

mujeres entre los 55 y los 59 años de edad que estuvieron trabajando entre los 12 y 14, y cuyo trabajo durante este periodo en sus vidas no es explicado utilizando la información de retrospectiva de los individuos actualmente dentro de la fuerza de trabajo, no es mayor del 12%.

La participación de los jóvenes en la fuerza de trabajo parece haberse vuelto menos común en Bogotá, de acuerdo a la información presentada en "Participación en la Fuerza de Trabajo de Niños y Jóvenes en Colombia". Esto implica que mientras solamente el 1.7% de los actuales residentes de Bogotá entre los 12 a los 14 años se encuentran trabajando actualmente, una mayor proporción de jóvenes fueron económicamente activos en el pasado. Para poder explicar esto en el cálculo presentado en el párrafo anterior, hay que doblar la cifra de 1.7%. Ajustando ésta a 3.4%, la porción total de mujeres entre los 55 y 59 años que estuvieron trabajando de jóvenes no captada por los datos retrospectivos de la EPCVB es de 24%. Esta es una cifra máxima a medida en que considera uno de los grupos de mujeres de mayor edad.

Mientras un 24% no es un porcentaje insignificante, es probable que incluya un gran número de niños que realizaron trabajos de medio tiempo o de temporada. La mayoría de aquella gente joven que se encuentran actualmente fuera de la fuerza de trabajo, pero que trabajaron anteriormente, reconocieron que dejaron de trabajar para estudiar, o que tuvieron un empleo temporal. Casi el 70% de niños y niñas entre los 12 y 14 años de edad que se encontraban fuera de la fuerza de trabajo en 1992 y que han trabajado antes, dijeron haber dejado de trabajar para poder estudiar. Casi el 21% de los niños y el 16% de las niñas dijeron que habían completado un periodo de trabajo temporal. Una pequeña proporción de niños mencionaron enfermedades, accidentes y condiciones laborales no satisfactorias como razones para salir de la fuerza de trabajo. En general, las cifras implican que mientras algunos de los empleos más dañinos pueden no estar incluidos en la base de datos, el trabajo infantil y de jóvenes contabilizado en los datos retrospectivos en la EPCVB incluyen mucho del trabajo de tiempo completo y el trabajo que no se combina con la escolaridad. La mayor parte del trabajo que "se pierde" utilizando los datos retrospectivos de los participantes de la fuerza de trabajo en la EPCVB es de temporada o de medio tiempo.

Los datos sobre el historial de empleo de aquellos que se encuentran fuera de la fuerza de trabajo o desempleados se encuentra más desagregada en el cuadro 6. Solamente una pequeña porción de la gente entre los 18 y 24 que se encontraban de la fuerza de trabajo en 1992, comenzaron a trabajar antes de los 15 años de edad. Aún más, solamente el 1.5% de las mujeres entre los 18 y 19 años, y ningún hombre, han trabajado entre 2 y 5 años, antes de estar fuera de la fuerza de trabajo. Para las mujeres de 20 a 24 años, 5.6% trabajaron de 2 a 5 años antes de estar fuera de la fuerza de trabajo, y 2.4% más de 5 años antes de salir también. Entre los hombres de la misma edad, 1.5% trabajaron de 2 a 5 años antes de salir de la fuerza de trabajo. Para una gran proporción de los desempleados entre los 15 y los 19 años esta no era su primera búsqueda de empleo. La mayoría de los hombres, y más de un tercio de las mujeres, han trabajado antes de ser desempleados. De nuevo, los desempleados conforman una pequeña porción de los jóvenes que trabajan.

