
DETERMINANTES DE LOS INGRESOS

Mariana Marchionni

mariana@depeco.econo.unlp.edu.ar

CEDLAS – Universidad Nacional de La Plata

Ecuaciones de Mincer (1974)¹

$$(1) \quad \ln Y = \alpha + \beta \text{educación} + \gamma \text{experiencia laboral} + \delta Z + \mu$$

- Interpretación: modelo de precios hedónicos. El bien *mano de obra* tiene múltiples características. Su precio en equilibrio (el salario de mercado) puede interpretarse como que surge de la interacción de diferentes mercados implícitos, cada uno correspondiente a una de esas características.
- El análisis empírico de las ecuaciones de Mincer permite estimar la contribución de cada una de las características del trabajador a su salario.
- ¿Cuáles son las características que determinan el salario? Capacidades adquiridas como consecuencia de la inversión en capital humano (educación formal, experiencia laboral), capacidades debidas a habilidades innatas del individuo (“inteligencia”), tipo de empleo (rama de actividad, por ejemplo), base de contrato part time o full time, género y otros (pertenencia a unión sindical, raza, región).

¹ Mincer (1974). Schooling, experience, and earnings. National Bureau of Economic Research. Columbia University Press. New York.

El coeficiente β tiene nombre: Retorno a la educación
 derivando (1) con respecto a la educación se obtiene:

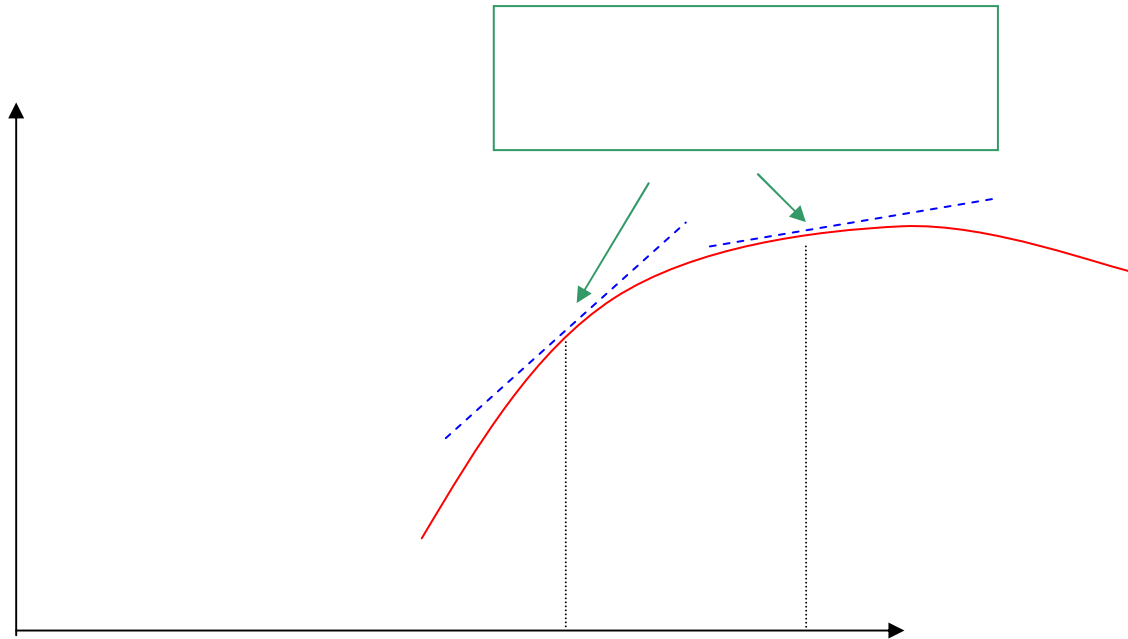
$$(2) \quad \beta = \frac{\partial \ln Y}{\partial \text{educación}} = \frac{\partial Y}{\partial \text{educación}} \frac{1}{Y} \cong \frac{dY/Y}{d(\text{educación})} = \text{retorno a la educación}$$

- Esta interpretación es válida si
 - (i) el único costo de la educación es el ingreso no percibido y
 - (ii) el incremento del ingreso debido a un año adicional de educación es constante.
- Con relación a (ii): podría pensarse que la verdadera relación entre $\ln Y$ y educación no es lineal. Por ejemplo, podría ser del tipo:

$$(3) \quad \ln Y = \alpha + \beta_1 \text{educación} + \beta_2 \text{educación}^2 + \gamma \text{experiencia} + \delta Z + \mu$$

En este caso el retorno a la educación no es constante, depende de la “cantidad” de educación ya adquirida: $\beta_1 + 2\beta_2 \text{educación}$

Ejemplo: $\beta_1 > 0$, $\beta_2 < 0$



- La función de ingresos ya no es lineal: crece a tasa decreciente a medida que aumenta la educación.
- El retorno a la educación no es constante: es positivo (al menos hasta cierto nivel educativo) pero decreciente con la educación.

