables dicotómicas que acompañan a los diferentes años de las versiones de CASEN. En este caso no se puede rechazar la hipótesis nula de que el coeficiente para el año 2000 y el 2003 sean iguales. Sin embargo ambos coeficientes resultan estadísticamente diferentes del coeficiente para el año de 1998. En un lapso de tiempo mayor al comprendido entre 2000 y 2003 los salarios reales dejaron de crecer para adaptarse a la crisis de 1998. La Figura 19 muestra el comportamiento del pronóstico de los salarios y no exhibe ninguna caída a lo largo de la serie.

Tabla 10 Estimaciones CASEN (90 – 06) – SL Pooled

In_yrl	Coef.	Rob. Std. Err.	t	P > t	[95% Conf. Interval]	
esc	0.0408	0.0008	49.61	0.000	0.0391	0.0424
edad	0.0455	0.0005	85.61	0.000	0.0444	0.0465
edad2	-0.0004133	0.0000	-60.91	0.000	-0.0004	-0.0004
ln_hr	0.3625	0.0039	93	0.000	0.3548	0.3701
sexo	-0.2644	0.0021	-128.09	0.000	-0.2684	-0.2604
regiones	-0.1798	0.0020	-92.05	0.000	-0.1837	-0.1760
ϕ prim	0.0241	0.0016	14.83	0.000	0.0209	0.0273
ϕ secund	0.0453	0.0022	21.02	0.000	0.0411	0.0495
ϕ univ	0.0955	0.0041	23.14	0.000	0.0874	0.1036
primaria	0.0394	0.0039	10.1	0.000	0.0317	0.0470
secundaria	0.1067	0.0047	22.75	0.000	0.0975	0.1158
TnoU	0.0803	0.0060	13.3	0.000	0.0684	0.0921
UNI	0.3877	0.0085	45.49	0.000	0.3710	0.4044
1992	0.1799	0.0047	37.92	0.000	0.1706	0.1892
1994	0.2549	0.0046	54.96	0.000	0.2458	0.2639
1996	0.2478	0.0048	51.87	0.000	0.2384	0.2571
1998	0.2949	0.0048	60.89	0.000	0.2854	0.3043
2000	0.3090	0.0047	65.11	0.000	0.2997	0.3183
2003	0.3073	0.0047	65.61	0.000	0.2981	0.3165
2006	0.4126	0.0047	87.24	0.000	0.4033	0.4219
alfa	8.1778	0.0231	354.34	0.000	8.1325	8.2230

Figura 19 Ingreso Estimado y Edad/Tiempo – SL Pooled CASEN (90 – 06)



Modelo de Splines Lineales – Estimaciones Anuales

Las estimaciones para el modelo de SL para cada año se presentan en la Tabla 11. Igualmente, todos los modelos en su conjunto son estadísticamente significativos siguiendo las pruebas F individuales. El R cuadrado oscila en un rango más amplio, desde 0.33 en 1990 hasta 0.51 para el 2003. La mayoría de los coeficientes resultan significativos. Al comienzo de la serie, en 1990 el premio de la educación universitaria y el de la educación primaria no son significativos, sin embargo todos los coeficientes de ϕ lo son. Para el resto de la serie, el coeficiente del premio a la educación universitaria nunca vuelve a ser negativo y de hecho se vuelve el más impactante de todos los premios. En el caso del premio por la educación primaria, vuelve a dejar de ser significativo en 1996. Los cambios en la pendiente en el retorno a la escolaridad medidos con ϕ son significativos a lo largo de la serie: seguir estudiando siempre tiene un impacto sobre los salarios y cada vez mayor.