Otro medio de validar los datos retrospectivos de la EPCVB es comparar las tasas retrospectivas de la participación en la fuerza de trabajo para los periodos en que los grupos dados habrían comenzado a trabajar. El grupo de 30 a 34 años de la EPCVB hubiera estado aproximadamente en el rango de edad de 12 a 14 años a principios de los 70. Las Encuestas Nacionales a Familias de 1972 sugieren que aproximadamente 10% de los que se encuentran de los 12 a 14 años eran económicamente activos. Para los hombres el porcentaje es aproximadamente 12% y para las mujeres cerca del 8%. Los datos

retrospectivos de la EPCVB muestran que 20% de los hombres y 14% de las mujeres comenzaron a trabajar entre los 12 y 14 años, y que 16% de las mujeres y 31% de los hombres ya habían comenzado a trabajar a los 14 años. Dado que la información para la participación de niños y jóvenes en la fuerza de trabajo durante los meses de vacaciones escolares es dos veces mayor que durante los meses de clases, estos datos retrospectivos no son sorprendentes. Sugieren que aún dado el hecho de que los datos retrospectivos pierden algo del trabajo de temporada y de medio tiempo desempeñado por los jóvenes, una gran porción del trabajo que se incluye en estos datos es de temporada o no continuo. Es interesante que el artículo "Participación en la Fuerza de Trabajo de Niños y Jóvenes" muestra que más del 40% de los jóvenes que trabaja que se encuentra de los 12 a 14 años combinaron el trabajo y la escuela en las áreas urbanas de Colombia en 1992, y que los datos fueron similares para 1972.

Los resultados presentados anteriormente sugieren que el análisis de regresión presentado posteriormente en este artículo tenderá a errar en el sentido de sobrestimar los efectos negativos de la experiencia laboral temprana. La sobrestimación dependerá del grado en el que el trabajo que se pierda en los datos retrospectivos de la EPCVB sea menos intenso, estacional o de medio tiempo que el trabajo que es medido. Es más probable que el trabajo que se pierde sea compatible con la escolaridad y menos probable que sea dañino para el joven. La conclusión de que el sesgo es hacia el sobrestimar un efecto negativo está motivando ambos supuestos que son utilizados en el artículo para desarrollar estimaciones del *trabajo mientras se estudia*, y en reducir la posibilidad de que los resultados no reflejen las formas más intensas de participación de trabajo infantil y de jóvenes.

Un elemento cualitativo y cuantitativo más serio es el que la información disponible de la EPCVB sobre la edad de entrada a la fuerza de trabajo excluye al trabajo no remunerado en el propio hogar. En otras palabras, la pregunta se redacta de tal manera que la persona que "trabaja" de tiempo completo cuidando su propio hogar o niños pequeños no se le considera como si trabajara. Esto sesga el análisis de este artículo, particularmente con respecto a las mujeres. Una base de datos ideal incluiría información acerca de la edad en la que la gente comenzó a trabajar en el hogar, y el tipo de trabajo y número de horas empleadas.

Muchas niñas y mujeres jóvenes pasan muchas horas trabajando en el hogar y cuidando niños. Como se discute en "Participación de Niños y Jóvenes en la Fuerza de Trabajo en Colombia", el 5% de las mujeres en las áreas urbanas de 12 a 14 años, y el 10% de 15 a 17 años, consideran el trabajo en el hogar como su actividad "primaria". El trabajo no remunerado en el hogar, a pesar de que sea insignificante si se realiza por periodos cortos bajo supervisión, puede involucrar muchas horas. Entre las mujeres de 12 a 14 años viviendo en Bogotá, el 75% reportan horas positivas utilizadas en el trabajo doméstico. Para el 75% que realizan este tipo de trabajo, el promedio de horas semanales trabajadas es 21. Entre estos trabajadores, el 27% ocupan 30 horas o más de trabajo por semana. Mientras que para los hombres el 41% no utilizan nada de tiempo realizando tareas domésticas, entre los que si lo hacen el promedio es de 15 horas por semana, con un 14% utilizando 30 horas o más por semana. Es menos probable que las mujeres jóvenes que realizan trabajo doméstico como actividad primaria asistan a la escuela, a diferencia de los que trabajan por un salario fuera del hogar. El trabajo en si mismo no puede proporcionar entrenamiento que lleve en el futuro a oportunidades en el mercado de trabajo.