Cuestiones prácticas relativas a la estimación de ecuaciones de Mincer (o del retorno a la educación)

- I) **¿Qué ingresos son relevantes?** → Los ingresos de fuente laboral
- Es más obvia la relación entre variables de capital humano y este tipo de ingresos.
 - Otros ingresos (especialmente los de capital) son captados con mayor error por las encuestas de hogares (subdeclaración)
- II) **¿En qué unidad de medida conviene expresarlos?** → Por hora.
- Usualmente se considera al salario horario como dado y a las horas trabajadas como elegidas por el individuo. De esta forma, el concepto relevante cuando lo que interesa medir son los retornos es el de salario horario.
 - Aún así, existen diferencias entre contratos en base part o full time, de manera que deberían usarse variables que controlen por el número total de horas trabajadas.
 - Una excepción: cuando hay desempleo puede ser conveniente usar el salario total. Suponer que el desempleo entre los que tienen educación primaria es del 50%, y el otro 50% que consigue trabajo recibe un salario de \$10 por hora. Por otra parte, no hay desempleo entre los que tienen educación secundaria y el salario horario para un trabajador de este grupo también es de \$10. El retorno a la educación es del 100%, pero un análisis basado en el salario horario mediría un retorno nulo por no considerar la diferencia en probabilidades de conseguir trabajo.

III) ¿Cómo se mide la educación?

- Lo ideal sería contar con una medida del aporte de la educación formal al capital humano del individuo. Esto guarda relación con la calidad de la educación.
- En las encuestas de hogares, fuente habitual de información sobre estas variables, sólo se cuenta con mediciones de la cantidad de tiempo que el individuo pasó dentro del sistema educativo.
- Puede contarse con información de los años de educación formal o con información sobre completitud o no de los distintos niveles educativos.
- Consideremos las siguientes variables *dummy* o binarias para indicar el máximo nivel educativo alcanzado por el individuo:

pric =1 si máx. nivel. educativo alcanzado es primario completo, =0 en caso contrario

seci =1 si máx. nivel. educativo alcanzado es secundario incompleto, =0 en c.c.

secc =1 si máx. nivel. educativo alcanzado es secundario completo, =0 en c.c.

supi =1 si máx. nivel. educativo alcanzado es superior incompleto, =0 en c.c.

supc =1 si máx. nivel. educativo alcanzado es superior completo, =0 en c.c.

- Si todos los niveles = 0, el individuo tiene un nivel educativo menor que primario completo (categoría base)

- Cuando se usan estas categorías educativas, el modelo queda especificado como sigue:

$$(4) \quad \ln Y = \alpha + \beta_1 \text{pric} + \beta_2 \text{seci} + \beta_3 \text{secc} + \beta_4 \text{supi} + \beta_5 \text{supc} + \gamma \text{experiencia} + \delta Z + \mu$$

Notar: como la categoría omitida en este caso es educación primaria incompleta, todos los coeficientes de las *dummies* educativas deben interpretarse con relación a esta categoría.

¿Cómo se interpreta β_3 , por ejemplo?

$\beta_3 \times 100$ es el cambio porcentual en el ingreso horario que deberíamos esperar obtenga un individuo que aumenta su educación desde primaria incompleta a secundaria completa. La interpretación de retornos sigue siendo válida pero debe ser readaptada a cambios discretos en la educación.

Ejemplo:

Supongamos que $\beta_1=0.06$ y $\beta_3= 0.10$.

Un individuo que todavía no terminó la escuela primaria debería esperar que su ingreso aumente un 6% si completa ese nivel. Si continúa sus estudios hasta graduarse del secundario debería esperar obtener un aumento del 10% en su ingreso por hora.

Con retornos a la educación positivos, esperaríamos que $0 < \beta_1 < \beta_2 < \beta_3 < \beta_4 < \beta_5$.

Ventaja adicional por usar categorías educativas en vez de años de educación: permitimos que la relación entre $\ln Y$ y la educación sea no lineal sin restringir a una forma cuadrática como en el modelo (3).

IV) ¿Qué es la experiencia?

