

## 4. Los Datos

La muestra que se utiliza para el análisis es la *Encuesta Nacional de Empleo Urbano* para el año 2004 (ENEU 2004). Dicha encuesta se levanta trimestralmente y de manera continúa a partir del año 2000 por el *Instituto Nacional de Estadística Geografía e Informática* (INEGI). La ventaja de utilizar la información estadística contenida en esta encuesta es que ofrece datos sobre la estructura ocupacional (como por ejemplo, ingreso, horas trabajadas, afiliación sindical) relacionado con las características sociodemográficas (edad, género, educación, número de miembros en el hogar).

Además, los datos provienen de localidades o zonas con 100 000 o más habitantes por lo que el vínculo hacia el mercado laboral es estrecho. Asimismo, debido a que se busca encontrar evidencia empírica del impacto de la extensión de los años de escolaridad sobre los ingresos, es necesario tomar en consideración lugares donde posiblemente haya ocurrido algún tipo de ajuste por la descentralización del sistema educativo, pues de no ser así, se estarían mitigando efectos por rezagos en la implementación de tal cambio institucional.

La encuesta tiene un tamaño de 198 353 individuos o 105 584 hogares. Sin embargo, ésta disminuyó a 9802 observaciones pues los individuos tenían que ser i) “hijos”<sup>1</sup> ii) entre las edades de 15 a 65 años de edad, iii) que hayan reportado ingresos positivos, iv) que hayan recibido un salario a cambio del trabajo y v) que hayan trabajado más de cero horas en la semana. Asimismo, a nivel características individuales vi) se tomaron personas que ya no asistieran a la escuela, y vii) que pertenecen al género masculino, para evitar casos de discriminación por género.<sup>2</sup>

---

<sup>1</sup> Para la construcción de la muestra se consideró como hijos a todos aquellos individuos hijos del jefe de hogar o del conyugue del jefe. Es decir, que tuvieran al mismo tiempo un padre. De no haber padre, la construcción de la relación familiar se vuelve relativa pues si el jefe de hogar es hombre y no reporta hijos, entonces será hijo siempre y cuando exista en el hogar “ascendiente del jefe”, por ejemplo.

<sup>2</sup> El mínimo de 15 años de edad garantiza que los individuos hayan terminado al menos el nivel de escolaridad mínima (secundaria) para así poder a comenzar a trabajar. Implícitamente se requiere que la gente en la parte baja de la distribución de edad no trabaje y estudie al mismo tiempo, pues se estaría generando un tipo de sesgo en la estimación. De

## 4.1 Las variables

Como fue mencionado en el capítulo anterior, las variables a utilizar son el logaritmo natural del ingreso ( $\ln Y$ ), la educación ( $S$ ), reforma o tratamiento ( $D$ ), habilidad ( $H$ ), restricción de liquidez ( $L$ ), un vector de características laborales ( $X$ ) y una serie de interacciones entre estas variables que permitirán captar la prima salarial existente por poseer uno o más atributos. Por ello son importantes las variables dicotómicas  $D$ ,  $H$  y  $L$  pues son controles clave para probar las hipótesis planteadas.

Para la construcción de la primera variable se tomó el logaritmo natural del ingreso ( $\ln Y$ ) mensual corriente del individuo, como en la función de ingresos minceriana<sup>3</sup>. La educación ( $S$ ), por su parte, está en años continuos, de tal forma que 6 años de escolaridad representa el nivel primaria completa, 9 secundaria, 12 preparatoria, 17 profesional, 20 maestría y 26 doctorado completos.

Debido a la falta de información estadística en relación a quién participó en la reforma o no, la variable reforma o tratamiento ( $D$ ) es una variable proxy que identifica por un corte de edad si los individuos fueron tratados o no. Esta variable toma el valor de 1 si el individuo estaba entre el rango de edad 15 – 24 en el año de la reforma (1992) tomando como base el año de la encuesta.

Para el caso de la variable “restricción de liquidez” ( $L$ ), se construyeron 4 variables diferentes. Las dos primeras variables se basan en si el individuo tiene un papa cuyo ingreso mensual por número de miembros en el hogar está por encima de 1 línea de pobreza (primera variable) o por encima de 2 líneas de pobreza (segunda variable). La construcción de las

---

la misma forma se necesita que la gente obtenga un ingreso por su trabajo, pues de no ser así se estarían subestimando los ingresos o la primas salariales. Por ello, el máximo de edad considerada es de 65 años.