Tabla 11 SL – Modelos Anuales. Estimaciones CASEN 90 – 06

ln_yrl	1990	1992	1994	1996	1998	2000	2003	2006
S	0.0459	0.0342	0.0478	0.0481	0.0391	0.0377	0.0405	0.0337
Rob. Std. Err.	0.004	0.002	0.002	0.002	0.002	0.002	0.002	0.003
P> t	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
edad	0.0609	0.0537	0.0498	0.0521	0.0390	0.0400	0.0328	0.0382
	0.003	0.001	0.001	0.001	0.002	0.001	0.001	0.001
	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
edad2	-0.000608	-0.000511	-0.000469	-0.000482	-0.000338	-0.000344	-0.000269	-0.000339
	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
ln_hr	0.304	0.342	0.299	0.189	0.344	0.479	0.583	0.385
	0.033	0.013	0.011	0.008	0.009	0.010	0.009	0.010
	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
sexo	-0.3043	-0.2818	-0.2799	-0.2827	-0.2704	-0.2300	-0.1861	-0.2780
	0.010	0.006	0.005	0.006	0.006	0.006	0.005	0.005
	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
regiones	-0.0520	-0.1578	-0.2207	-0.2533	-0.1866	-0.1580	-0.1789	-0.1337
	0.008	0.005	0.005	0.005	0.006	0.006	0.005	0.005
	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
∅ prim	0.0223	0.0212	0.0083	0.0460	0.0284	0.0236	0.0199	0.0220
	0.007	0.004	0.004	0.005	0.005	0.004	0.004	0.005
	0.003	0.000	0.045	0.000	0.000	0.000	0.000	0.000
∅ secund	0.0746	0.0551	0.0556	0.0289	0.0398	0.0496	0.0576	0.0299
	0.009	0.006	0.006	0.006	0.006	0.006	0.005	0.005
	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
∅ univ	0.1659	0.1704	0.1503	0.0981	0.0725	0.0572	0.0222	0.0662
	0.030	0.011	0.010	0.012	0.012	0.012	0.009	0.011
	0.000	0.000	0.000	0.000	0.000	0.000	0.019	0.000
primaria	-0.0276	0.0437	0.0340	0.0104	0.0539	0.0362	0.0491	0.0797
	0.018	0.010	0.010	0.011	0.011	0.010	0.010	0.011
	0.129	0.000	0.000	0.353	0.000	0.000	0.000	0.000
secundaria	0.1052	0.1487	0.1328	0.0479	0.0985	0.1084	0.1015	0.0959
	0.021	0.012	0.012	0.013	0.014	0.013	0.012	0.013
	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
TnoU	-0.1494	0.0821	0.0402	0.0568	0.0869	0.0911	0.0862	0.1676
	0.033	0.018	0.015	0.015	0.018	0.017	0.015	0.015
	0.000	0.000	0.008	0.000	0.000	0.000	0.000	0.000
UNI	0.0087	0.3244	0.3179	0.2595	0.4728	0.4672	0.4804	0.5498
	0.034	0.025	0.022	0.023	0.026	0.025	0.022	0.021
	0.796	0.000	0.000	0.000	0.000	0.000	0.000	0.000
_cons	8.1452	8.3483	8.6903	9.1841	8.6991	7.9752	7.5704	8.6468
	0.182	0.072	0.061	0.051	0.059	0.058	0.051	0.058
	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