- El coeficiente γ que acompaña a la variable *experiencia* en la ecuación de ingresos se interpreta como el retorno de un año adicional de experiencia laboral.
- Idealmente: la experiencia laboral debería medir la contribución al capital humano producto de aprender a hacer el trabajo por la simple práctica o por la capacitación recibida en el lugar de trabajo (*on-the-job training*).
- Imposible de medir en la práctica. Por eso, en aplicaciones empíricas, suele aproximárselo al tiempo o antigüedad en el empleo (o empleos similares).
- En la Encuesta Nacional de Empleo e Ingresos 2004 (ENEI) se preguntó por la cantidad de años que el individuo lleva trabajando en “esa” empresa (la actual). Al preguntar por la antigüedad en el empleo actual se subestima la verdadera experiencia laboral, problema que muy posiblemente se agrave con la edad del trabajador.
- En la Encuesta Nacional de Condiciones de Vida 2000 (ENCOVI) se preguntaba sobre el tiempo total que ha trabajado en “esa” ocupación (la actual) aunque no se haya desempeñado todo el tiempo en la misma empresa.

- Una práctica usual es considerar lo que se ha dado en llamar “**experiencia laboral potencial**”, que se define como $exp = edad - años\ de\ educación - 6$, y mide el tiempo que, potencialmente, un individuo puede haber trabajado.
- Cuando se emplea esta definición de experiencia puede demostrarse la equivalencia, para valores adecuados de los parámetros, entre los dos modelos siguientes:

$$(1) \quad \ln Y = \alpha + \beta \text{educación} + \gamma \text{experiencia} + \delta Z + \mu$$

$$(5) \quad \ln Y = A + B \text{educación} + G \text{edad} + \delta Z + \mu$$

Reemplazando $experiencia = edad - educación - 6$ en (1)

$$\ln Y = \alpha + \beta \text{educación} + \gamma (edad - educación - 6) + \delta Z + \mu$$

Reordenando se obtiene:

$$\ln Y = (\alpha - 6\gamma) + (\beta - \gamma) \text{educación} + \gamma \text{edad} + \delta Z + \mu$$

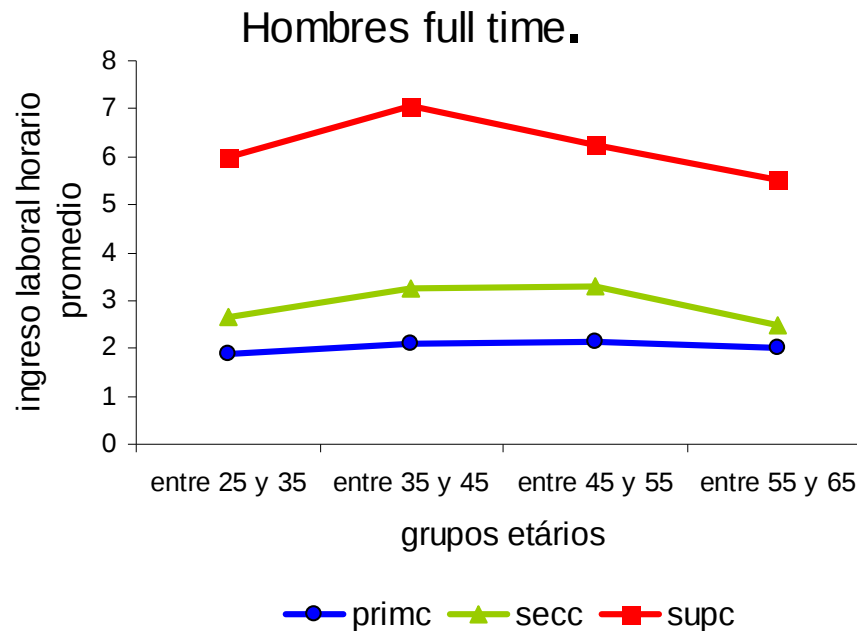
que equivale a (5) si se define: $A = \alpha - 6\gamma$, $B = \beta - \gamma$ y $G = \gamma$.

Supongamos que estimamos el modelo (5). Cómo interpretamos los resultados?

- El retorno a la edad es idéntico que el retorno a la experiencia (luego mirar G en la ecuación (5) equivale a mirar γ de la ecuación (1).
- Supongamos que tanto la educación como la experiencia laboral contribuyen a aumentar el salario ($\gamma > 0$ y $\beta > 0$).
 - Entonces $\beta > B (= \beta - \gamma)$
 - Intuición: β mide el efecto bruto de la educación sobre el ingreso, que no contempla el hecho de que cada año adicional de educación implica un año menos de experiencia. Con un retorno positivo sobre la experiencia el efecto neto de la educación sobre el ingreso debe ser menor que β .
 - Conclusión: si lo que nos interesa es conocer el retorno bruto a la educación (β) pero usamos el modelo (5) (que incluye edad entre las variables explicativas), no deberíamos mirar al estimador de B sino computar $B + G = B + \gamma = \beta$.