<sup>3</sup> El utilizar el logaritmo natural del ingreso más que al ingreso como variable dependiente sirve para obtener un coeficiente que represente la tasa de crecimiento del ingreso por un cambio en una variable independiente y además el comportamiento de la distribución de ingreso se estandariza y tiende a la normalidad pues cambios relativamente grandes en el ingreso son cambios no tan grandes en esta forma.

líneas de pobreza se hizo mediante la toma del promedio del tipo de cambio interbancario diario durante el período de levantamiento de la ENEU 2004.<sup>4</sup>

La tercera variable toma el valor de 1 si el individuo tiene un papa con un ingreso por encima de la mediana<sup>5</sup> de ingreso de los papas y por último, una cuarta variable toma como liquidez aquellos individuos con papas con un ingreso superior al promedio de ingreso de los padres.

Por su parte, la variable “habilidad” ( $H$ ) es también una variable dicotómica proxy. Al igual que la variable restricción de liquidez, se construyeron 4 variables de ésta para observar las variaciones causadas sobre el ingreso por el cambio en la habilidad. La primer variable toma el valor de 1 si el individuo tiene papa con educación de preparatoria y más. La segunda variable toma como hábil a individuos con papa con educación mayor a preparatoria. La tercera, que tanto el papa, como mama tengan niveles en promedio mayores a preparatoria. La cuarta variable tomó a la educación promedio de la familia, también superior a preparatoria, como hábil. Lo mismo sucede con la quinta variable (educación mama mayor a preparatoria) y sexta variable (educación de hermana mayor a preparatoria).

En relación al vector de características  $X$ , se incluyeron variables como el logaritmo natural de horas trabajadas al mes, experiencia laboral y su cuadrado. Estas variables están relacionadas con el mercado laboral y sirven para dar un mejor perfil a los ingresos. Para la construcción de la primera variable se tomó el logaritmo natural de las horas trabajadas al mes. En cuanto a la segunda, a falta de una variable que cuantifique la experiencia laboral, se generó una variable que toma los años de vida del individuo menos los años de escolaridad menos seis.

En cuanto a la primera interacción,  $SD$ , representa el nivel de escolaridad de un individuo que participó en la reforma, pues esta variable presenta años continuos de escolaridad siempre y cuando  $D$  sea igual a 1. Algo similar sucede con la variable  $SL$  y  $SH$ , pues presentan años continuos de escolaridad si el individuo tiene una restricción de liquidez baja ( $L=1$ ) o si el individuo es hábil ( $H=1$ ).

---

<sup>4</sup> De acuerdo con la descripción de la *Encuesta Nacional de Empleo Urbano* 2004, el período de levantamiento es de abril a junio. Dado que el tipo de cambio promedio en ese período fue de 11.3984 pesos por dólar y debido a que el ingreso es mensual, la línea de pobreza (1 dólar diario) fue multiplicada por 30 días.

<sup>5</sup> La media es una medida que está en función de la distribución de los datos, mientras que la mediana depende únicamente del número que se encuentra en la mitad del total de observaciones

Cuando la interacción es entre variables dicotómicas como *DL* o *DH* simplemente sirven como indicador de que un individuo participó en la reforma y además son líquidos o hábiles. Estas interacciones son clave si se quiere captar el impacto de la liquidez o habilidad sobre el cambio en el ingreso por la reforma o tratamiento. Una tercer interacción especificada en el modelo es *LH*. Esta variable es 1 si el individuo tiene una restricción de liquidez baja y habilidad alta, simultáneamente. Esta variable es indicador de ese subgrupo.

Las interacciones entre tres variables como *SDL* *SDH* *SLH* denotan los años de escolaridad continuos cuando los individuos participaron en la reforma y además tienen una restricción de liquidez baja (*SDL*) o en el caso de la variable *SDH*, los años de escolaridad de los individuos que participaron en la reforma y son hábiles. Por su parte, la tercera interacción son los años de educación, también de los individuos hábiles con una restricción de liquidez baja (*SLH*). Una última interacción es entre tres variables discontinuas como la variable *DLH* que toma el valor de 1 si el individuo tiene una restricción de liquidez baja, es hábil y además participo en la reforma. La última variable *SDLH* representa la escolaridad para aquellos que cumplen con las características  $D=1, L=1, H=1$ .