Una dificultad con el análisis de regresión presentado posteriormente, es el que la variable sobre la edad de entrada a la fuerza de trabajo es una variable endógena. La decisión de ingresar a trabajar puede estar determinada por los salarios o por la recompensa a la entrada temprana que el joven o la familia esperaban obtener. Un problema similar puede surgir con la educación, a medida que la persona o su

familia eligen su nivel de educación. Al no tener los instrumentos adecuados para referirse a estos problemas, el análisis empírico supone que ambas variables deben tratarse como exógenas.

Otra manera de considerar este problema analítico es la heterogeneidad individual. Sin un experimento natural, los individuos que eligen llevar a cabo ciertas actividades no serían una muestra aleatoria comparados con los grupos que no realizan las mismas actividades. Bajo este escenario, no se sabe cual sería el efecto de, por ejemplo trabajar de niños o combinar el trabajo y la escuela, sobre la gente que no realizó estas actividades. De manera similar, no se sabe cuales serían los resultados, en este caso los salarios de los adultos, para los individuos que sí trabajaron si no lo hubieran hecho.

Dada la naturaleza de corte transversal de los datos, es imposible controlar las muchas causas del sesgo asociados con la heterogeneidad de los individuos. Aún más, los datos no incluyen una medida de la habilidad, y el sesgo por la omisión de una variable que la represente puede llevar a una sobrestimación de los rendimientos de la escolaridad.¹⁹ Aún así, mientras que la base de datos de la EPCVB no es un panel, si incluye varias variables relacionadas con características personales y del empleo, y de antecedentes familiares. A pesar de que esto es solamente una solución parcial a los problemas de selectividad y de heterogeneidad individual, el incluir estas variables en el análisis empírico proporciona una indicación sobre el grado en el que los resultados son influidos por otros factores relacionados con los antecedentes familiares.

Además de incluir las variables de control, también es posible evaluar la dirección en la heterogeneidad individual podría estar influyendo a los resultados presentados posteriormente. Bajo ciertos escenarios, si las variables disponibles no controlan lo suficiente la heterogeneidad individual, puede existir un sesgo hacia arriba en la estimación de los coeficientes que miden la experiencia laboral temprana. Esto llevaría a una sobrestimación de el verdadero efecto de la entrada temprana a la fuerza de trabajo sobre los ingresos como adulto. En otras palabras, la entrada temprana parecería tener efectos más positivos (o menos negativos) sobre los ingresos que lo que es en realidad el caso. Esto sería verdad si los niños que ingresaron a la fuerza de trabajo fueron aquellos que era más probable que se beneficiaran de esta experiencia temprana, posiblemente en el sentido de ser más capaces físicamente de combinar el trabajo y la escuela. Una familia puede elegir enviar a sus miembros más sanos a trabajar, bajo el supuesto de que serán capaces de funcionar más exitosamente y ganar más. También es posible que los niños que combinan el trabajo y la escuela son los más capaces, y por lo tanto son los que pueden hacer esto de manera exitosa. Los niños menos capaces, o aquellos que provienen de ambientes más pobres, pueden no estar preparados intelectual o emocionalmente para sobresalir en ambas actividades al mismo tiempo.

Un punto relacionado es que los ingresos de los dueños de negocios y los auto empleados pueden incluir un componente, que es el rendimiento del capital físico. Esto puede causar una sobrestimación de los rendimientos del capital humano en estos sectores. Debido a estas diferencias, las funciones de ingreso para los trabajadores asalariados y no asalariados son a veces estimadas por separado, como se hace en muchas regresiones en este artículo (Schultz, 1990; Tenjo, 1993).