- Miremos cómo son los perfiles de ingresos laborales de acuerdo a la edad.

Perfiles de ingresos-edad.



- Esta evidencia surge de 2 cosas.
 - Posibilidad de **no linealidades en el efecto de la experiencia potencial** (o edad) sobre los ingresos laborales.
 - El efecto de la experiencia difiere con el grupo educativo considerado

- Una especificación de la ecuación de Mincer que contempla estas posibilidades viene dada por:

$$(6) \quad \ln Y = \alpha + \beta \text{educación} + \gamma_1 \text{experiencia} + \gamma_2 \text{experiencia}^2 + \delta Z + \mu$$

Notar que:

$$\text{experiencia}^2 = \text{edad}^2 + \text{educación}^2 - 2 \times \text{edad} \times \text{educación} - 12 \times \text{edad} + 12 \times \text{educación} + \text{cte}$$

Reemplazando en (6) y reordenando obtenemos:

$$(7) \quad \ln Y = \Delta + \Gamma_1 \text{edu} + \Gamma_2 \text{edu}^2 + \Sigma_1 \text{edad} + \Sigma_2 \text{edad}^2 + \Omega \times \text{edu} \times \text{edad} + \delta Z + \mu$$

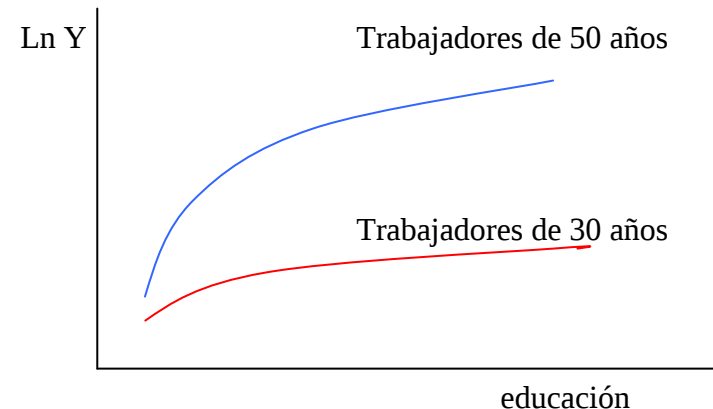
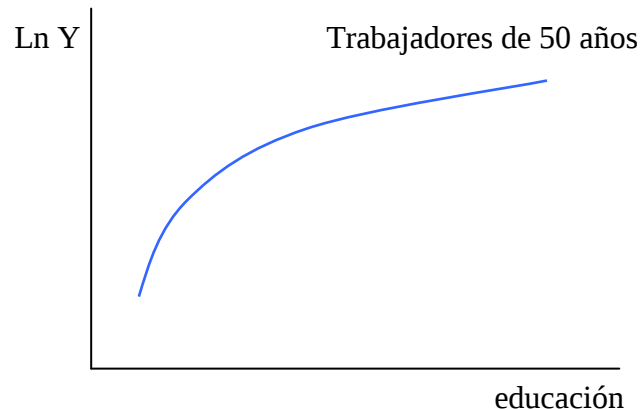
Luego, la inclusión de un término cuadrático para la *experiencia potencial* en (6) implica 3 cosas con respecto a los determinantes de los ingresos:

- (1) el término edu^2 permite estimar retornos a la educación no lineales.
- (2) el término edad^2 permite estimar retornos a la experiencia (edad) no lineales.
- (3) el término de interacción entre educación y edad ($\text{edad} \times \text{educación}$) permite que el modelo prediga perfiles de ingreso-educación no paralelos para distintos grupos etarios.

Retornos a la educación en el modelo (6)

$$\begin{aligned}\frac{\partial \ln Y}{\partial edu} &= \Delta + 2 \times \Gamma_2 \times edu + \Omega \times edad \\ &= \Delta + 2 \times \gamma_2 \times edu - 2 \times \gamma_2 \times edad\end{aligned}$$

- Supongamos que $\gamma_1 > 0$ y $\gamma_2 < 0 \Rightarrow$ el retorno a la experiencia crece a una tasa decreciente.
- Dada la *edad*, si $\gamma_2 < 0$ a mayor educación menor retorno. La pendiente del perfil *lnY-educación* se reduce a medida que aumenta la educación. Si $\gamma_2 < 0$ la pendiente del perfil *lnY-educación* es mayor cuanto mayor la edad



Algunas estimaciones de ecuaciones de Mincer

- Encuesta Permanente de Hogares, Argentina 2002 (EPH)
- Retornos a la educación ¿constantes o variables?

