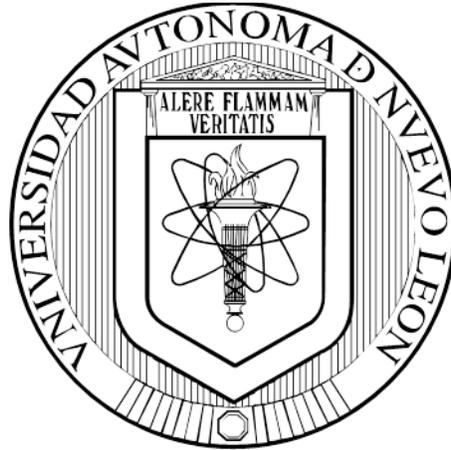


**UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN
FACULTAD DE ECONOMÍA**



TESIS

**“LA EXPERIENCIA Y SU SESGO EN LOS MODELOS DE CAPITAL
HUMANO: EL CASO DE MÉXICO”**

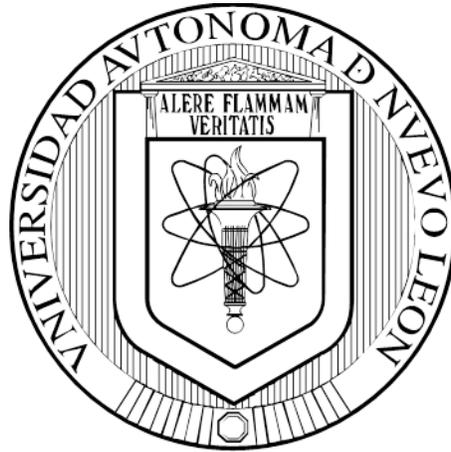
POR

AGUSTÍN TORRES ZAVALA

**PRESENTA COMO REQUISITO PARCIAL PARA
OBTENER EL GRADO DE MAESTRÍA EN ECONOMÍA
CON ORIENTACIÓN EN ECONOMÍA INDUSTRIAL**

DICIEMBRE, 2015

**UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN
FACULTAD DE ECONOMÍA
DIVISIÓN DE ESTUDIOS DE POSGRADO**



TESIS

**“LA EXPERIENCIA Y SU SESGO EN LOS MODELOS DE CAPITAL
HUMANO: EL CASO DE MÉXICO”**

POR

AGUSTÍN TORRES ZAVALA

**PRESENTA COMO REQUISITO PARCIAL PARA
OBTENER EL GRADO DE MAESTRÍA EN ECONOMÍA
CON ORIENTACIÓN EN ECONOMÍA INDUSTRIAL**

DICIEMBRE, 2015

Contenido

	Página
Introducción.....	1
<i>Capítulo 1: Antecedentes y Revisión de literatura.....</i>	2
<i>Capítulo 2: Marco teórico</i>	6
<i>2.1 El modelo.....</i>	6
<i>Capítulo 3: Análisis descriptivo.....</i>	8
<i>Capítulo 4: Estrategia empírica.....</i>	13
<i>4.1 Modelo de inferencia para experiencia real.....</i>	16
<i>Capítulo 5: Estimaciones y resultados.</i>	18
<i>Capítulo 6: Conclusiones</i>	28
<i>Bibliografía.....</i>	30
<i>Anexo 1.....</i>	34
<i>Anexo 2.....</i>	35

Introducción.

El modelo propuesto por Mincer (1974) permite generar una medición del beneficio o premio en el ingreso relacionado con la acumulación de capital humano, siendo el principal referente para la mayoría de los artículos del tema. A pesar de su gran aceptación, dicha representación tiene algunos problemas en su estimación por la relación que existe entre el salario y la educación, conocida como endogeneidad, así como otros sesgos, tales como: autoselección, errores de medición en el salario, educación y experiencia.

Este trabajo se enfoca en la medición de la experiencia. La medida comúnmente utilizada es la experiencia potencial propuesta por Mincer. Debido a que generalmente no se conoce la historia laboral del trabajador, se infiere que su experiencia es la edad menos los años de educación aprobados y sustrayendo los años antes de iniciar la educación preescolar, regularmente seis años. Este supuesto implica sesgos en la medición, especialmente si el trabajador tuvo salidas y entradas del mercado laboral, por lo que, al ignorar estos movimientos laborales se le estarían atribuyendo un mayor número de años de experiencia que los que realmente tiene. Por tanto, el uso de una medida potencial incorpora un sesgo de medición que afectaría el coeficiente estimado de la educación y de la misma experiencia. El objetivo del presente estudio es proponer una alternativa para reducir el sesgo generado por la medida de experiencia utilizada, teniendo en cuenta como medida alternativa la experiencia real del individuo.

La Encuesta Nacional de Ocupación y Empleo (ENOE), aporta características de los individuos que nos ayudan a medir (e inferir, para algunos casos) la experiencia laboral real de una persona. Con base en dicha información, para el año 2014, se proponen ejercicios para generar medidas alternativas de experiencia que mejoren el ajuste de los modelos analizados. Los resultados sugieren que con dicha corrección en la medida de experiencia, se obtienen estimaciones más robustas en el modelo, lo cual representa una alternativa importante para sustituir la experiencia potencial.

La estructura del trabajo es la siguiente: El capítulo 1 resume la literatura existente referente al tema. El capítulo 2 presenta un marco teórico del modelo a seguir en el análisis. El capítulo 3 consiste en una descripción de los datos y criterios utilizados para la medida propuesta y una visión descriptiva del sesgo con los propios datos. En el

capítulo 4 se presenta la estrategia empírica a seguir. En el capítulo 5 se presentan las estimaciones y resultados, mientras que en el capítulo 6 se exponen las conclusiones.

Capítulo 1: Antecedentes y Revisión de literatura

La teoría del capital humano, en general, argumenta que una mayor escolaridad y el nivel de capacitación en el trabajo incrementa la productividad de las personas al adquirir conocimientos y habilidades, lo que implica un incremento esperado en su ingreso futuro. Dada esta perspectiva se considera que la educación, en cualquiera de sus formas, es una inversión, que en el mediano y largo plazo se ve reflejado en ingresos más elevados (Becker, 1964). La mayor parte de los conceptos relacionados a los estudios recientes al tema, se basan en los trabajos desarrollados por Mincer (1958), Schultz (1961) y Becker (1964), pioneros en el tema.

Mincer (1958) muestra que la inversión en capital humano mediante la experiencia adquirida en el trabajo sí influye en las diferencias salariales entre trabajadores y, por tanto, en cómo se determina la desigualdad y el sesgo de dichas ganancias. Schultz (1961) también explica las brechas salariales y de productividad entre individuos por la inversión en capital humano de éstos, mientras Becker (1964) aborda el estudio sobre la inversión de capital humano como la influencia que ésta ejerce sobre los ingresos, estudiando las diferentes formas de inversión (adquisición de grado de escolaridad superior, formación en el trabajo, etc.).

La ecuación de ingresos propuesta formalmente por Mincer en 1974 es la más aceptada y utilizada para la estimación de los rendimientos del capital humano. A pesar de la adopción como estándar de dicho modelo, existen diversos autores que señalan carencias relacionadas en la medición de las variables contenidas en él; esto implica resultados y conclusiones erróneas cuando no se emplea la corrección adecuada o la medición que mejor se ajuste a la realidad.

Griliches (1977) hace un análisis de los posibles sesgos en el planteamiento y resolución de la relación estudiada utilizando Mínimos Cuadrados Ordinarios (MCO), identificando un sesgo por las variables omitidas en el modelo Minceriano (como la habilidad del individuo, por ejemplo). Además de ello cuestiona el cómo se mide la educación y el supuesto de exogeneidad de la misma, lo cual implica una estimación

poco precisa de la tasa de retorno de ésta en el modelo, generando estimadores sesgados e inconsistentes.

El método más común para corregir estos problemas es utilizar Variables Instrumentales (VI). Por ejemplo, existe una corriente en los años 90's en que se utilizan como VI las características de los sistemas escolares. El primer trabajo al respecto fue el de Angrist y Krueger (1991), donde se utiliza el trimestre de nacimiento de los individuos como instrumento, argumentando que las personas que nacen al iniciar el año tienen en promedio una escolaridad menor, ya que alcanzan la edad mínima de conclusión "permitida" antes que los que nacen a final de año. Sus resultados indican que utilizar MCO y su concepción de VI apunta a resultados similares en los rendimientos a la educación mostrados, aunque sugieren cierto sesgo que subestima los estimadores MCO.

Kane y Rouse (1993), por su parte, estudian los retornos de asistir algunos años a la universidad u obtener el grado de educación universitaria. Cuando controlan por antecedentes familiares y habilidad de la persona encuentran que, en promedio, los que estudiaron 2 o 4 años en la universidad ganan 5% más que los graduados de la preparatoria. Utilizando como instrumentos la distancia de la escuela más cercana y las colegiaturas, intentan corregir el error de medición y el sesgo de selección.

En la literatura sobre el tema, dicho error de medición ha recibido especial atención ya que se ha trabajado con muestras de personas vinculadas familiarmente como hermanos, gemelos que tienen distintos niveles de escolaridad. Al suponer que son individuos genéticamente similares o idénticos se busca controlar el sesgo generado por la habilidad de los individuos, que no puede ser medido en una variable. Ashenfelter y Krueger (1994) y Ashenfelter y Rose (1998), utilizan dicha teoría para realizar ejercicios al respecto y corregir el sesgo generado por la omisión de variables importantes.

Diferentes instrumentos se han utilizado para corregir dicho sesgo, como son la escolaridad de alguno de los padres, número de hermanos, empleo del jefe de familia, entre otros. Para Carneiro, Heckman y Vytlacil (2002), la mayor parte de éstos son inválidos, argumentando que de cualquier modo existe correlación entre ellos y la variable omitida.

Otro error de medición abordado en la literatura y que motiva al presente estudio, es el sesgo que se genera al no utilizar la experiencia laboral de un individuo, cuyo valor, cuando no se conoce, se aproxima por la experiencia potencial propuesta por Mincer.