## **4.2 Estadística Descriptiva**

En esta sección se presenta una breve descripción sobre la distribución de las variables que fueron utilizadas. La idea es obtener un panorama general de cómo está distribuida la muestra y hacia donde está la concentración de observaciones. Debido a que la muestra se restringió, los supuestos aplicados pudieron sesgar la distribución hacia un cierto tipo de individuo. Por lo anterior, la descripción comienza comparando los primeros momentos de la muestra con los de la *ENEU* 2004. Después, la descripción estadística será únicamente para la muestra.

La Tabla 3 presenta el promedio y desviación estándar de variables como edad, educación e ingreso después de haber aplicado los supuestos a la *ENEU* 2004. Empezando por la edad, el promedio en la muestra tiene 22.923 años, claramente inferior al promedio de 35.198 años en la *ENEU* (véase Tabla A.1.1 y A.1.2 del Apéndice). Esto se debe principalmente a que la

muestra es sólo de hijos, pues si se compara el promedio de edad de la *ENEU* de los hijos, aproximadamente son iguales<sup>6</sup>. De hecho, la desviación estándar de la *ENEU* pasó de 17.17 a tan solo 5.24.

Tabla 3  
Características de la Muestra

Variable	Media		Desviación estándar		Observaciones	
	Hombre		Hombre	Mujer		
Edad	22.92338		5.244958		9802	
Escolaridad	10.28627		3.54781		9802	
Ingreso	3529.157		3037.232		8858	
Numero de hermanos	3.070394		1.038757		9802	
<b>Padres</b>	Hombre	Mujer	Hombre	Mujer		
Edad	49.8562	47.10368	7.596431	7.395269	7333	4591
Escolaridad	7.524274	7.210303	4.703406	4.344473	7333	4591
Ingreso	5307.785	3215.523	6773.428	3338.051	6492	4233
<b>Familia</b>						
Escolaridad	9.097038		3.29017		9802	
Ingreso	3917.561		3057.414		9475	
Número de familiares en el hogar	4.783718		1.522322		9802	

Fuente: Elaboración propia usando los datos de la *ENEU* (2004)

Nota: La muestra es de hombres mayores a 15 y menores a 65 años de edad, "hijos", que trabajan y reciben ingresos positivos

En relación a la educación, el promedio en la *ENEU* es de 9.14 años comparado con 10.2 años en la muestra restringida. Ambos casos son superiores a los 8.2 años reportados en la literatura empírica, lo que confirma la tendencia de incrementos en los años promedio de escolaridad. Esto es aún más claro si se comparan los promedios de educación de los padres con el del hijo en la muestra, pues el promedio de los padres en la muestra es tan sólo de 7.52 (papa) y de 7.21 (mama).

En cuanto al ingreso, el mayor promedio en la muestra restringida lo tienen los padres con 5307.78 pesos mensuales, seguido por el del hijo con 3529.157 y por último de las madres (3215.523). A comparación de la *ENEU*, el ingreso del padre es el que más se aproxima al de la encuesta (5182.55 pesos), pero el del hijo y el de la madre incrementó bastante después de haber aplicado los supuestos<sup>7</sup>.

<sup>6</sup> La edad promedio en la *ENEU* 2004 de los hijos del jefe de hogar es de 21.558 para hombres y 21.867 para mujeres.

<sup>7</sup> Una posible explicación es que el supuesto de ingresos positivos, junto con el de horas positivas de trabajo hayan cortado la parte baja de la distribución del ingreso. También el supuesto de que los individuos ya no asistan a la escuela pudo haber tenido un efecto similar en la distribución, pues en edades tempranas el trabajo a tiempo parcial se practica mucho más en relación a trabajos de tiempo completo. Para el caso de las madres puede ir más allá de los supuestos. Podría ser que su ingreso promedio sea menor, porque no tienen tanta experiencia laboral como los hombres, por discriminación salarial por género o porque no trabajan.

Por otro lado, es interesante resaltar que la proporción de individuos que participan en la reforma es de un 67.34%, comparado con un 32.66% de individuos que no participan, por lo que la composición relativa de la muestra es principalmente gente que participó en la reforma (véase Cuadro A.1.3 en el Apéndice 1).