Aún así, la medida de los salarios e ingresos totales, como se manejan en este artículo, tiene importantes ventajas sobre las medidas que separan los salarios de los ingresos. Primero, el tamaño de la muestra se

¹⁹ Tenjo (1993), utilizando una encuesta de 1988 de 2000 trabajadores en Bogotá, encuentra que el incluir medidas sobre la habilidad y el conocimiento reduce los rendimientos de la educación de 10.2% a 8.6% para los hombres, y de 9.8% a 8.1% para las mujeres. Los rendimientos de la habilidad y el conocimiento son más altos para las mujeres que para los hombres.

incrementa substancialmente al incluir a los auto empleados. Segundo, la exclusión de ciertos grupos de educación del análisis de salarios limita el rango de ocupaciones y sectores que se incluyen, e introduce una causa adicional de selectividad en el modelo que puede ser importante tanto para hombres como para mujeres. Mientras que los auto empleados y los trabajadores no remunerados pueden no conformar una proporción grande de participantes en la fuerza de trabajo en los países con ingresos altos, en los países en desarrollo estas categorías pueden constituir la mitad de los hombres adultos empleados y dos tercios de las mujeres adultas (Schultz, 1991). Relativamente pocos estudios toman en cuenta las diferencias entre los hombres y las mujeres de los sesgos en los rendimientos de la educación generados por utilizar una muestra selectiva de generadores de ingreso (Schultz, 1991). En los datos de la EPCVB, la selectividad del sector salario (aún sin corregir la exclusión de los trabajadores familiares no remunerados) prueba ser especialmente importante para las mujeres, ya que el restringir la muestra puede excluir ocupaciones más flexibles y dar mayor ponderación en la categoría de baja remuneración de los trabajadores domésticos.

Que los resultados sean *robustos* se prueba corriendo cada regresión utilizando tres especificaciones: una formulación simple que incluye solamente medidas de la educación, experiencia y edad al entrar; otra con variables explicativas, controlando los antecedentes familiares y características personales; y una con variables dummy adicionales para la ocupación y la industria. Las variables sobre antecedentes familiares incluyen: la educación y ocupación del padre del jefe de familia donde vive la persona; la educación de la madre; y si la principal ocupación de la madre del jefe de familia era el trabajo. En algunas ecuaciones es posible identificar los distintos efectos de las distintas ocupaciones de la madre.²⁰ En estos casos, se incluye un conjunto más completo de control. Las variables dummy también se incluyen para indicar: si la persona está relacionada con el jefe de familia (no un inquilino, empleado doméstico, invitado, etc.); si es padre o suegro del jefe de familia; si está involucrado en actividades sindicales o políticas; si está involucrado en actividades de la comunidad; si es miembro de un club deportivo o cultural; si el jefe de familia vivió en áreas rurales durante su infancia; si la familia emigró a Bogotá; y si la persona trabaja en una empresa no sindicalizada. Los coeficientes de la variable de control se discuten en el texto, y un conjunto completo de los resultados está disponible en el Apéndice 1.

El sesgo de selección es un problema importante en las ecuaciones de ingresos, especialmente para las mujeres. A pesar de que la base de datos se describe con mayor detalle más adelante, el cuadro 7 proporciona una visión general de la selectividad. Los jefes de familia que son hombres se incluyen en el estudio como un grupo aparte debido a que la mayoría tienen salarios e ingresos observables. Los resultados para este grupo ayudan a identificar la naturaleza de la selectividad para los hombres. Las funciones de ingresos similarmente proveen información sobre la selectividad en las ecuaciones de salarios.