Garvey y Reimers (1980) detectaron que el uso de la experiencia potencial genera un sesgo por dos razones: 1) La gente no necesariamente termina un grado de escolaridad por año; y 2) pueden decidir no entrar al mercado laboral en algún punto de su vida, por lo que admitir un historial laboral continuo desde su último grado escolar resulta inadecuado, sobre todo para el caso de las mujeres. Utilizando como base la experiencia potencial y un modelo lineal de experiencia real, muestran el sesgo en el modelo y generan una predicción de experiencia que mejora el ajuste del modelo Minceriano.

Utilizando dos conjuntos de datos diferentes *National Longitudinal Survey of Youth* (NLSY) y el *Panel Study of Income Dynamics* (PSID), Filer en 1993 propone ampliar la estrategia utilizada por Garvey y Reimers (1980), implementando un modelo para estimar la experiencia real de las mujeres trabajadoras con la información disponible de los que pueda calcularse la experiencia e imputarla a los que no. Concluye que, para el caso de las mujeres, el uso de la experiencia potencial puede sesgar el impacto estimado en factores como la escolaridad y la raza de la persona, sugiriendo que la experiencia real y su predicción de experiencia generan mejores resultados.

Antecol y Bedard (2004) en un intento por explicar los diferenciales entre mujeres jóvenes blancas y sus pares mexicanas y de raza negra, puntualizan que el uso de experiencia real en el mercado laboral es trascendental para estimar con mayor precisión dicha disparidad. Utilizando la NLSY, encuentran que mexicanas y afroamericanas ganan 9% y 15% menos que sus pares de raza blanca.

Oaxaca y Reagan (2008) reconocen la existencia de sesgo al utilizar la medida de experiencia potencial propuesta por Mincer en los modelos de capital humano, donde enfatizan el efecto que dicha medida genera sobre las tasas de rendimiento de la escolaridad (especialmente en el caso de las mujeres). Utilizando tres conjuntos de datos para los cálculos (NLSY79, PSID e IPUMS), proponen un modelo para estimar la experiencia real de los historiales laborales que no pueden calcularse. Concluyen que la experiencia potencial introduce sesgo en la predicción del modelo Minceriano, especialmente en los rendimientos a la educación, tanto en hombres como mujeres, y en el diferencial salarial entre ambos.

En la literatura sobre capital humano, la antigüedad o el tiempo de permanencia en un cargo o empleo es llamado *tenure* y se considera parte del complemento de la experiencia. Autores como Williams, N. (1991), Bronars, S. y Famulari, M. (1997), Bratsberg, B. y Terrel, D. (1998) y Burdett, K. y Coles, M. (2010) utilizan dicho historial como complemento de medición en el modelo Minceriano, es decir, visualizan tres componentes de capital humano en la ecuación: La educación formal medida por los años de escolaridad obtenidos, la experiencia potencial propuesta por Mincer y la antigüedad o *tenure* en dicho empleo.

Para México existen estudios sobre el tema con diferentes enfoques, sin embargo el tema de sesgo en la medida de experiencia no ha sido el tema central en alguno de ellos. Bracho y Zamudio (1994) estiman los rendimientos a la escolaridad mediante la ecuación de Mincer para 1989, con datos de la ENIGH. Resaltan la importancia de diferenciar en el análisis los efectos del género, la zona de residencia, el grupo de edad o el nivel educativo, encontrando rendimientos promedio de 11.7% por año adicional de escolaridad.

Barceinas (2001) utiliza la ENIGH de los años 1994 y 1996 para medir la tasa de rendimiento educativo. Utilizando distintas alternativas metodológicas en la estimación (directo, ecuación de Mincer y elaborado), encuentra tasas de rendimiento que van de 13.5% a 15% al emplear el método tradicional Minceriano.

Utilizando la Encuesta Nacional de Empleo Urbano (ENEU) para el periodo 1988-2002, López-Acevedo (2004) agrega variables como región, sector a la ecuación Minceriana y utiliza el método de regresión cuantílica o *quantile regression* para la medición. Dado que la regresión cuantílica implica retornos a la educación por cuantiles de ingreso, la autora demuestra que la desigualdad educativa explica una parte importante en la desigualdad del ingreso para los trabajadores mexicanos, independientemente si está en el segmento más “rico” o más “pobre” de la distribución.

Rodríguez-Oreggia (2005), buscando explicar las diferencias salariales y obtener los rendimientos de la educación entre regiones en el país, utiliza los datos de la ENEU de 1987-2002 para calcular la ecuación Minceriana. Teniendo en cuenta los efectos de la ubicación geográfica y posibles factores institucionales que afectan la producción, el

autor encuentra que los rendimientos crecen desde 1987 hasta el periodo donde se implementó el TLCAN, decreciendo en el periodo posterior para todas las regiones.

Caamal-Olvera (2013) analiza la tendencia del rendimiento educativo en el país de 1988 a 2013, teniendo como base la Encuesta Nacional de Empleo Urbano (ENEU) y la ENOE. Con base en la técnica de regresión cuantílica, los resultados muestran una tendencia cóncava a través del tiempo, siendo menos pronunciada cuando el modelo es ajustado con variables de regionalización y tipo de industria. Caamal-Olvera (2014) encuentra que los rendimientos de la escolaridad son mayores para las mujeres en relación con los hombres, además, la tendencia a lo largo del tiempo es más decreciente para los hombres que para las mujeres.

Capítulo 2: Marco teórico

2.1 El modelo

Como se mencionó, se sabe que los modelos de capital humano pueden tener distintos tipos de sesgo. El objetivo de esta investigación es determinar si la medida de experiencia laboral elegida tiene algún efecto sobre los parámetros del modelo, principalmente los relacionados al capital humano, al comparar los cálculos derivados de utilizar la medida de experiencia tradicional (potencial) contra los obtenidos utilizando la medida de experiencia a proponer. Teóricamente, dicho sesgo puede verse reflejado en el término de error del modelo, lo que implica menor precisión en los valores estimados.

Para dimensionar la importancia del sesgo por la experiencia, a continuación se explica teóricamente la forma en que está presente. Los ingresos laborales de un individuo están relacionados directamente con ciertas características innatas y, a su vez, con algunas habilidades desarrolladas a lo largo de su vida, siendo trascendentales el grado de educación y la experiencia en el mercado laboral, por lo que dichas variables son esenciales al estimar funciones de ingreso. Por limitación de información, generalmente se utilizan aproximaciones de la experiencia laboral, aunque es preferible tener información real; en la mayoría de los casos, esto no se da.

En este trabajo se emplea como base teórica el marco de la ecuación de Mincer para analizar el efecto de la experiencia elegida sobre las variables de acumulación de capital

humano, en alguna de las formas definidas para este contexto, mediante una versión simplificada.

Suponga que la experiencia real de un trabajador, dadas las características de una encuesta, no puede obtenerse. Cuando se tiene información sobre la edad del individuo y sus años de escolaridad se puede generar una aproximación de la experiencia laboral, que en la literatura sobre el tema se conoce como experiencia potencial (X^*). Por simplicidad, se asume que la expresión que define a la experiencia potencial puede ser:

$$(1) X_i^* = X_i + \varphi_i$$

Donde φ_i es la diferencia entre ambas medidas. Dicho término puede tener correlación con la experiencia real, sin embargo, dada la definición de experiencia potencial, no puede considerarse probable que tenga media de cero, como generalmente se supone en los términos de error. Despejando para la experiencia real y sustituyendo en el modelo de Mincer, el modelo sería:

$$(2) Y_i = \beta_0 + \beta_1 S_i + \beta_2 X_i^* + \beta_3 X_i^{*2} + \sum_{i=1}^k \alpha_i H_i + \varepsilon_i^*$$

$i = 1, \dots, N$ siendo N es el tamaño de la muestra

Donde:

- Y = Logaritmo natural del salario por hora
- S = Años de escolaridad alcanzados por el individuo
- X^* = Experiencia de trabajo potencial
- H = Conjunto de variables de control consideradas en el modelo (Ejemplos: Estado civil, identificadores de región, entre otros.)

$$(3) \varepsilon_i^* = \varepsilon_i - \beta_2 \varphi_i - 2\beta_3 X_i \varphi_i - \beta_3 \varphi_i^2$$

Nótese que el término de error ahora se compone por tres valores además del error tradicional, por lo que la media de los estimadores sería el parámetro poblacional más (o incluso menos) la media de los valores mencionados; lo anterior implica que utilizar la medida de experiencia potencial genera un sesgo implícito en el cálculo.

Capítulo 3: Análisis descriptivo

La ENOE es una encuesta trimestral publicada por el INEGI que busca obtener información sobre las características ocupacionales de la población en nivel nacional, generando también variables demográficas y económicas que permiten profundizar en el análisis de los aspectos laborales. La muestra sigue por cinco periodos consecutivos a una misma familia, alternando cada trimestre al 20% de los individuos. Sin embargo, en este estudio se considera solamente un punto en el tiempo, debido a que el módulo ampliado es el que proporciona la información necesaria para el cálculo de la medida de experiencia propuesta, la cual se nombrará *experiencia real*. Los datos utilizados para realizar las estimaciones provienen del primer trimestre del 2014.

Adicionalmente, la muestra se acota a individuos cuya edad oscile entre 16 y 70 años, que hayan reportado ser empleados subordinados con pago diferente de cero, que hayan trabajado la semana previa a la entrevista y reporten trabajar consecutivamente todos los meses del año, esto con el fin de contar con personas que puedan considerarse trabajadores activos laboralmente estables. Además, no se tomaron en cuenta las observaciones que, en el cálculo de experiencia potencial, mostraron valores inconsistentes (valores negativos), los cuales representaron el 0.03% de la muestra inicialmente elegida (13 observaciones). Dado el objetivo de estudio, se eligió al grupo de los hombres, ya que, al menos teóricamente, se considera más estable cuando se trata de tiempo entre ingreso-salida del mercado laboral, ya que la mujer probablemente tenga periodos de desempleo (voluntario o involuntario) más grandes.