(1) $\ln Y = \alpha + \beta \text{educación} + \gamma \text{experiencia} + \delta Z + \mu$

```
regress logW aedu exp hombre fulltime if muestra==1
```

Source	SS	df	MS	Number of obs = 20231		
Model	2946.25164	4	736.562911	F(4, 20226) = 1633.93		
Residual	9117.72123	20226	.45079211	Prob > F = 0.0000		
Total	12063.9729	20230	.596340725	R-squared = 0.2442		
				Adj R-squared = 0.2441		
				Root MSE = .67141		

logW	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
aedu	.1140719	.0014272	79.93	0.000	.1112746	.1168693
exp	.0146693	.0004667	31.43	0.000	.0137545	.0155841
hombre	.1380559	.010091	13.68	0.000	.1182767	.1578351
fulltime	-.2281073	.0110691	-20.61	0.000	-.2498037	-.2064108
_cons	-.5774771	.0242377	-23.83	0.000	-.6249849	-.5299693

$$(3) \quad \ln Y = \alpha + \beta_1 \text{educación} + \beta_2 \text{educación}^2 + \gamma \text{experiencia} + \delta Z + \mu$$

```
. regress logW aedu aedu2 exp hombre fulltime if muestra==1
```

Source	SS	df	MS	Number of obs = 20231		
Model	3081.19922	5	616.239843	F(5, 20225) = 1387.48		
Residual	8982.77365	20225	.444142084	Prob > F = 0.0000		
Total	12063.9729	20230	.596340725	R-squared = 0.2554		
				Adj R-squared = 0.2552		
				Root MSE = .66644		

logW	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
aedu	.0017736	.0065964	0.27	0.788	-.0111558	.014703
aedu2	.0052886	.0003034	17.43	0.000	.0046939	.0058833
exp	.0135891	.0004674	29.08	0.000	.0126731	.0145052
hombre	.1434822	.0100211	14.32	0.000	.1238399	.1631244
fulltime	-.2168684	.0110061	-19.70	0.000	-.2384413	-.1952956
_cons	-.0464652	.038818	-1.20	0.231	-.1225516	.0296213

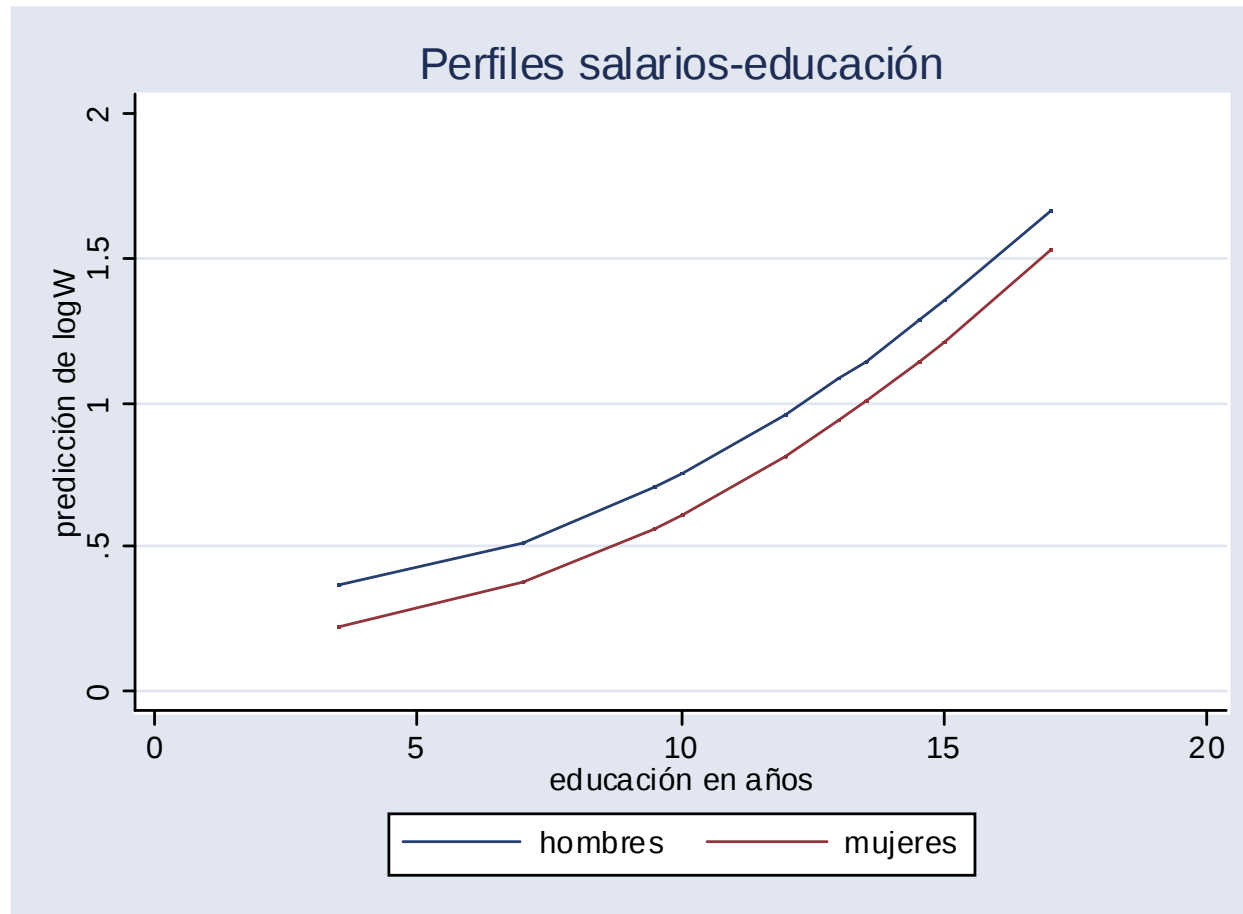
La estimación del modelo (3) sugiere que el retorno a la educación es **significativamente no lineal** (porque el coeficiente de *aedu2* es significativamente distinto de cero).