Para proporcionar información adicional sobre un posible sesgo en la selección, se comparan los resultados de los siete modelos de MCO de hombres y mujeres incluyendo antecedentes familiares y características personales con los resultados utilizando un modelo de Heckman para corregir la selectividad. El modelo implica primero estimar una ecuación Probit que prediga la probabilidad de

²⁰ La información sobre los padres del jefe de familia no es un sustituto de los antecedentes familiares de los otros miembros de la familia. Todavía, para los cónyuges, hijos y hermanos del jefe de familia, estos datos pueden dar un indicio de la historia familiar. Los miembros que no pertenecen a la familia se les asigna un valor de cero en las variables de antecedentes familiares, y se incluye una variable dummy adicional para absorber algunas de las diferencias entre ellos y los miembros de la familia. Similarmente, se les da a los abuelos un valor de cero y se incluye una variable dummy adicional para ellos.

tener un valor diferente de cero para el logaritmo de salarios o ingresos. Los resultados del Probit se usan para desarrollar un término de corrección de selección de la muestra en la estimación de las ecuaciones de ingresos y salarios. Las ecuaciones de Probit se identifican utilizando una serie de variables que se excluyen de las funciones de ingresos/salarios. Incluyen variables dummy para indicar el estado civil de la persona, una sola dummy para indicar si existe algún infante de 4 años o manos en la casa, y la cantidad total de ingreso familiar no ganado. Las ecuaciones Probit también incluyen una medida de la educación (que corresponde a la ecuación de salarios/ingresos), un conjunto simple de dummies para la ocupación del padre y de la madre del jefe de familia, y controla si la persona es inmigrante o nació en áreas rurales.

Existe dificultad en utilizar los modelos de selección que son específicos de este artículo sobre la entrada temprana a la fuerza de trabajo. Dado que la información sobre la edad de entrada al mercado de trabajo solamente está disponible para los empleados, cualquier variable derivada de la edad de entrada no puede ser utilizada en las ecuaciones Probit. Esto implica que la experiencia (total) y la combinación de trabajo y escuela deben ser excluidas del Probit en el modelo 2 al 5. Los modelos que incluyen la experiencia total también incluyen la edad y su cuadrado, una variable disponible para toda la muestra, en las ecuaciones Probit. En la ausencia de una variable que especifique la edad y la experiencia, es probable que las variables dummy para el estado civil absorban algo de la variación relacionada al ciclo de vida como un determinante de la participación en la fuerza de trabajo. Aún más, el incluir las variables de antecedentes familiares y personales en las ecuaciones de salarios parece absorber buena parte del sesgo de selección. Como se discute más adelante, al incluir estas variables, los rendimientos de la educación se reducen así como también la significancia de los términos de selección en los modelos de Heckman. El incluir la edad en el Probit no es una solución ideal al problema de no tener información acerca de la edad de entrada, y el Apéndice 2 discute más detalladamente la sensibilidad de los resultados al excluir e incluir la edad y las medidas de la experiencia en el Probit.²¹

En general, las variables de identificación tienen el signo correcto y son significativas en las ecuaciones Probit. Los resultados se reportan en el Apéndice 1. Es menos probable que los hombres trabajen si son solteros o viudos, sugiriendo que estas variables recogen efectos sobre el estilo de vida. Las mujeres sin pareja es más probable que trabajen. El ingreso familiar no ganado afecta negativamente la probabilidad de trabajar tanto para hombres como para mujeres.

La inestabilidad de los modelos de selección de la muestra es muy conocida (Mroz, 1987; Falaris, 1995). Otra dificultad con estos modelos es que la selección involucra más que solamente las diferencias entre los participantes en la fuerza de trabajo y los no participantes (Schultz, 1991). En efecto, los modelos de selección en este artículo deben también tomar en cuenta las personas sin ingresos y que se encuentran en el mercado de trabajo (trabajadores familiares y los desempleados), y en el caso de las ecuaciones de salarios, los trabajadores no asalariados (auto empleados y patrones). Teóricamente, esto debe considerarse como un proceso múltiple de selección con diferentes medios de estimar cada proceso. Como la selectividad dentro del empleo remunerado no es el enfoque principal de este artículo, esto debe mantenerse en mente y los resultados de las regresiones de Heckman debe utilizarse como puntos de

²¹ Un conjunto de datos nacionales obtenidos en 1993 debe estar disponible durante 1995 y proporcionarán una visión adicional. Los datos nacionales incluyen la edad de entrada a la fuerza de trabajo para todos los jefes de familia, sin importar su status de empleo. Con estos datos, será posible calcular el tamaño del sesgo en las correcciones de Heckman en las ecuaciones que incluyen los jefes de familia que son hombres.

referencia. La sensibilidad de los modelos de Heckman ante cambios en la especificación se exploran con mayor profundidad en el Apéndice 2.