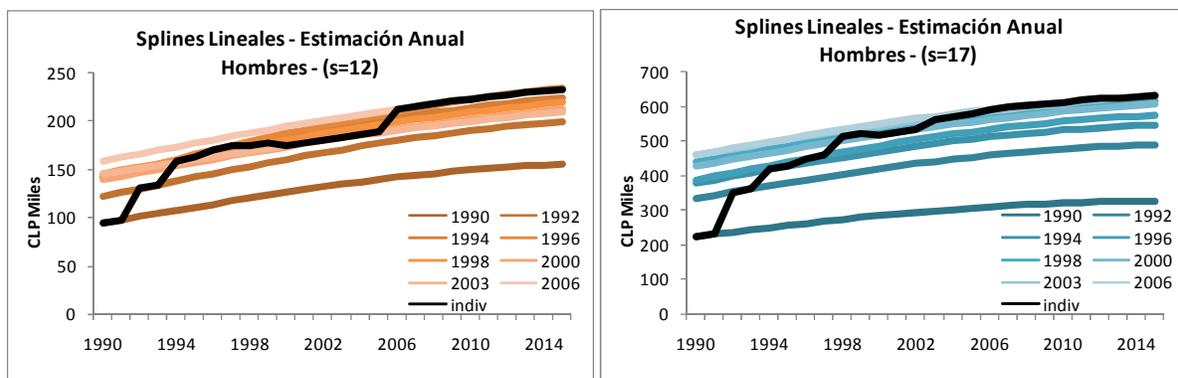
La tendencia se mantiene para varias de las mediciones de correlación que representan los coeficientes. Ya se mencionaba la evolución del premio universitario con un registro bajo en al comienzo de la serie, pero

que se convierte en el de mayor impacto, y constantemente en aumento. Para la educación superior no universitaria la tendencia también ha sido al alza pero no en la magnitud de la educación universitaria.

El aumento en los años de escolaridad ha hecho que el retorno a la educación caiga. La sumatoria de todos los ϕ , que representan el retorno a la educación para un año de educación post-universitario ha caído de 0.31 a 0.15. La misma medición para un año universitario – todos los ϕ sin incluir el universitario – pasó de 0.14 a 0.8 y la medición para el retorno de un año de secundaria también cayó.

El diferencial en el ingreso entre géneros fue desapareciendo paulatinamente a lo largo de los últimos años para reaparecer en la versión de 2006. La diferencia en los salarios de las regiones y de la Región Metropolitana se amplió hasta 1996 y a partir de entonces se ha venido cerrando. Ambas diferencias siguen siendo ampliamente significativas.

Figura 20 Ingreso Estimado y Edad/Tiempo – SL Pooled CASEN (90 – 06)



Al observar, en la Figura 20, los pronósticos para los individuos definidos, graduados de la educación secundaria y de la educación universitaria, existe una sutil diferencia en las curvas. Ambas curvas caen en 1998 pero la recuperación es más lenta en el pronóstico de los graduados de la secundaria, que no se da completo hasta el 2006. Por otro lado, el rebote en el salario de los graduados universitarios se da en 2003, que es la siguiente observación disponible.

Conclusiones

El presente artículo tenía como objetivo la estimación de los salarios de la población chilena diferenciando por nivel de estudio, indagar sobre sus determinantes y su comportamiento. En la presentación de los datos se establece la relación riesgo beneficio que se espera de los niveles más altos de educación: Con una mayor dispersión, la educación superior tiene mayores ingresos esperados y, en términos relativos

concentra una mayor parte de los subgrupos de ingresos elevados. El presente estudio recomienda seguir estudiando formas de distribución de los salarios y mediciones de riesgo de inversiones en capital humano.

En cuanto a la estimación determinística, se analizaron dos tipos de funciones que pretendían incorporar el significado económico de la educación. La primera ecuación, Minceriana Modificada, incorporaba la forma clásica de las ecuaciones de Mincer con algunos cambios que se han introducido a lo largo del tiempo, para controlar por efectos como el género, la ubicación y la elasticidad a la oferta de trabajo. La segunda ecuación, de Splines Lineales, desagregó el coeficiente de años de educación para los diferentes niveles educativos del sistema chileno, midiendo sus premios a graduación y re calculando las tasas de retorno exponencial para cada nuevo nivel comenzado. Independiente de la metodología econométrica utilizada, el modelo de SL se encuentra en mejores condiciones para explicar la varianza en los salarios, y sus estimaciones se acercan más a los valores medios de los subgrupos educativos. En el caso de MM difícilmente se pueden explicar salarios sobre CLP500,000 para recién graduados universitarios, sin tener en cuenta la enorme varianza, mientras que en el modelo de SL lineales se puede encontrar una forma funcional que explique salarios hasta por CLP1,000,000 de forma determinística para los graduados de la educación post-universitaria.