Criterios para generar medida propuesta.

Para establecer el criterio en la medición de experiencia laboral *real*, se toman como base para el cálculo las siguientes preguntas del cuestionario ampliado:

- Año en que inició a trabajar por primera vez (p3r)
- Desde entonces, ¿ha trabajado todos los años para su actual empresa o patrón? (p3s)
- ¿Alguna vez se quedó sin trabajo o negocio y tuvo que buscar otro? (p9)

La pregunta 9 permitirá saber si los entrevistados reportaron antecedentes laborales. Cuando el individuo menciona nunca haber roto el vínculo con la misma unidad

económica¹, se puede considerar que el historial de experiencia completo puede ser calculado (caso 1). Para confirmar la consistencia de la información, se verifica si ha trabajado todos los años para su actual empresa desde el año que menciona haber entrado por primera vez (pregunta 3r). El historial laboral en este caso puede ser considerado como *tenure*, ya que el individuo no ha cambiado de empleo en su vida laboral.

Cuando el individuo acepta haber renunciado o perdido su trabajo en algún punto del tiempo, es posible rescatar una parte de su historial laboral mediante algunas preguntas del cuestionario; sin embargo, el alcance en ese sentido es limitado al trabajo inmediato anterior (Caso 2).

En el siguiente cuadro se observa la distribución de la muestra (sin factor de expansión), en el cual se puntualiza si el trabajador inició sus actividades en la empresa donde labora en el año de la encuesta, el año previo a éste o hace dos años o más.

Cuadro 1: Distribución porcentual de personas por periodo o año de ingreso por primera vez a su actual empresa, 2014

	Personas	Porcentaje
Total	44,608	100.0
Año de reporte	1,034	2.3
Año anterior al de reporte	4,983	11.2
Dos años anteriores o más al de reporte	38,518	86.4
No sabe	73	0.2

Fuente: Cálculos propios de ENOE para cada año.

Como puede observarse, una pequeña parte del grupo de trabajadores entrevistados reporta haber ingresado a su actual empresa por primera vez en dicho año o un año anterior (13.5%), mientras que poco más del 86% mencionan haber ingresado hace dos años o más. Es importante mencionar que, dada la estructura del cuestionario, la pregunta sobre la estancia consecutiva en la actual empresa desde su primer ingreso a ésta, sólo aplica para este último grupo de trabajadores.

¹ Ver Manual del entrevistador de la ENOE, INEGI. Pg. 344-346. Recuperado de http://www.inegi.org.mx/est/contenidos/espanol/metodologias/enoe/manuales/Manual_entrev_ENOE_1.pdf

Dado lo anterior, el cruce para generar de las variables mencionadas - muestra la distribución de únicamente de 38,518 personas en 2014, información que se muestra en términos relativos en el cuadro 2.

Cuadro 2: Personas según estancia consecutiva en el actual empleo desde su entrada a éste por primera vez, y pérdida de empleo en algún punto de su vida laboral, 2014.

Trabajo consecutivo en su empresa actual desde su entrada a ésta por primera vez	Alguna vez perdieron un empleo a lo largo de su vida laboral			
	Porcentaje respecto al total.			
	Si	No	N.S.	Total
Total	33.15	66.39	0.45	38,518
Si	32.49	65.95	0.45	98.90
No	0.65	0.43	0.00	1.08
N.S.	0.01	0.01	0.00	0.02

N.S. – No Sabe

Nota: La suma porcentual puede ser mayor a 100% debido al redondeo de cifras.

Fuente: Cálculos propios con base en la ENOE para el año señalado.

Como se puede observar, un alto porcentaje de personas reportan no haber abandonado el mercado laboral por la pérdida de empleo, con un valor relativo 66.39% respecto al total de la muestra. Visualizando la propuesta de este estudio de los casos para el rescate de la experiencia real, se observa que 25,404 personas en 2014 cuentan con las respuestas al cuestionario contempladas para generar el caso 1 (experiencia laboral completa o *tenure*, en este caso). Dichos casos equivalen a 65.95% del total muestral.

El caso 2 se compone de los individuos que alguna vez se quedaron sin trabajo y tuvieron que buscar otro, los cuales representan el 33.15% de la muestra. Es importante resaltar que existe un pequeño porcentaje de personas de los que no se tiene información concreta en alguna de las preguntas, y representan únicamente el 0.02%.

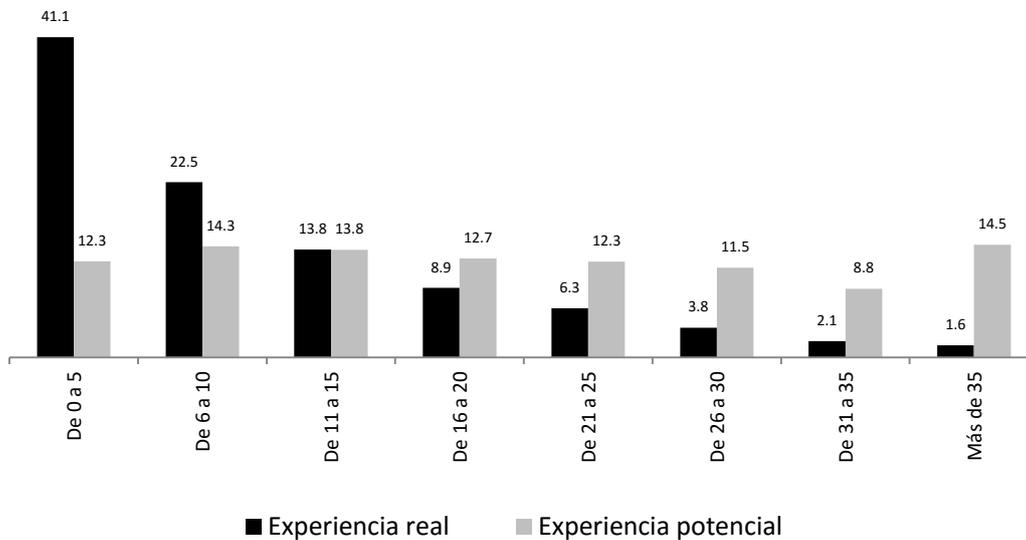
Además, existe un pequeño grupo (0.43%) cuyas respuestas pueden resultar inconsistentes dado el cruce de preguntas, ya que responden nunca haber perdido un empleo en su vida laboral y, al mismo tiempo, no haber trabajado consecutivamente en la empresa en la que actualmente lo hacen desde que entraron por primera vez a dicha institución.

La experiencia potencial se define como:

$$(4) \text{ Exp_pot} = \text{Edad} - \text{Años de escolaridad} - 6$$

Considerando la información calculada de la experiencia real, para el caso 1, se genera un comparativo entre los tipos de medición de la experiencia para las sub-muestras en este ejercicio. Utilizando rangos quinquenales hasta los 35 años para su clasificación se generan las distribuciones de ambas medidas²:

Gráfica 1: Distribución muestral de experiencia real vs experiencia potencial por rangos 2014.



Fuente: Cálculos propios con base en la ENOE 2014 (INEGI)

Al graficar las distribuciones de ambas medidas para el mismo número de observaciones en el que puede rescatarse la experiencia real, se percibe claramente el sesgo que se ha mencionado teóricamente: La experiencia real exhibe una tendencia marcada en los rangos de experiencia más bajos, específicamente los rangos “de 0 a 5” y “de 6 a 10” años, lo que implicaría que aproximadamente 10 mil personas en la muestra, cifra cercana al 40%, se les han asignado más años de experiencia que los realmente laborados. En tanto, la experiencia potencial, dada su naturaleza, presenta un comportamiento similar a una distribución uniforme.

²En el cálculo de los datos del caso 1 (*tenure*), se logró generar la medida para todos los individuos considerados. Para el cálculo de la experiencia potencial de las mismas personas, no se lograron obtener el 0.15% y 0.19% de la muestra debido a limitaciones en el reporte de los años de escolaridad o la edad en la muestra.

Considerando los porcentajes base de la gráfica anterior, se tiene que los rangos de experiencia más elevados reúnen un porcentaje de individuos de 3 a 9 veces mayor en la experiencia potencial respecto a la experiencia real, los cuales se redistribuyen en los rangos más bajos de la medida real. Para ambos años, el porcentaje de personas con experiencia “de 0 a 5” años es tres veces mayor en la experiencia real respecto a la potencial.

El cuadro 3 se presenta para consolidar la información anterior, es importante conocer las características de los trabajadores en los rangos de la medida de experiencia calculados, específicamente de las variables consideradas en el cálculo de la experiencia potencial (edad y escolaridad), con el fin de tener un panorama general sobre qué tan diferentes son realmente los grupos y si sus diferencias son estadísticamente significativas.

Cuadro 3: Promedio de edad según rango de años de experiencia, por año de la encuesta y tipo de experiencia utilizada para generar el rango, 2014.

Años de experiencia	Real	Potencial	Prueba de medias
De 1 a 5	30.7	20.8	Significativa
De 6 a 10	35.3	25.0	Significativa
De 11 a 15	39.5	29.8	Significativa
De 16 a 20	43.3	34.5	Significativa
De 21 a 25	47.0	39.0	Significativa
De 26 a 30	50.5	43.7	Significativa
De 31 a 35	53.1	48.4	Significativa
Más de 35	58.7	55.9	Significativa

Nota: La prueba de medias se realizó para un nivel de confianza de 95%.
Fuente: Cálculos propios con base en la ENOE 2014 (INEGI)

Para el caso de los promedios de edad en los distintos rangos de experiencia manejados, es claro que existe una redistribución de los trabajadores en la muestra de los niveles más elevados en dichos rangos a los inferiores, lo cual implica que no se da por hecho la participación inmediata de la persona en el mercado de trabajo, ya que puede haber casos en lo que, por diversas razones, el individuo deje de trabajar o ingrese a éste en un lapso posterior de haber terminado el último grado aprobado de estudio.