Si se comparan las medias de los grupos que participaron en la reforma contra los que no, resalta que el nivel de ingreso promedio del primer grupo es inferior al segundo por 1700 pesos o un 56.7 % menos como proporción del ingreso de las personas que participaron. Respecto a la educación, los niveles promedio dentro del grupo con participación presentan 9.7 años de escolaridad, también inferior al grupo sin participación (11.41). Destaca además, que el individuo promedio en el grupo con participación tiene 3 hermanos contra 2 en el grupo sin participación (véase Tabla 4).

Tabla 4  
Características de la Muestra condicionada en Participación en la Reforma

Variable	Media		Desviación estándar		Observaciones	
	Con Participación en Reforma	Sin Participación en Reforma	Con Participación en Reforma	Sin Participación en Reforma		
Edad	19.98803	28.97657	2.580993	4.011361	6601	3201
Escolaridad	9.738524	11.41581	3.031469	4.207418	6601	3201
Ingreso	2999.38	4702.743	2136.887	4189.287	6103	2755
Número de hermanos	3.064081	1.762887	1.024391	.9282091	6601	3201

Fuente: Elaboración propia usando los datos de la ENEU (2004)

En el Apéndice sección A.1 de este estudio, se presentan algunas estadísticas descriptivas sobre la distribución de las variables en la muestra. En esa sección de los cuadros A.1.4 a A.1.7 se presenta la distribución de las observaciones sobre si tienen una Restricción de Líquidez o no según el criterio de construcción de la variable. Para las 6 variables Habilidad se presentan también unos cuadros de frecuencia (véase cuadros A.1.8- A.1.13).

En relación a las variables de Restricción de Líquidez ( $L$ ) se puede observar que la muestra está compuesta por individuos con ingresos promedio del hogar del padre relativamente mayores a 1 y 2 líneas de pobreza, pues el 94.3% y 74.9% de la muestra tiene padres cuyos ingreso rebasan este nivel. Con las variables  $L III$  (mayor a mediana) y  $L IV$  (mayor a media) mejora la distribución, especialmente con la última pues se tiene un 52.7% por encima del promedio de ingresos de los papas.

En relación a las variables de Habilidad, la proporción es más o menos equitativa. Sin embargo esta cambia conforme se incrementa el nivel de educación, pues de un 43.13% de individuos con padres con niveles educativos con preparatoria y más (Hábil) pasó a un 35.19% cuando los individuos tienen padres con educación superior a preparatoria (véase frecuencia de variable *H II*). Sorprendentemente un 57.62% de la muestra tiene madres cuyos niveles de escolaridad están por encima de niveles educativos de preparatoria (Cuadro A.1.12). Esto podría deberse a la construcción de las variables familiares en la muestra.<sup>8</sup>

El cuadro A.1.14 del Apéndice, presenta los promedios de escolaridad e ingreso por cada variable de habilidad. Como es de esperarse, las personas que presentan un mayor nivel de habilidad, aquellos donde *H* toma valores de 1, reciben un ingreso mensual mayor que las personas no hábiles en todos los casos. Los niveles de educación también son superiores en todos los casos para las personas que tienen el atributo de habilidad en la muestra.

Asimismo, al observar los promedios de ingreso y educación con los niveles de Restricción de Liquidez (Cuadro A.1.15), es posible identificar que no existe una diferencia significativa, en los años promedio de escolaridad, de encontrarse en la primer restricción de liquidez y encontrarse en la segunda. Ya que en promedio los años de escolaridad de un individuo que se encuentre sobre una línea de pobreza (10.35 años), son aproximadamente iguales a estar en sobre dos líneas de pobreza (10.60 años). Esto confirma que la muestra tiene niveles promedio tanto de ingreso como de educación superiores al promedio de la *ENEU 2004* (2457.11 pesos y 9.14 años de escolaridad, véase Tabla A.1.1 y A.1.2). Para todas las restricciones de liquidez, los individuos que se encuentran por encima de una determinada restricción de liquidez, presentan mayores niveles de escolaridad promedio, en comparación con los individuos que no lo localizan en la misma restricción de liquidez.

---

<sup>8</sup> La ENEU ofrece una variable de parentesco pero en relación al jefe de hogar. La idea es que ser jefe de hogar no implica ser padre, de ahí que una hija del jefe de hogar, puede ser madre siempre y cuando tenga hija y el jefe de hogar no sea mujer. De ser hombre el jefe de hogar se inclinó a tomar como madre al “ascendiente del jefe de hogar” mujer siempre y cuando dicho jefe no tuviera “conyugue de jefe” mujer.