Las estimaciones se llevaron a cabo mediante dos metodologías. La primera, MCO que ante el rechazo de la hipótesis nula de la normalidad de los errores, se utilizó el cálculo de errores estándar robustos (RSE) para las pruebas de significancia de los coeficientes. La otra metodología fue la sugerida por Heckman para la corrección de sesgo de selección y tratando de incorporar algunas características del hogar al hecho de que una persona participe o no activamente en el mercado laboral. Se encontró que se justificaba la corrección del sesgo para el modelo de SL.

Dentro del modelo de participación se encontró que el estado civil y la relación con el jefe de hogar son determinantes de la probabilidad de encontrar trabajo. Tomando a los casados como referencia, los solteros tienen una posibilidad menor de participar en el mercado laboral. Lo anterior se esperaba por el hecho de que la mayoría de personas solteras son estudiantes que asisten a establecimientos educativos. Dentro de las relaciones de parentesco, cualquier persona que no sea jefe de hogar parte con una probabilidad menor de estar en el mercado laboral.

Finalmente, se corrió el MM y el de SL sin la corrección del sesgo para cada una de las versiones de CASEN (90'06). Se encontró que la dinámica del retorno a la educación se puede separar en dos partes, antes de la crisis asiática donde el retorno de la educación superior aumentó para cada nueva versión, y después de la crisis asiática (98-00), cuando el retorno a la educación se estancó y perdió su ritmo de crecimiento. Causas que exploren el por qué de este comportamiento serían una continuación de este estudio.

Referencias

- Beyer, Harald (2000). "Educación y Desigualdad de Ingresos una Nueva Mirada". Centro de Estudios Públicos. Versión modificada de la Serie Documentos de Trabajo N°297, Agosto 1999.
- Guataqui, Juan C. ; García, Andrés; Guerra, José Alberto y Maldonado Darío (2009). "Beyond the Mincer Equation: The internal Rate of Return To Higher Education in Colombia". Bogotá. Universidad del Rosario. Serties Documentos de Trabajo No 68.
- Heckman, James J.; Lochner, Lance; Todd, Petra (2008). "Earnings Functions and Rates of Return". Cambridge, Massachusetts. *NBER* .Working paper 13780.
- Lemieux, T. (2006), "The "Mincer Equation" Thirty Years after Schooling, Experience and Earnings" en Grossbar, S. (Ed.) Jacob Mincer: A Pioneer of Modern Labor Economics, Springer.
- Meller, Patricio y Rappoport , David (2004). "Comparaciones Internacionales de la Dotación de Profesionales y la Posición Relativa Chilena".
- Mizala, Alejandra y Romaguera, Pilar (2003) "Remuneraciones y Tasas de Retorno de los Profesionales Chilenos", mimeo, CEA
- Prada, Carlos F. (2006). "¿Es Rentable la Decisión de Estudiar en Colombia?". Bogotá. Pontificia Universidad Javeriana. (Artículo presentado para obtener el título como Economista en la Pontificia Universidad Javeriana).
- Todd, Petra (2003). "Fifty Years of Mincer Earnings Regressions". Bonn. *IZA*: Discussion Paper Series No. 775.
- Sapelli, Claudio (2009). "Los Retornos de la Educación en Chile: Estimaciones por Corte Transversal y por Cohortes". P.U. Católica de Chile Documento de trabajo 349.
- Wooldridge, Jeffrey (2001). "Econometric Analysis of Cross Section and Panel Data". The MIT Press, Cambridge Massachusetts.