Las diferencias en dichos promedios son más notorias en los rangos más bajos de experiencia, donde se observan diferencias de hasta 10 años en el promedio de edad en los tres rangos de experiencia inferiores. Dicha brecha disminuye conforme se avanza a los rangos más elevados, siendo el rango “de más de 35” en donde se observa la menor

diferencia, la cual es de 2.8 años. Es importante tener en cuenta que dichas diferencias en medias son estadísticamente diferentes, corroborando que los promedios al utilizar la experiencia real como medición no acotan mediante la edad la experiencia en el mercado laboral.

Cuadro 4: Promedio de escolaridad según rango de años de experiencia, por año de la encuesta y tipo de experiencia utilizada para generar el rango, 2014.

Años de experiencia	Real	Potencial	Prueba de medias
De 1 a 5	10.0	11.6	Significativa
De 6 a 10	10.1	11.1	Significativa
De 11 a 15	9.9	10.8	Significativa
De 16 a 20	9.9	10.5	Significativa
De 21 a 25	9.9	10.0	No significativa
De 26 a 30	10.5	9.8	Significativa
De 31 a 35	9.2	9.5	No significativa
Más de 35	7.0	6.6	No significativa

Nota: La prueba de medias se realizó para un nivel de confianza de 95%.

Fuente: Cálculos propios con base en la ENOE 2014 (INEGI)

El cuadro 4 presenta los promedios de escolaridad distribuidos por rango de años de experiencia, se observan pequeñas variaciones entre los tipos de medida de experiencia. Al utilizar los rangos generados con la experiencia potencial se percibe que el promedio de escolaridad decrece conforme se incrementan los años de experiencia, desde 11.6 hasta 6.6 años, siendo éste el promedio más bajo en toda la distribución.

Por el contrario, cuando se observa la información de dicho promedio con base en la medida de experiencia real, no existe un patrón como el descrito anteriormente; los extremos para este caso son 7.0 y 10.5 años. En general, el promedio de escolaridad para las diferentes medidas es muy similar, aunque el comparativo demuestra que no necesariamente los promedios de escolaridad más altos se encuentran necesariamente en los rangos más bajos de experiencia, como la experiencia potencial sugiere.

Capítulo 4: Estrategia empírica

El método comúnmente utilizado para calcular los parámetros que ajustan a un modelo de regresión lineal es el de Mínimos Cuadrados Ordinarios (MCO), el cual se basa en la media de la variable dependiente de la regresión. Al utilizar dicho proceso, se asume que las probables diferencias en términos de impacto de las variables exógenas en la distribución condicional son poco significativas, lo cual puede ser un argumento débil para defender su aplicación en segmentos específicos de la distribución.

Koenker y Bassett (1978) propusieron por primera vez un modelo para obtener estimadores en distintos rangos de una distribución, alternativo al estándar tradicional. Los cuantiles en estadística se refieren a aquellos valores de la variable que, en orden ascendente, separan la distribución condicional en ciertas características, segmentando las frecuencias de acuerdo al número de cuantiles que se requieran.

En general, el modelo de regresión cuantílica se basa en un modelo lineal para el τ -ésimo cuantil condicional de Y ; la diferencia básica con el modelo de regresión lineal tradicional es que, para obtener los estimadores, se busca la minimización de la suma de los residuos asimétricamente ponderados en vez de minimizar la suma de cuadrados de los residuos.

En términos formales, la población que conforma el τ -ésimo cuantil de la variable Y (Y_τ) puede verse como sigue: Se tiene una variable aleatoria Y que se caracteriza por tener una función de distribución $F(y) = P(Y \leq y)$. Cuando y es continua, para valores $0 < \tau < 1$, el valor de dicho cuantil es la y que satisface $\tau = P(Y \leq y_\tau) = F(y_\tau)$; entonces $Y_\tau = F^{-1}(\tau) = \inf(y: F(y) \geq \tau)$. Al igual que una función de distribución, cada cuantil proporciona una representación completa de la variable Y .

El modelo de regresión tradicional estima una función lineal de la media condicional $E(Y/X = x) = x'\beta$, mediante la minimización del cuadrado de los errores. Lo anterior con el objetivo de encontrar un estimador de la media no condicionada de Y , es decir $E(Y) = \hat{\mu}$ que resulta del problema de minimización:

$$(5) \hat{\mu} = \min_{\mu \in R} \sum (y_i - \mu)^2$$

Donde y y μ representan una muestra aleatoria y la media de ésta, respectivamente. Dado que el objetivo es la estimación de la función de la media condicional, el problema de minimización es el siguiente:

$$(6) \hat{\beta} = \min_{\mu \in R} \sum (y_i - \mu(x, \beta))^2$$

Dicho razonamiento también aplica para estimar la función lineal de cuantil condicional. Con base en la minimización de la suma de los errores asimétricamente ponderados, de acuerdo al signo, se obtienen los cuantiles:

$$(7) \hat{\xi}(\tau) = \min_{\xi \in R} \sum \varphi_{\tau}(y_i - \xi)$$

Donde ξ es la muestra del cuantil τ (donde $0 < \tau < 1$) y φ_{τ} es una función lineal de los residuos de dicha muestra que solucionan el problema. Replicando el proceso que se siguió anteriormente para la regresión clásica, se obtienen la siguiente expresión para obtener la estimación de las funciones condicionales de cuantiles:

$$(8) \hat{\beta}_{\tau} = \min_{\xi \in R^k} \sum \varphi_{\tau}(y_i - \xi(x'_i \beta))$$

Donde φ_{τ} puede ser τ_{ε} si $\varepsilon \geq 0$ y $(\tau - 1)_{\varepsilon}$ si $\varepsilon < 0$.

Dado lo anterior, se define la τ -ésima regresión cuantílica como la solución al problema:

$$(9) \hat{\beta}_{\tau} = \min_{\beta \in R^k} \left[\sum_{i: y_i \geq x'_i \beta} \beta_i \tau |y_i - x'_i \beta| + \sum_{i: y_i < x'_i \beta} \beta_i (1 - \tau) |y_i - x'_i \beta| \right]$$

Donde $0 < \tau < 1$.

Para obtener los parámetros que minimizan la ecuación anterior se tienen que utilizar métodos de optimización iterativos y/o métodos de programación lineal ya que, la función como tal, no es diferenciable³.

Autores como Buchinsky (1994 y 1998), López Acevedo (2004), Caamal-Olvera (2013), han optado por utilizar la regresión cuantílica (*quantile regression*) como un método alternativo para generar estimadores más precisos, ya que este procedimiento permite analizar cada sección de la distribución y hacer la estimación para segmentos

³ Dado que no es objeto de este estudio desarrollar exhaustivamente dicha técnica se dejan los detalles específicos del desarrollo matemático como lectura complementaria en Koenker y Bassett (1978).

definidos de acuerdo a la necesidad del investigador, sin supuestos forzosos sobre la distribución del término de error, lo que implica que es un método semi-paramétrico.

4.1 Modelo de inferencia para experiencia real.

En la sección 2.1 se establece que, en teoría, el historial laboral de un individuo es conocido, por lo que el tiempo de experiencia en cualquier área puede ser recopilado para utilizarse en el modelo minceriano y, por tanto, generar estimadores insesgados; cuando se utiliza la medida de experiencia potencial se comprobó la existencia de un posible sesgo en los estimadores (Ecuación 3).

Cuando no se cuenta con la información necesaria para generar dicho historial, Oaxaca (2008) propone un modelo semi-logarítmico que genera estimadores de la medida que permitan, junto con las observaciones para las que se tiene certeza, obtener una medida de experiencia laboral más robusta que la tradicionalmente utilizada. La propuesta se basa en un modelo de regresión que toma en cuenta las observaciones en las que podemos obtener la experiencia laboral del individuo; el modelo en términos matriciales es el siguiente:

$$(10) \quad \ln x_i = Z_i \gamma_i + v_i$$

Donde:

- i = Representa cada individuo
- $\ln x_i$ = Logaritmo natural de la experiencia real
- Z_i = Conjunto de variables que identifican las características individuales y laborales de los individuos (género, edad, variables dummy sobre condiciones laborales relevantes, etc.)
- v_i = Término de error que satisface los supuestos estándar sin una distribución en específico

Dado lo anterior, podemos estimar la experiencia laboral al utilizar los resultados que arroja la regresión 10, calculada por MCO, como sigue:

$$(11) \quad \hat{\delta}_1 \hat{x}_i = \hat{\delta}_1 (Z_i \hat{\gamma}_i)$$

Donde:

- $\hat{\gamma}_i$ = Estimadores que se obtienen de la ecuación 2.
- $\hat{\delta}_1$ = Factor de escala con el que se busca que la media estimada sea igual a la media muestral.

El factor de escala se define como:

$$(12) \quad \hat{\delta}_1 = \frac{\sum_i x_i}{\sum_i \exp(Z_i \hat{\gamma}_i)}$$

Originalmente la idea es observar la razón de medias de los valores reales respecto a los estimados, lo que se resume en la ecuación anterior con la razón de las sumatorias, ya que los valores muestrales (n) sobre los que se llevan a cabo dicha suma son exactamente los mismos. El objetivo de ponderar dicha desviación es obtener valores predichos de la experiencia real lo más cercano posible a los verdaderos. Ese proceso de cálculo se desarrolla exclusivamente para complementar la información sobre experiencia que no pudiera ser obtenida directamente con las variables del conjunto de datos.

Con el antecedente que el cálculo por MCO de cualquier tipo de regresión tiene algunos problemas inevitables de sesgo, la idea de seguir esta metodología es generar una base de datos más grande, para estimar con más precisión los parámetros del modelo.