El retorno a la educación estimado es igual a $-0.0018 + 2 \times 0.005 \times aedu$.

Por ejemplo:

- el retorno estimado para un individuo con 7 años de educación (en Argentina **primaria completa**) es de 0.068 ($-0.0018 + 2 \times 0.005 \times 7$). Es decir, que si una persona que tiene educación primaria completa estudia 1 año adicional, debería esperar un aumento aproximado de su ingreso horario del **6.8%**.
- Por su parte, el retorno esperado para un individuo con **educación secundaria completa** (12 años) es del **11.8%** ($-0.0018 + 2 \times 0.005 \times 12 \times 100$)
- Y para uno con **educación universitaria completa** es casi del **16.8%** ($-0.0018 + 2 \times 0.005 \times 17 \times 100$).

➤ Gráficamente



- Otra alternativa es usar **categorías educativas** como en el modelo 4:

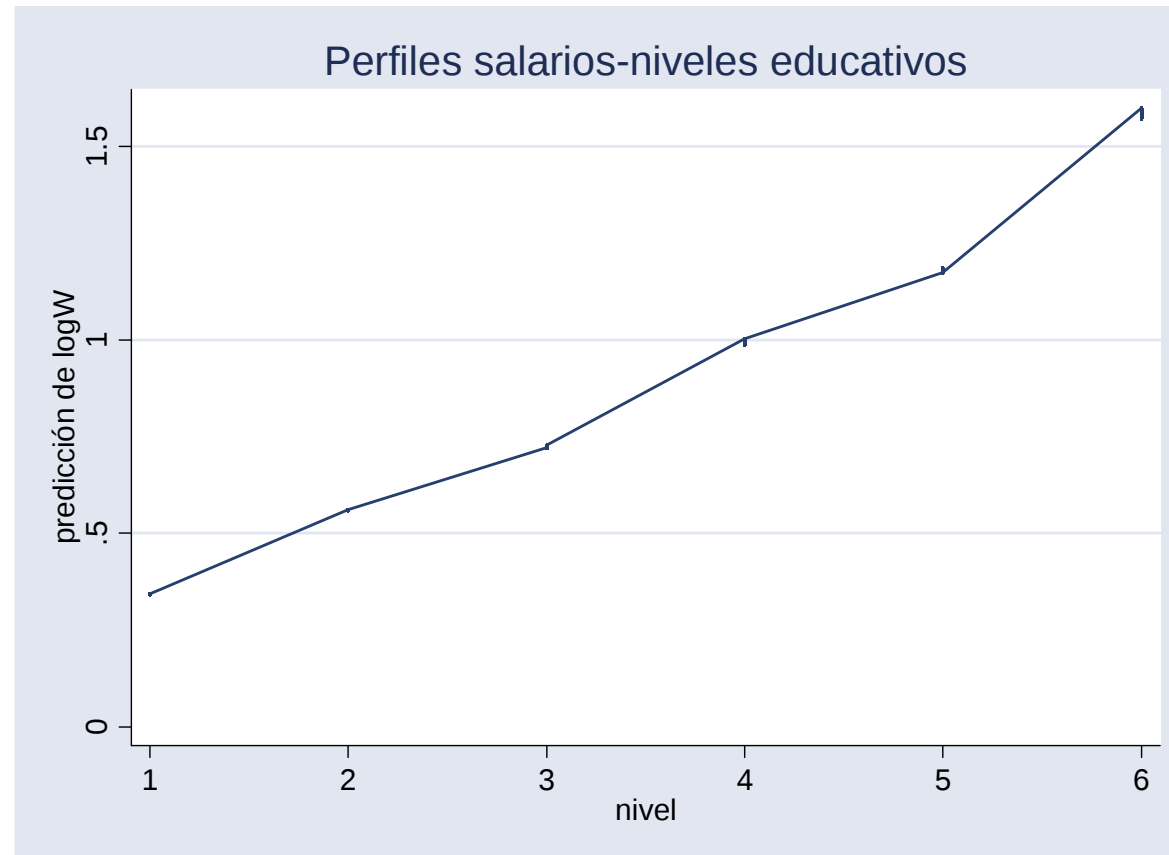
$$(4) \ln Y = \alpha + \beta_1 \text{pric} + \beta_2 \text{seci} + \beta_3 \text{secc} + \beta_4 \text{supi} + \beta_5 \text{supc} + \gamma \text{experiencia} + \delta Z + \mu$$

```
regress logW pric seci secc supi supc exp hombre fulltime if muestra==1
```