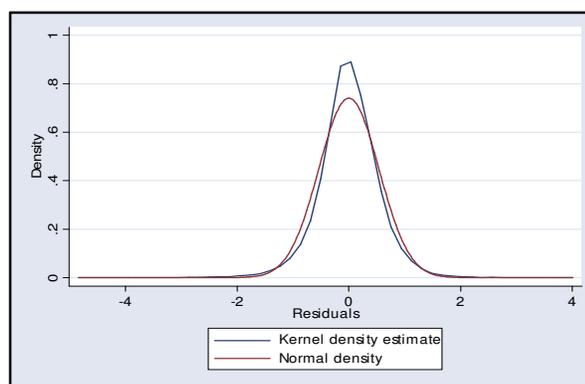
ANEXOS

A.1 Pruebas de Hipótesis – Modelo Minceriano Modificado

Tabla A.1 Pruebas de Hipótesis: Modelo Minceriano Modificado

Shapiro-Wilk W test for normal data			
Variable	Obs	z	Prob>z
Residual	70328	17.715	0
Cameron & Trivedi's decomposition of IM-test			
Source	chi2	df	p
Heteroskedast.	4467.99	24	0
Skewness	367.73	6	0
Kurtosis	237.19	1	0
Variance Inflation Factor (Multicollinearity Test)			
Variable	VIF	1/VIF	
edad2	40.64	0.024604	
edad	40.32	0.024804	
s	1.16	0.860175	
sexo	1.06	0.94618	
ln_hr	1.03	0.970774	
regiones	1.03	0.974112	

Figura A.1 Distribución Empírica Residuos



Los resultados de las pruebas de hipótesis sugieren existen problemas de violación de los supuestos de normalidad y heterocedasticidad. Se corrigieron con el uso de Errores Estándar Robustos (RSE).

A.2 Pruebas de Hipótesis – Modelo Splines Lineales

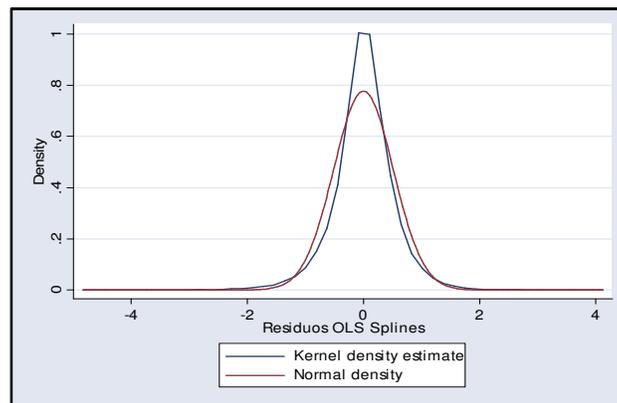
Tabla A.2 Pruebas de Hipótesis: Modelo Splines Lineales

Shapiro-Wilk test for normal data			
Variable	Obs	z	Prob>z
res_1S	70328	18.687	0

Cameron & Trivedi's decomposition IM-Test			
Source	chi2	df	p
Heteroced.	3869.65	77	0
Skewness	345.64	14	0
Kurtosis	258.71	1	0

Variance Inflation Factor (Multicollinearity Test)			
Variable	VIF	1/VIF	
edad2	41.65	0.0247	
edad	41.26	0.024237	
phi_pri	26.76	0.037374	
s	19.71	0.050729	
phi_sec	10.33	0.096845	
secundaria	6.5	0.153746	
UNI	4.34	0.230593	
primaria	4.05	0.24685	
IP	1.59	0.627495	
CFT	1.25	0.80138	
phi_uni	1.15	0.871148	
sexo	1.06	0.939641	
ln_hr	1.04	0.964551	
regiones	1.03	0.970665	

Figura A.2 Distribución Empírica Residuos : Modelo Splines Lineales



Los resultados de las pruebas de hipótesis sugieren existen problemas de violación de los supuestos de normalidad y heterocedasticidad. Se corrigieron con el uso de Errores Estándar Robustos (RSE).

A.3 Corrección de Sesgo por Géneros – Modelo Minceriano Modificado

La Tabla A.3 presenta las estimaciones del modelo de corrección de sesgo separando por género. Aunque no se muestran acá cabe mencionar que los dos modelos separadamente fueron estadísticamente significativos.