Para este caso, después de algunas pruebas con distintas variables, el modelo para el ajuste y predicción es el siguiente:

$$(13) \quad \ln x_i = \gamma_0 + \gamma_1 edad_i + \gamma_2 a_escola_i + \gamma_3 casado_i + \gamma_4 contra_esc_i + \gamma_5 medica_i + \gamma_7 afore_i + \varepsilon_i$$

Donde:

- edad = Edad en años
- a_escola = Años de escolaridad
- casado = Variable dummy que toma el valor de 1 si la persona está casada o en algún tipo de unión
- contra_esc = Variable dummy que toma el valor de 1 si la persona tiene un contrato por escrito con la institución donde trabaja

- medica= Variable dummy que toma el valor de 1 si la persona tiene algún tipo de servicio médico por parte de la empresa para la que presta sus servicios
- afore= Variable dummy que toma el valor de 1 si la persona tiene afore

En el modelo anterior, las variables elegidas para las estimaciones fueron seleccionadas con base en las características que identifican la estabilidad laboral que un trabajador promedio busca, ya que se pretende hacer inferencia de cuánta experiencia real tendrían si suponemos que hubieran estado en un solo empleo, situación de los individuos para el caso 1.

Asumiendo que el promedio de los años de experiencia laboral de una persona es independiente del cuantil de ingreso en el que ésta se clasifique, la experiencia promedio estimada de la muestra sería una estimación suficiente para dar un panorama general de dicho indicador. En este caso, dado el supuesto anterior, se optó por utilizar el método de Mínimos Cuadrados Ordinarios (MCO) para la estimación del modelo de inferencia.

La combinación de la experiencia real que se infiere del propio cuestionario y de este modelo, generan una variable de experiencia compuesta que en el estudio llamaremos experiencia *construida*. Los resultados de dicho modelo se presentan en la sección de anexos.

Capítulo 5: Estimaciones y resultados.

Como se ha mencionado en el capítulo 2, el modelo logarítmico del salario propuesto por Mincer (1974) es el más utilizado para realizar este tipo de mediciones. El planteamiento de un modelo sencillo, permite entender e identificar con mayor facilidad los beneficios o defectos que puede tener la medida de experiencia propuesta en las distintas variables relacionadas con capital humano utilizadas en la estimación.

Se plantea un modelo semi-logarítmico de inferencia para estimar la cantidad de experiencia real de las personas con historial laboral incompleto dentro de la muestra; con dicho proceso se espera obtener resultados más robustos para el modelo de Mincer

que, en general, pueda ser aplicado independientemente del objetivo de estudio para el que se utilice la ecuación.

Tipos de medidas de experiencia en el modelo.

De acuerdo a las características de la encuesta y dados los casos que se puntualizaron, existen dos tipos de individuo en la base: A los que se puede calcular su historial laboral completo (caso 1) y los que no (caso 2). En el segundo caso se observan personas que sólo tienen parte de su historial laboral, el cual puede calcularse para aquellos que tienen antecedentes laborales.

Utilizando el modelo semi-logarítmico de inferencia propuesto se puede imputar la experiencia laboral, de acuerdo a las características personales y laborales, lo cual permitiría hacer inferencia con una mayor cantidad de observaciones cuando se aplica el modelo de Mincer.

Dado lo anterior, y considerando todos los posibles escenarios, se definen tres medidas de experiencia a comparar:

- 1) ***Experiencia potencial***: Medición estándar utilizada en la literatura, propuesta por Mincer, que acumula los años de experiencia que un individuo tiene en el mercado laboral, asumiendo que ingresa en éste desde su egreso del último nivel de estudios reportado ($Exp = Edad - Años\ de\ escolaridad - 6$)
- 2) ***Experiencia real***: Medición de la experiencia laboral dadas las respuestas a las tres preguntas contenidas en el cuestionario ampliado de la ENOE. De acuerdo al resumen de del cuadro 1, se refiere a los individuos del caso 1, los cuales han permanecido en su empleo desde que iniciaron su trabajo.
- 3) ***Experiencia real construida***: Estimación de la experiencia laboral para los individuos en la encuesta, combinando las observaciones del caso 1 con los resultados del modelo semi-logarítmico de inferencia de la experiencia (ecuación 13).

Para tener un panorama general, el siguiente cuadro muestra un resumen de las distintas medidas de experiencia del conjunto de datos:

Cuadro 5: Estadística descriptiva para variables de experiencia en modelos de Mincer.

	Frecuencia	Total	
		Media	Desviación Estándar
Potencial	44,523	20.7	13.1
Real	25,404	10.3	8.8
Real construida	44,112	8.4	7.3

Fuente: Estimaciones propias con base en la ENOE.

Como se puede notar, la media de la experiencia tradicional para el total de las observaciones resulta relativamente alta respecto a las medidas propuestas. En promedio, la experiencia potencial estimada para la muestra es de 20.7 años en 2014, con una desviación estándar de 13.1 años. La población a la que puede calcularse su historial laboral completo (experiencia real), muestra un promedio de 10.3 años de experiencia, con un rango menor de desviación estándar (8.8 años). La experiencia real construida muestra 8.4 años, con una desviación más ajustada (7.3 años).

El hecho de que, en promedio, se asignen de 10 a 12 años más de experiencia laboral a los individuos cuando se utiliza la medida potencial, implica un sesgo inherente por la elección en el tipo de medición, lo cual se transmite a los resultados de la estimación del modelo de Mincer en todos sus coeficientes.

Es importante señalar que para la estimación del modelo se tomará en cuenta todos los datos que contengan información de la medida de experiencia utilizada en cada ejercicio, por lo que no se verán la misma cantidad de personas en cada modelo; ello ya se refleja en el cuadro descriptivo 5. Incluso, es probable los números de dicho cuadro presenten diferencias respecto a las observaciones utilizadas para generar las regresiones del siguiente tema, esto por la falta de información de una o varias variables explicativas.

Estimación del modelo de Mincer

Con base en el modelo original de Mincer (1974), la propuesta para la ecuación es la siguiente:

$$(14) \quad Y_i = \beta_0 + \beta_1 a_escola_i + \beta_2 x_i + \beta_3 x_i^2 + \beta_4 casado_i + \beta_5 contra_esc_i + \beta_6 región_i + \varepsilon_i$$

Donde⁴:

- Y_i = Logaritmo natural del salario por hora (*lnsalario*)
- a_escola = Años de escolaridad
- x_i = Medida de experiencia potencial, experiencia real o experiencia real construida (*exp_pot*, *exp_real1*, *exp_real*)
- $casado$ = Variable dummy que toma el valor de 1 si la persona está casada o en algún tipo de unión
- $contra_esc$ = Variable dummy que toma el valor de 1 si la persona tiene un contrato por escrito con la institución donde trabaja
- $region$ = Se toman en cuenta las regiones geográficas en las que se divide el país, agregando 5 variables dummy y teniendo como base la región “México” conformada por el Estado de México y el Distrito Federal⁵.

En el cuadro 6 se muestra el resumen de los coeficientes de los rendimientos en las variables relacionadas con capital humano: Educación y experiencia, según cuantil y de acuerdo a la medida de experiencia utilizada para el cálculo.

Autores como Garvey y Reimers (1980), Barceinas (2001), Antecol y Bedard (2004) y López-Acevedo (2004) coinciden en que conforme el individuo acumule más experiencia y/o escolaridad su nivel salarial se incrementa, situación observada en los resultados con estimaciones positivas de los coeficientes relacionados; también, cuando se emplea el cuadrado la experiencia, el efecto esperado para alcanzar un nivel máximo de ingreso, dada la edad y la experiencia del trabajador, se da: Signo negativo en el coeficiente.

⁴ Las variables que no se mencionan en la enumeración de este apartado se definen igual que en el modelo de inferencia propuesto para obtener la experiencia real del individuo en el apartado 4.

⁵ Las seis regiones son: Frontera Norte (Baja California, Coahuila, Chihuahua, Nuevo León, Sonora y Tamaulipas), “Centro-Norte” (Aguascalientes, Durango, Guanajuato, San Luis Potosí y Zacatecas), “Centro” (Hidalgo, Morelos, Puebla, Querétaro y Tlaxcala), “Pacífico” (Baja California Sur, Colima, Chiapas, Guerrero, Jalisco, Michoacán, Nayarit, Oaxaca y Sinaloa), “Golfo” (Campeche, Quintana Roo, Tabasco, Veracruz y Yucatán) y “México” (Estado de México y Distrito Federal).

Cuadro 6: Coeficientes de variables relacionadas al capital humano por cuantil, según medida de experiencia utilizada en el modelo de regresión, 2014.

Cuantil	Educación			Experiencia		
	Potencial	Real	Construida ¹	Potencial	Real	Construida ¹
q10	0.04678 (0.00148)	0.05146 (0.00212)	0.04491 (0.00168)	0.01231 (0.00149)	0.01379 (0.00152)	0.01465 (0.00146)
q25	0.05915 (0.00098)	0.05986 (0.00174)	0.05406 (0.00146)	0.01431 (0.00069)	0.01863 (0.00119)	0.02068 (0.00143)
q50	0.07291 (0.00082)	0.06885 (0.00118)	0.06431 (0.00098)	0.01741 (0.00084)	0.02274 (0.00091)	0.02748 (0.00118)
q75	0.08769 (0.00094)	0.07599 (0.00091)	0.07249 (0.00065)	0.02081 (0.00106)	0.02408 (0.00134)	0.03243 (0.00145)
q90	0.09731 (0.00117)	0.08177 (0.00120)	0.07980 (0.00085)	0.02177 (0.00113)	0.01871 (0.00268)	0.03209 (0.00131)

¹ La experiencia construida es el resultado de la combinación entre los valores de experiencia real y las estimaciones de experiencia realizadas con el modelo del apartado 4.1

Nota: Los coeficientes son significativos al 99% de confianza.