Source	SS	df	MS	Number of obs = 20231		
Model	2978.46827	8	372.308534	F(8, 20222) = 828.66		
Residual	9085.5046	20222	.449288132	Prob > F = 0.0000		
Total	12063.9729	20230	.596340725	R-squared = 0.2469		
				Adj R-squared = 0.2466		
				Root MSE = .67029		

logW	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
pric	.2638004	.0188653	13.98	0.000	.2268228	.3007779
seci	.4639453	.0204766	22.66	0.000	.4238096	.5040811
secc	.7713009	.0205492	37.53	0.000	.7310229	.811579
supi	.9760496	.0240361	40.61	0.000	.9289368	1.023162
supc	1.408272	.021825	64.53	0.000	1.365494	1.451051
exp	.0132858	.0004733	28.07	0.000	.012358	.0142136
hombre	.1659392	.010139	16.37	0.000	.1460659	.1858124
fulltime	-.2000454	.0110841	-18.05	0.000	-.221771	-.1783198
_cons	-.0274044	.0243504	-1.13	0.260	-.0751331	.0203243

Gráficamente: predicción del logW para hombres de 40 años, trabajadores fulltime.



- ¿Hay **diferencias por género** en los retornos?

En las especificaciones anteriores se permite que el efecto del género sobre los ingresos opere sólo a través de la constante. Esto impone una forma particular en los perfiles de ingresos: hace que los perfiles de *lnY-educación* de hombres y mujeres sean paralelos.

Es decir, si bien puede diferir el nivel de salario, el retorno a la educación es el mismo para ambos géneros.

Para evaluar la posibilidad de que el género también afecte las pendientes (los retornos), pueden estimarse modelos separados para hombres o mujeres, o elegir una especificación que incluya interacciones entre el género y las demás variables