Tabla A.3 Heckman por Géneros – MM

Modelo Probit Participacion					
particip	Hombres			Mujeres	
	Coef.	P> z		Coef.	P> z
Esposo(a)/Pareja	-0.2722	0.000	>	-0.5848	0.000
Hijo(a) de ambos	-1.2089	0.000	<	-0.9914	0.000
Hijo(a) jefe	-0.9066	0.000	<	-0.7703	0.000
Hijo(a) pareja	-1.0721	0.000	<!	-0.8880	0.000
Padre o Madre	-0.2739	0.112	>	-0.4239	0.000
Suegro(a)	-0.7582	0.000	==	-0.6826	0.000
Yerno o nuera	-0.1305	0.055	>	-0.8456	0.000
Nieto(a)	-1.9222	0.000	<	-1.5331	0.000
Hermano(a)	-0.0565	0.290	>	-0.3126	0.000
Cuñado(a)	-0.2132	0.010	>	-0.4098	0.000
Otro familiar	-0.6618	0.000	==	-0.6901	0.000
No familiar	-0.1372	0.035	>!	-0.2376	0.000
Serv. Doméstico	NA	NA	<	2.7077	0.000
Union Libre	0.3537	0.000	>	0.2315	0.000
Anulado(a)	-0.5013	0.014	<	0.3754	0.006
Separado(a)	0.0484	0.292	<	0.6395	0.000
Divorciado(a)	0.1629	0.551	<!	0.4631	0.004
Viudo(a)	-0.7426	0.000	<	-0.3337	0.000
Soltero(a)	-0.7231	0.000	<	0.2407	0.000
NS/NR	-0.0725	0.856	==	-0.2370	0.601
Numper	0.0335	0.000	>	0.0111	0.001
educ_jefe	-0.0882	0.000	<	-0.0377	0.000
s	0.1394	0.000	<	0.1540	0.000
edad	0.0094	0.000	<	0.0130	0.000
regiones	-0.1515	0.000	>	-0.2339	0.000
_cons	0.4851	0.000	>	-1.5665	0.000
Modelo de Ingresos (MM)					
In_y	Hombres			Mujeres	
	Coef.	P> z		Coef.	P> z
s	0.0928	0.000	<	0.1144	0.000
edad	0.0400	0.000	>	0.0236	0.000
edad2	-0.0003	0.000	<	-0.0001	0.000
In_hr	0.3981	0.000	<	0.5614	0.000
regiones	-0.1369	0.000	>	-0.1983	0.000
alfa	8.1863	0.000	>	7.0580	0.000
mills					
lambda	-0.0697	0.000	<	0.0999	0.000

Desde el modelo de participación se pueden apreciar diferencias importantes entre los géneros, que no se podían apreciar al considerar hombres y mujeres en el mismo grupo. La Tabla A.3 muestra cual de los estimadores es mayor al considerar los intervalos de confianza de cada uno: “=” representa la presencia del estimador de un género en el intervalo de confianza del otro. El signo de admiración representa que aunque hay intersección en los intervalos de confianza, los coeficientes de un género no se encuentran dentro del intervalo de confianza del otro.

Las mayoría de las características socio demográficas acá consideradas en el modelo de participación parecen ejercer una influencia diferente sobre hombres y mujeres. Por mencionar un caso la educación del jefe del hogar influye negativamente sobre la participación de un individuo, probablemente porque el jefe quiere que los miembros de su hogar persigan estudios más avanzados; sin embargo, en el caso de las mujeres el coeficiente es mucho menor, lo que podría indicar que a un determinado nivel educativo, los jefes quieren que las mujeres de su hogar también trabajen. El efecto de la educación del individuo también ejerce un mayor efecto en el caso de las mujeres. Ambos resultados se deben interpretar sobre la base de que el modelo de participación parte de una constante diferente para hombres (0.49) y mujeres (-1.57), así, todo lo demás constante, los hombres están más dispuestos a participar en el mercado laboral.