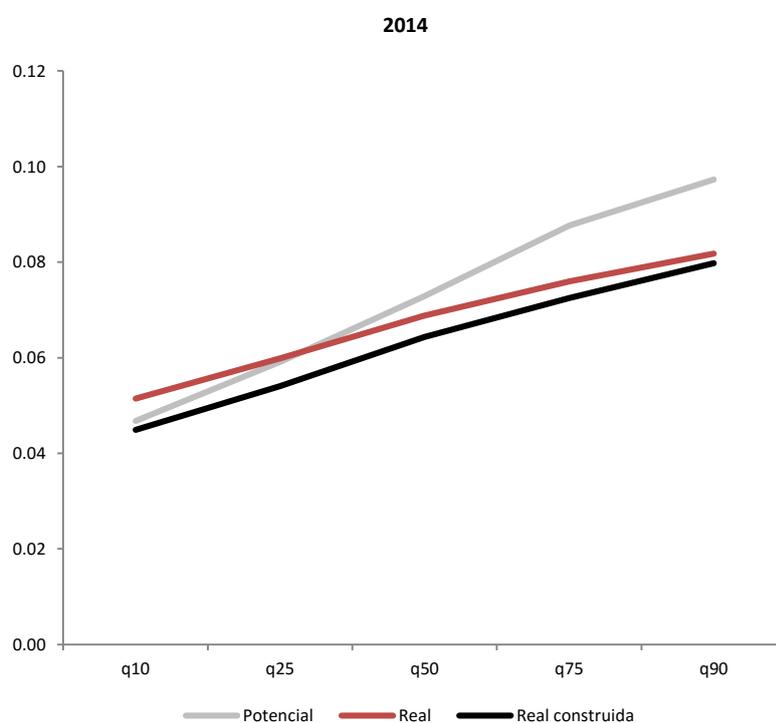
Fuente: Cálculos propios basados en el modelo de regresión y la ENOE.

En el caso de los rendimientos a la educación, el efecto teórico esperado de incremento en el nivel de ingreso por año de escolaridad adicional, es acompañado de los resultados empíricos, resultando una tendencia de incremento porcentual conforme el cuantil de ingreso es más elevado. Los cuantiles q10 y q25 son los que muestran valores menores en este ejercicio. Para el cuantil q10 se observa un incremento en los rendimientos de la escolaridad respecto al escenario base al utilizar experiencia real, pasando de 4.68% a 5.15%, mientras que utilizando experiencia real construida el valor del coeficiente es más bajo que el escenario base, llegando hasta 4.50%.

Mientras tanto, en el escenario base de q25 se observa un valor de 5.915%; similar a los resultados del cuantil más bajo, al utilizar las medidas alternativas de experiencia propuestas se presentan resultados similares respecto al escenario base. Por ejemplo, el coeficiente al utilizar experiencia real se ubica en 5.986%, lo cual implica que en promedio en dicho cuantil el rendimiento es 1.2% más elevado. Dicho incremento se revierte al utilizar experiencia real en los cálculos, obteniendo un valor de 5.406%, lo cual es menor en 8.6%.

A partir del cuantil que representa a la mediana de la distribución, q50, se muestra una tendencia clara, ya que para ambos años las medidas de experiencia propuestas generan valores más pequeños que el escenario base planteado, situación particularmente visible en los cuantiles más altos. En la gráfica 2 se muestra un ejercicio para visualizar dicha discrepancia, observando que el coeficiente de la educación en el modelo efectivamente incrementa su brecha conforme los cuantiles son más elevados, teniendo su máximo en el cuantil más alto considerado para los cálculos.

Gráfica 2: Rendimientos educativos por año y cuantil, según medida de experiencia utilizada en el modelo de regresión.



Fuente: Cálculos propios, con base en ENOE 2014. INEGI.

Es importante resaltar que, para este caso, los resultados al utilizar experiencia real y construida, tienen un comportamiento similar entre cuantiles, siendo la segunda ligeramente menor en todos los casos y para ambos años; en conjunto, al utilizar cualquiera de los dos tipos de medida propuestas, se puede admitir que el concepto de experiencia elegido para el modelo sí tiene implicaciones importantes para el componente de escolaridad medido en años aprobados para los cuantiles por encima de q50. Este ejercicio se realizó también utilizando la muestra para 2013, la tendencia fue similar a la observada en 2014, incluso siendo marginalmente un poco menores a los de 2014.

Comparando los resultados con ejercicios similares para México, los coeficientes relacionados al rendimiento educativo sufren cambios notables, sobre todo en los cuantiles más bajos. Por ejemplo, respecto al último año de información en Caamal-Olvera (2013) utilizó un polinomio en la variable de edad, y al comparar con esta metodología se encuentran diferencias al utilizar experiencia real, ya que los resultados para los cuantiles q10, q25 y q50 resultan ser más altos. El cuantil más elevado (q90) en el presente modelo presenta valores más bajos, mientras que en q75 no se encuentra una tendencia clara al respecto. A su vez, los resultados en la mediana de nuestro modelo son más bajos que los obtenidos por Rodríguez-Oreggia (2005) al utilizar el modelo de Mincer básico y también cuando corrige dicho modelo con el método de Heckman.

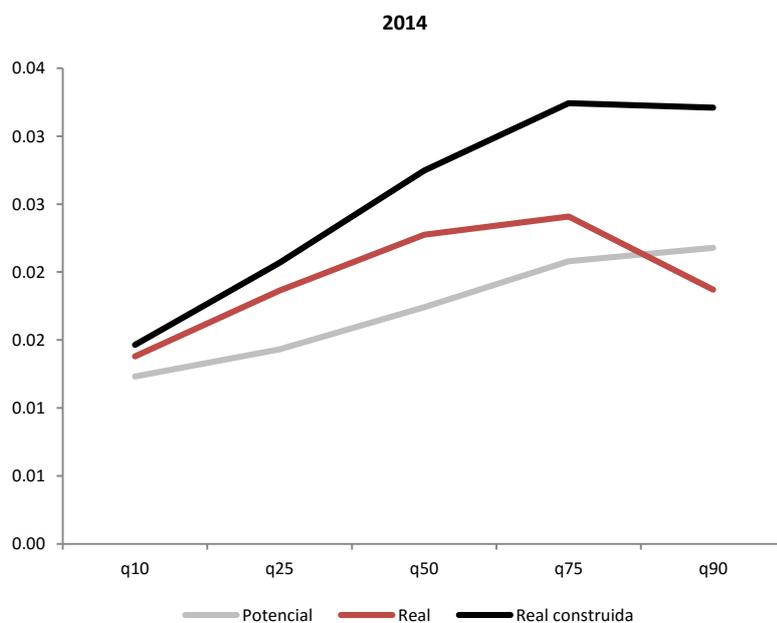
Al utilizar la experiencia real construida los resultados guardan una relación similar a la anterior: Los cuantiles q10, q25 y q50 resultan ser más altos, mientras q75 y el q90 son más bajos respecto a Caamal- Olvera (2013).

El segundo componente respecto a capital humano en el modelo es el coeficiente de la experiencia, cuyos resultados se muestran en la gráfica 3. Como se observa, al utilizar experiencia potencial los valores van desde 1.23% a 2.18% para 2014, mostrando que el rendimiento por año de experiencia adicional crece conforme un individuo se encuentra en un cuantil más alto. Para los casos donde se utiliza la experiencia real o construida dicha tendencia es relativamente parecida entre cuantiles, en la mayoría de los casos con valores más altos respecto al estándar de medición de una regresión minceriana (excepto en el cuantil más alto).

Los valores obtenidos al utilizar la experiencia denominada real son claramente más elevados que el escenario base para los primeros cuatro cuantiles del ejercicio, mostrando una tendencia a la baja en el último cuantil, donde el 1.87% reflejado es inferior al 2.18% del escenario base. En la gráfica 3, se observa dicho comportamiento, lo cual implica que el rendimiento por año adicional de experiencia tomando en cuenta la medición real de esta es más bajo en los cuantiles más altos comparado con la medición tradicional, sobreestimándola, y más alto en los primeros cuatro cuantiles, subestimando la medida.

En el caso de la experiencia real construida, la tendencia creciente entre cuantiles es muy similar a la medida potencial, con cierta disminución en los cuantiles más altos de la distribución (similar a la experiencia real); en cada año se observa que los rendimientos para esta son subestimados al utilizar la *proxy* de experiencia clásica Minceriana. Es de resaltar que, en 2014, los cuantiles Q75 y Q90 exhiben las mayores brechas entre ambas medidas, con diferencias en puntos porcentuales de 1.16% y 1.03%, respectivamente. Para 2013, el comparativo entre experiencia potencial y real construida mantiene una brecha menor que en el 2014, mientras que el cálculo con experiencia real muestra un comportamiento similar al escenario base de ese año en los cuantiles más altos.

Gráfica 3: Rendimientos de experiencia por cuantil, según medida de experiencia utilizada en el modelo de regresión.



Fuente: Cálculos propios, con base en ENOE 2014. INEGI.

El cuadro 7 presenta un resumen de las pseudo R cuadrada⁶ para cada cuantil, de acuerdo al tipo de experiencia utilizada. Los valores observados, por definición técnica, no pueden ser interpretados como porcentajes de ajuste de las variables independientes, lo cual implica que, en términos exactos, no se puede saber qué tanto más (o menos) es

⁶ $pseudoR^2 = 1 - \frac{\ln \hat{L}(M_{completo})}{\ln \hat{L}(M_{intercepto})}$, donde $\ln \hat{L}(M_{completo})$ =Modelo con predictores, $\ln \hat{L}(M_{intercepto})$ =Modelo sin predictores y \hat{L} = Verosimilitud estimada. Dicho valor puede estar entre 0 y 1. Entre más cercano esté a 1, mejor será el ajuste del modelo y viceversa (No tiene interpretación porcentual).

mayor (o menor) el ajuste entre regresiones, sin embargo sí se puede comparar el ajuste entre las ecuaciones propuestas.

Comparando los resultados del modelo que utiliza experiencia potencial contra el de experiencia real, se observa un ajuste más robusto en esta última medida para todos los cuantiles, lo cual es consistente con el planteamiento teórico del capítulo 2. Dichas diferencias de ajuste, en promedio, son similares entre cuantiles.