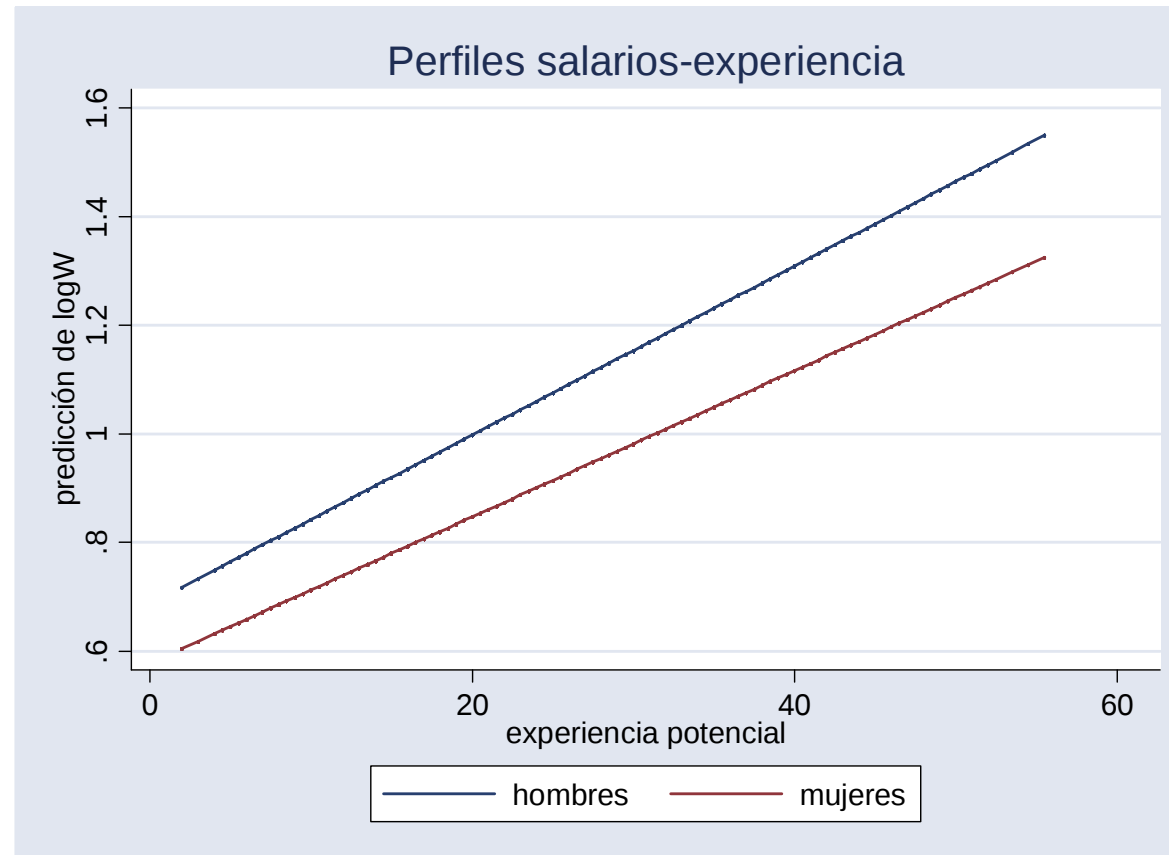
- Comparamos los resultados de estimar modelos para hombres y mujeres por separado (muestra estratificada en base al género) con los de un único modelo que incorpore interacciones del género con todas las demás variables explicativas
- Las dos estrategias son equivalentes

Var. dep. logaritmo del ingreso horario	Modelo para		
	hombres	mujeres	interacciones
aedu	0.114	0.113	0.113
exp	0.016	0.013	0.013
fulltime	-0.188	-0.258	-0.258
hombre			0.029*
aedu x hombre			0.001*
exp x hombre			0.002
fulltime x hombre			0.070
_cons	-0.497	-0.526	-0.526
Obs.	11594	8637	20231
R2	0.23	0.26	0.24

Todos los coeficientes son significativos al 1% o 5% excepto *.

- Según esta evidencia no habría diferencias significativas en el retorno a la educación entre géneros. Por ejemplo: el coeficiente de años de educación en el modelo para hombres es 0.114. El mismo se puede obtener de la tercer columna de la tabla haciendo *coef. años de edu + coef. años de educación × hombre = 0.113 + 0.001*

- Gráficamente: trabajadores fulltime con 12 años de educación formal



Estimamos el efecto interacción entre edad y educación.

$$(7) \ln Y = a + b_1 \text{educación} + b_2 \text{educación}^2 + c \text{educación} * \text{edad} + g_1 \text{edad} + g_2 \text{edad}^2 + \delta Z + \mu$$

```
regress logW aedu aedu2 edad edad2 aedu_edad hombre fulltime if muestra==1
```

Source	SS	df	MS	Number of obs = 20231		
Model	3191.49588	7	455.927982	F(7, 20223) = 1039.19		
Residual	8872.47699	20223	.438731988	Prob > F = 0.0000		
Total	12063.9729	20230	.596340725	R-squared = 0.2645		
				Adj R-squared = 0.2643		
				Root MSE = .66237		

logW	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
aedu	-.0631077	.0088197	-7.16	0.000	-.080395	-.0458204
aedu2	.005681	.0003049	18.63	0.000	.0050833	.0062787
edad	.0456534	.0041538	10.99	0.000	.0375117	.0537951
edad2	-.0005007	.0000436	-11.49	0.000	-.0005861	-.0004153
aedu_edad	.0010109	.0001203	8.40	0.000	.0007751	.0012467
hombre	.1449541	.0099721	14.54	0.000	.1254079	.1645003
fulltime	-.2199202	.0109435	-20.10	0.000	-.2413704	-.1984701
_cons	-.4837848	.1092831	-4.43	0.000	-.6979886	-.269581