Las medias y las desviaciones estándar de las variables utilizadas en los modelos se presentan en los cuadros 8a, 8b y 8c. Los resultados del trabajo empírico se presentan en los cuadros 9a a 18c. Cada modelo tiene tres cuadros (a, b, y c), uno para hombres, mujeres y jefes de familia. Cada cuadro tiene dos conjuntos de datos, el primero para los salarios y el segundo para los ingresos. Dentro de cada conjunto existen cuatro columnas, cada columna correspondiendo a una regresión. La primera regresión en cada conjunto incluye solamente las variables básicas de educación, experiencia, edad de entrada, y trabajo mientras se estudia. La segunda columna se refiere a las regresiones que incluyen los controles de los antecedentes familiares y características personales. Las terceras columnas de cada conjunto de los cuadros para los hombres y mujeres (no jefes de familia que son hombres) incluyen los resultados utilizando las correcciones de sensibilidad. Las regresiones en este modelo pueden ser comparadas con las segundas columnas con las regresiones que incluyen el control de los antecedentes familiares y las características personales. Las últimas columnas dan los resultados de las regresiones con estos controles de antecedentes y variables personales, así como las dummies para la industria y la ocupación. Solo los coeficientes relacionados a la educación, experiencia y edad de entrada se reportan en los cuadros con texto. Todos los coeficientes se reportan en el Apéndice 1.

5. Entrada temprana a la fuerza de trabajo y el entrenamiento

Los signos y magnitudes de los resultados de las regresiones utilizando la medida "tradicional" de la experiencia (modelo 1) son como se esperaba son comparables a los resultados de otras fuentes. La cantidad de varianza explicada es razonable para este tipo de regresiones, aunque un poco bajas si se hubieran utilizado las Encuestas Nacionales a Familias.²² El rendimiento de un año de educación para los hombres utilizando la especificación simple sin variables de control, es 11.4% en las ecuaciones de salarios y 11.3% en las ecuaciones de ingresos. Para las mujeres, los rendimientos son muy similares y para los jefes de familia que son hombres son un poco más bajos. El coeficiente en la variable dummy para la educación post universitaria es mayor para las mujeres. Como se esperaba, los rendimientos de la experiencia tienden a ser más altos y mostrar mayor curvatura para los hombres tanto en las ecuaciones de ingresos como de salarios (cuadros 9a, 9b y 9c).²³

El añadir los controles tiene el resultado esperado de reducir la magnitud de los rendimientos de la educación ya que los antecedentes familiares y las características personales están probablemente correlacionadas con el logro educativo. De manera similar, el tipo de empleo que una persona puede esperar obtener es en parte determinado por su nivel de educación. Una porción de los rendimientos de la educación surgen en las diferencias en los sectores, ocupaciones y condiciones bajo las cuales la gente trabaja. En el modelo tradicional con los controles, los rendimientos de la educación para las mujeres tiende a ser un poco más baja que para los hombres. Los rendimientos de la educación post universitaria son consistentemente más significativos y más altos para las mujeres.

²² La robustez de los resultados y la comparabilidad a otras encuestas es motivante. La EPCVB es uno de los primeros ejemplos de una encuesta sobre la pobreza y estándares de vida utilizado en Colombia. El cuestionario y muestra son relativamente nuevos y experimentales.

²³ Los diferenciales por sexo en los rendimientos de la educación post universitaria se discuten en detalle en *Opposite Ends of the Spectrum*.