Cuadro 7: México. Pseudo R cuadrada por año y cuantil, según medida de experiencia utilizada en el modelo de regresión

Año	Cuantil	Tipo de experiencia		
		Potencial	Real	Construida
2014	MCO	0.2895	0.3356	0.2908
	q10	0.0970	0.1184	0.0984
	q25	0.1203	0.1489	0.1221
	q50	0.1596	0.1914	0.1607
	q75	0.2083	0.2366	0.2063
	q90	0.2256	0.2463	0.2204

Fuente: Cálculos propios basados en el modelo de regresión y la ENOE de cada año.

Al contrastar el modelo base con el que utiliza experiencia construida los resultados varían entre cuantiles: En el año 2014, se observa una mejora marginal en los tres cuantiles más bajos, lo cual indica cierta reserva y ambigüedad en la interpretación que puede darse a los resultados, ya que no hay claramente una estimación mayor o menor. En los cuantiles más altos, el ajuste es favorable para el modelo base (experiencia potencial) y en el que se utiliza la experiencia real. En el año 2013, los cuantiles q10 y q25 muestran una ligera mejoría en el ajuste al utilizar experiencia construida frente al modelo con experiencia potencial, situación que se revierte para los tres segmentos restantes de la distribución de ingresos.

Por último, las medidas de experiencia propuestas en este estudio al ser comparadas arrojan el resultado esperado: Utilizar datos considerados puntualmente reales en el modelo para la experiencia genera mejores resultados que, incluso, utilizar medidas alternativas con un posible mejor ajuste. Al comparar cada cuantil, la respectiva bondad

de ajuste en los modelos, se observan diferencias claras para aseverar que los mejores resultados están relacionados con el uso de experiencia real en el modelo, lo cual también es consistente con la teoría presentada sobre el sesgo inherente al utilizar *proxys* de experiencia en el modelo minceriano.

Capítulo 6: Conclusiones

El creciente interés en los estudios relacionados a la teoría del capital humano y los avances generados los últimos años en materia de legislación educativa en el país, enfatizan la importancia de los resultados a obtener en los distintos ejercicios relacionados a las variables de capital humano representado en cualquier modelo clásico o actual propuesto por el (los) investigador (es) en turno.

Una de las variables más cuestionadas históricamente en dichos modelos, específicamente el modelo Minceriano de ganancias por capital humano, es la experiencia potencial, cuyo uso genera un sesgo implícito en los resultados obtenidos. En el presente trabajo se proponen 2 formas alternativas para capturar la experiencia laboral real de una persona para México, acorde a la información disponible en las fuentes oficiales.

Utilizando la base de datos de la ENOE 2014 en el trimestre donde se aplica el cuestionario “ampliado”, se encuentra que para una gran cantidad de individuos puede rescatarse perfiles de experiencia completos. Con base en ello, se propusieron las alternativas de experiencia real y experiencia real construida, buscando solucionar los problemas mencionados.

Utilizando el modelo de Mincer y la técnica de regresión cuantílica para las estimaciones, los resultados muestran que la elección del tipo de medida de experiencia a utilizar en el modelo tiene implicaciones importantes en los resultados, ya que, estadísticamente, hay diferencias importantes, sobretodo en el coeficiente que mide el rendimiento a la educación, que sobreestima dichos valores en su modelo base (al utilizar experiencia potencial), observando diferencias en puntos porcentuales de entre 1% y 2% para los cuantiles más altos.

En el caso de la experiencia se observa que, utilizando una mejor medida de ésta, el sesgo puede reducirse, teniendo que, con base en la experiencia real, hay una subestimación del aporte a los rendimientos respecto al modelo base en los cuantiles más bajos y una sobreestimación en los cuantiles altos. Al utilizar la experiencia real

construida se observa una subestimación marcada para todos los cuantiles en ambos años.

Finalmente, de acuerdo a las pruebas realizadas a los resultados, se sugiere que, con dichas medidas, en especial la medida de experiencia real, el sesgo inherente al utilizar experiencia potencial se reduce, lo cual representa una alternativa importante para sustituir la experiencia potencial; No obstante lo anterior, es importante destacar que, aunque no se ataca directamente en este estudio el sesgo por habilidad generado en este tipo de mediciones, el hecho de utilizar la regresión cuantílica lo modera de alguna manera, separando en la distribución los más hábiles de los menos hábiles.

Bibliografía

Angrist, J. y A. Krueger (1991), “Does compulsory school attendance affect schooling and earnings?”, *The Quarterly Journal of Economics*, vol. 106, No. 4, pp. 979-1014.

Antecol, H. y Bedard, K. (2002) “The relative earnings of young mexican, black, and white women”. *Industrial and Labor Relations Review*, 56(1), pp. 122–135

Antecol, H. y Bedard, K. (2004) “The racial wage gap: the importance of labor force attachment differences across black, mexican, and white men” *Journal of Human Resources*, 39(2), pp. 564–583

Ashenfelter, O. y Krueger, A. (1994), “Estimates of the economic return to schooling”, *The American Economic Review*, vol. 84, No. 5, pp. 1157-1173.

Barceinas, F. (2001), “Capital humano y rendimientos de la educación en México”, Tesis Doctoral, Universidad Autónoma de Barcelona

Becker, G. (1993), “Human Capital, 3rd edition”, *Chicago: University of Chicago Press*.

Behrman et. Al. (1980), “Socioeconomic success: A study of the effect of genetic endowments, family environment and schooling (Contributions to economic analysis)”.

Bracho, T. y A. Zamudio (1994), “Los rendimientos económicos de la escolaridad en México, 1989”, *Economía Mexicana. Nueva Época*, vol. III, N° 2, pp. 345-377.

Bratsberg, B y Terrel, D. (1998), “Experience, tenure, and wage growth of young black and white men”, *The Journal of Human Resources*, vol. 33, No. 3, pp. 658-682

Caamal-Olvera, C.G. (2013). “Effects of manufacturing dynamics on returns to education in the North of Mexico,” *Políticas Públicas*, Vol. 1, No. 1, pp. 43-58.

Caamal-Olvera, C.G. (2014). “Gender Differences in the Returns to Education in Mexico,” *Journal of Business and Policy Research*, Vol. 9. No. 1, pp. 65 – 81.

Carneiro, P., J. Heckman y E. Vytlačil (2001), "Estimating the returns to education when it varies among individuals", Working Paper, *University of Chicago*.

Carneiro, P. y J. Heckman (2002), "The evidence on credit constraints in post-secondary schooling", *Economic Journal*, vol. 112, No. 482, pp. 705-734.

Carnoy, M. (1967), "Rates of returns to schooling in Latin America", *The Journal of Human Resources*, vol. 2, N° 3, pp. 359-374.

Card, D. y Krueger, A. (1992) "Does school quality matter? Returns to education and the characteristics of public schools in the United States", *Journal of Political Economy*, 100, February 1992.

Card, D. (1995) "Earnings, schooling and ability revisited", in: S. Polachek, ed. *Research in Labor Economics*, vol. 14, pp. 23-48

Card, D. y A. Krueger (1996), "School resources and student outcomes: An overview of the literature and new evidence from North and South Carolina", *Journal of Economic Perspectives*, vol. 10, pp. 31-50.

Card, D. (1999), "The causal effect of education on earnings", chapter in the *Handbook of Labor Economics*, Volume 3A.

Dougherty, C. (2005), "Why Are the Returns to Schooling Higher for Women than for Men?", *Journal of Human Resource*. Vol. 40, No. 4 (Autumn, 2005), pp. 969-988

Filer, R.K. (1993) "The usefulness of predicted values for prior work experience in analyzing labor market outcomes for women", *Journal of Human Resources* 28(3), pp. 519-537

Garvey, N. y Reimers, C. (1980) "Predicted vs. potential work experience in an earnings function for young women", *Research in labor economics* 3. JAI, Greenwich, pp. 99-127.

Griliches, Z. y Manson, W. (1972), "Education, income and ability", *Journal of Political Economy*, 80[3], May-June 1972 Part 2, S74-S103.

Griliches, Z. (1977), "Estimating the returns to schooling: Some econometric problems", *Econometrica*, vol. 45, pp. 1-22.

Hashimoto, M. y Raisian, J. (1985) "Employment tenure and earnings profiles in Japan and the United States", *American Economic Review*, vol. 75, No. 4, pp. 721-735.

Heckman, J. y Xuesong L. (2004), "Selection bias, comparative advantage and heterogeneous returns to education: Evidence from China in 2000", *Pacific Economic Review*, vol. 9, No. 3, pp. 155-171.

Kane, T.J. y Rouse, C.E. (1993), Labor market returns to two- and four- year colleges: is a credit and do degrees matter?, working paper no. 4268 (N.B.E.R. Cambridge, Massachusetts).

López-Acevedo, G. (2004), "Mexico: Evolution of earnings inequality and rates of returns to education (1988-2002)", *Estudios Económicos*, vol. 19, No. 2, pp. 211-284.

Mincer, J. (1974), "Schooling experience and earnings", *National Bureau of Economic Research*, Columbia University Press.

Moulton, B.R. (1986), "An analysis of female work experience data derived from social security records", *Journal of Economic and Social Measurement*, 14(1), pp. 66-75

Oaxaca, R. y Regan, T. (2009) "Work experience as a source of specification error in earnings models: implications for gender wage decompositions", *Journal of Population Economics*, 22, pp. 463-499

Oreopoulos, P. (2006) "Estimating average and local average treatment effects of education when compulsory schooling laws really matter", *American Economic Review*, 96(1), pp. 152-175.

Rodriguez-Oreggia, E., "Institutions, Geography and the Regional Development of Returns to Schooling in Mexico," Reporte técnico, Instituto de Investigaciones sobre Desarrollo Sustentable y Equidad Social, Universidad Iberoamericana 2005.

Spence, A.M. (1973) "Job market signaling", *Quarterly Journal of Economics*, Vol. 87, pp. 355-374.

Stevens, M. (2003), "Earnings functions, specific human capital, and job matching: tenure bias is negative", *Journal of Labor Economics*, vol. 21, No. 4, pp. 783-805

Straiger, D. y Stock, J.H., "Instrumental variables regression with weak instruments", *Econometrica*, vol. 65, pp. 557-586

Wellington, A. (1993), "Changes in the male/female wage gap, 1976-85", *The Journal of Human Resources*, vol. 28, No. 2, pp. 383-411.