En el modelo de salarios los coeficientes también aparecen con diferencias significativas: Los años de educación tienen un mayor impacto sobre los salarios de las mujeres que sobre el de los hombres. Por otro lado, el salario de los hombres crece a una tasa superior en los hombres, pero esa tasa se desacelera más rápido que en el caso de las mujeres. Finalmente, se hace mención al hecho de que al considerar las

regresiones separadas por géneros el coeficiente γ_1 que acompaña a lambda es significativo para los dos modelos considerados.

En conclusión, es mucho lo que falta por analizar en cuestión de género en los resultados del mercado laboral en Chile, de cómo las relaciones dentro del hogar influyen en la decisión de participar de las mujeres a diferencia de los hombres.

A.4 Corrección de Sesgo por Géneros – Modelo Splines Lineales

El mismo ejercicio anterior se corrió para el modelo de SL, la mayoría de los coeficientes de las variables socio demográficas consideradas mantuvieron su nivel y dirección, en comparación con el modelo MM. La Tabla A.4 presenta los coeficientes del modelo de selección y del modelo de salarios para aquellos coeficientes relacionados con las variables específicas de los splines y los años de educación, para ver si esta especificación arroja nuevas luces sobre los efectos sobre la participación y los ingresos.

Tabla A.4 Heckman por Géneros – SL

Modelo Probit Participacion					
particip	Hombres			Mujeres	
	In_y	Coef.		P> z	Coef.
s	0.0233421	0	<	0.0575337	0
primaria	0.2839389	0	>	0.1253783	0
secundaria	0.6797651	0	>	0.5117616	0
terciaria	0.7997891	0	==	0.8081841	0
constante	1.283127	0	>	-0.8591183	0
Modelo de Ingresos (MM)					
In_y	Hombres			Mujeres	
	Coef.	P> z		Coef.	P> z
s	0.0322353	0	==	0.0296935	0
phi_pri	0.0299473	0	==	0.0264252	0
phi_sec	0.0176102	0.001	<	0.0446126	0
phi_uni	0.0975423	0	>	-0.0160067	0.225
primaria	0.0749197	0	>!	0.0202256	0.329
secundaria	0.0676092	0	<	0.1025293	0
CFT	0.1258259	0	==	0.1042225	0
IP	0.1912495	0	==	0.1985103	0
UNI	0.6148109	0	==	0.5877585	0
alfa	8.626183	0	>	7.95702	0
lambda	-0.1019108	0		0.0351052	0.009

El acabar un nivel educativo aumenta la probabilidad de los hombres participen en el mercado laboral en mayor medida que en el caso de las mujeres. La excepción la constituye el caso de la educación terciaria: obtener un título de ese nivel afecta la decisión de participar en igual medida a hombres y a mujeres (c. 0.8 desviaciones estándar). Para la educación terciaria se encuentra también que es el nivel que más influye en la decisión de participar, tanto en hombres como en mujeres. Un año más de educación aumenta la probabilidad de participar de las mujeres en mayor proporción que en el caso de los hombres; este resultado es igual al del modelo MM, pero acá se está controlando por el nivel cursado.

En el modelo de ingresos, en el caso de las mujeres, la pendiente de los años de educación no parece cambiar de manera significativa para el nivel universitario, cuando se estudia una maestría o un postgrado. Los resultados también sugieren que las mujeres obtienen un

premio igual al de los hombres por terminar la educación superior, independiente del nivel. En el caso de finalizar la educación primaria, genera un premio mayor para los hombres pero los intervalos de confianza de los estimadores se intersecan. En el caso de la educación secundaria el premio es mayor para las mujeres. Se presentan el intercepto del modelo de participación y del modelo de salarios, pues las anteriores comparaciones sólo se pueden hacer después de corregir la diferencia que existe en la base sobre la que se parte a hacer la estimación: el intercepto.