Willis, R. y Rosen, S. (1979) "Education and self selection", *Journal of Political Economics*, 87[5] Part 2, Oct 1979: S7-S36.

Gujarati, Damodar: *Econometría*, Mc Graw Hill, 1992.

Maddala, G. S.: *Introduction to Econometrics*, Macmillan, 2a ed., New York, 1992.

Kmenta, Jan: *Elements of Econometrics*, 2a. ed. , Macmillan, New York, 1986.

www.inegi.org.mx

Anexo 1: Resultados de la regresión para generar la medida de experiencia construida.

Cuadro A1.1: México. Estadística descriptiva para variables del modelo de inferencia para experiencia, según año.

2014		
N= 25,007		
Variable	Promedio	Desviación Estándar
ln_x	1.98	0.85
eda	36.78	12.10
a_escola	9.98	3.98
casado	0.71	0.45
contra_esc	0.57	0.49
atnmedica	0.61	0.49
afore	0.54	0.50

Fuente: Estimaciones propias con base en la ENOE para el año correspondiente.

Cuadro A1.2: México. Resultados del modelo de inferencia para experiencia real, según año.

ln_x	2014
eda	0.0366 (0.0004)
a_escola	0.0053 (0.0013)
casado	0.1758 (0.0105)
contra_esc	-0.0872 (0.0167)
atnmedica	-0.0320 (0.0196)
afore	0.0953 (0.0165)
_cons	0.4771 (0.0182)
R ²	0.3124
F(6, 26110)	1,852.54
p-value	0.00
N	25,007

Fuente: Cálculos propios basados en el modelo de regresión y la ENOE para el año correspondiente.

Anexo 2: Resultados de regresiones cuantílicas según tipo de experiencia utilizada.

Cuadro A2.1: Resultados regresión cuantílica utilizando experiencia potencial, 2014						
ln_x	MCO	Q10	Q25	Q50	Q75	Q90
a_escola	0.0771 (0.0009)	0.0468 (0.0021)	0.0592 (0.0013)	0.0729 (0.0014)	0.0877 (0.0014)	0.0973 (0.0017)
exp_pot	0.0168 (0.0007)	0.0123 (0.0015)	0.0143 (0.0008)	0.0174 (0.0007)	0.0208 (0.0009)	0.0218 (0.0014)
exp_pot2	-0.0002 (0.0000)	-0.0002 (0.0000)	-0.0002 (0.0000)	-0.0002 (0.0000)	-0.0002 (0.0000)	-0.0002 (0.0000)
casado	0.0447 (0.0064)	0.0591 (0.0131)	0.0570 (0.0077)	0.0505 (0.0065)	0.0354 (0.0085)	0.0189 (0.0098)
medica5c	0.0340 (0.0031)	0.0418 (0.0051)	0.0427 (0.0049)	0.0342 (0.0036)	0.0291 (0.0037)	0.0179 (0.0047)
contra_esc	0.1577 (0.0067)	0.1719 (0.0092)	0.1502 (0.0068)	0.1511 (0.0074)	0.1399 (0.0105)	0.1416 (0.0118)
region_fn	0.0838 (0.0117)	0.0434 (0.0154)	0.0430 (0.0097)	0.0540 (0.0089)	0.1118 (0.0165)	0.1743 (0.0221)
region_cn	-0.0111 (0.0120)	-0.0575 (0.0170)	-0.0447 (0.0097)	-0.0264 (0.0082)	0.0227 (0.0144)	0.0603 (0.0140)
region_c	-0.0368 (0.0122)	-0.0771 (0.0122)	-0.0677 (0.0097)	-0.0476 (0.0113)	0.0017 (0.0163)	0.0323 (0.0155)
region_p	0.0596 (0.0113)	-0.0695 (0.0170)	-0.0018 (0.0119)	0.0659 (0.0085)	0.1356 (0.0128)	0.1792 (0.0114)
region_g	-0.0117 (0.0120)	-0.1112 (0.0177)	-0.0716 (0.0102)	-0.0270 (0.0108)	0.0629 (0.0142)	0.1160 (0.0152)
_cons	2.0458 (0.0157)	1.8681 (0.0334)	1.9642 (0.0180)	2.0703 (0.0190)	2.1647 (0.0210)	2.3681 (0.0228)
R ² (Pseudo R ²)	0.2895	0.0970	0.1203	0.1596	0.2083	0.2256

Fuente: Cálculos propios basados en el modelo de regresión y la ENOE para el año correspondiente.

Cuadro A2.2: Resultados regresión cuantílica utilizando experiencia real, 2014						
ln_x	MCO	Q10	Q25	Q50	Q75	Q90
a_escola	0.0711 (0.0010)	0.0515 (0.0021)	0.0599 (0.0017)	0.0689 (0.0012)	0.0760 (0.0009)	0.0818 (0.0012)
exp_real1	0.0202 (0.0011)	0.0138 (0.0015)	0.0186 (0.0012)	0.0227 (0.0009)	0.0241 (0.0013)	0.0187 (0.0027)
exp_real2	-0.0002 (0.0000)	-0.0002 (0.0000)	-0.0003 (0.0000)	-0.0003 (0.0000)	-0.0003 (0.0000)	-0.0001 (0.0001)
casado	0.1037 (0.0079)	0.0874 (0.0141)	0.1024 (0.0108)	0.1090 (0.0083)	0.1018 (0.0111)	0.1072 (0.0149)
medica5c	0.0417 (0.0042)	0.0468 (0.0057)	0.0482 (0.0046)	0.0374 (0.0040)	0.0342 (0.0069)	0.0206 (0.0068)
contra_esc	0.1826 (0.0090)	0.1869 (0.0145)	0.1775 (0.0098)	0.1783 (0.0084)	0.1787 (0.0118)	0.1902 (0.0173)
region_fn	0.1063 (0.0138)	0.0285 (0.0177)	0.0493 (0.0184)	0.0723 (0.0122)	0.1471 (0.0192)	0.2152 (0.0265)
region_cn	-0.0177 (0.0140)	-0.0733 (0.0224)	-0.0563 (0.0185)	-0.0419 (0.0106)	0.0253 (0.0145)	0.0700 (0.0283)
region_c	-0.0239 (0.0146)	-0.1081 (0.0192)	-0.0670 (0.0144)	-0.0406 (0.0136)	0.0251 (0.0213)	0.0684 (0.0263)
region_p	0.0544 (0.0129)	-0.0954 (0.0190)	-0.0135 (0.0170)	0.0485 (0.0113)	0.1395 (0.0155)	0.1981 (0.0275)
region_g	-0.0183 (0.0147)	-0.1542 (0.0192)	-0.1001 (0.0173)	-0.0355 (0.0120)	0.0740 (0.0168)	0.1314 (0.0271)
_cons	2.1446 (0.0174)	1.8703 (0.0318)	1.9879 (0.0280)	2.1602 (0.0186)	2.3422 (0.0156)	2.6193 (0.0320)
R ² (Pseudo R ²)	0.3356	0.1184	0.1489	0.1914	0.2366	0.2463

Fuente: Cálculos propios basados en el modelo de regresión y la ENOE para el año correspondiente.

Cuadro A2.3: Resultados regresión cuantílica utilizando experiencia real construida, 2014						
ln_x	MCO	Q10	Q25	Q50	Q75	Q90
a_escola	0.0667 (0.0008)	0.0449 (0.0017)	0.0541 (0.0015)	0.0643 (0.0010)	0.0725 (0.0007)	0.0798 (0.0009)
exp_pot	0.0256 (0.0010)	0.0147 (0.0015)	0.0207 (0.0014)	0.0275 (0.0012)	0.0324 (0.0014)	0.0321 (0.0013)
exp_pot2	-0.0004 (0.0000)	-0.0002 (0.0000)	-0.0003 (0.0000)	-0.0004 (0.0000)	-0.0005 (0.0001)	-0.0004 (0.0000)
casado	0.0849 (0.0061)	0.0821 (0.0099)	0.0889 (0.0060)	0.0894 (0.0071)	0.0865 (0.0064)	0.0734 (0.0115)
medica5c	0.0417 (0.0032)	0.0447 (0.0053)	0.0479 (0.0030)	0.0395 (0.0024)	0.0336 (0.0039)	0.0215 (0.0061)
contra_esc	0.1668 (0.0067)	0.1814 (0.0097)	0.1561 (0.0061)	0.1594 (0.0059)	0.1614 (0.0082)	0.1547 (0.0104)
region_fn	0.0980 (0.0117)	0.0508 (0.0156)	0.0477 (0.0106)	0.0718 (0.0142)	0.1237 (0.0140)	0.2038 (0.0206)
region_cn	-0.0109 (0.0120)	-0.0599 (0.0196)	-0.0508 (0.0101)	-0.0224 (0.0107)	0.0269 (0.0178)	0.0729 (0.0225)
region_c	-0.0347 (0.0122)	-0.0844 (0.0131)	-0.0674 (0.0117)	-0.0421 (0.0130)	-0.0039 (0.0143)	0.0541 (0.0194)
region_p	0.0648 (0.0113)	-0.0609 (0.0152)	-0.0026 (0.0114)	0.0673 (0.0121)	0.1411 (0.0145)	0.2071 (0.0203)
region_g	-0.0114 (0.0120)	-0.1132 (0.0148)	-0.0825 (0.0103)	-0.0219 (0.0152)	0.0585 (0.0161)	0.1303 (0.0219)
_cons	2.1716 (0.0144)	1.9017 (0.0242)	2.0347 (0.0218)	2.1774 (0.0132)	2.3587 (0.0162)	2.6076 (0.0244)
R ² (Pseudo R ²)	0.2908	0.0984	0.1221	0.1607	0.2063	0.2204

Fuente: Cálculos propios basados en el modelo de regresión y la ENOE para el año correspondiente.