

## Capítulo 3

### **Aplicación de métodos analíticos para la evaluación de impacto: estudio de caso\***

Este estudio de caso se basa en un programa hipotético contra la pobreza, PROSCOL, que proporciona transferencias en efectivo orientadas hacia familias pobres con hijos en edad escolar en una región de un determinado país en desarrollo. Con este caso se desean ilustrar los pasos analíticos que componen la puesta en práctica de una evaluación de impacto y las opciones que puede enfrentar un analista, en un proceso que es aplicable a cualquier tipo de programa contra la pobreza. Al examinar la forma de evaluar el impacto del programa, el analista de políticas incurre en diversos errores comunes a lo largo del proceso y busca información sobre temas específicos recurriendo a los conocimientos especializados de colegas: un estadístico, un economista, un profesor de econometría y un sociólogo.

Entre los pasos analíticos que el analista sigue en el caso están los siguientes: identificar las preguntas que se deben abordar en la evaluación del impacto, evaluar los datos disponibles, dar un primer vistazo a los datos, comprender los sesgos, aprender acerca del ingreso no percibido, agregar variables de control, comprender la importancia de la exogeneidad, explorar mejores formas de crear un grupo de comparación (correspondencia de puntuación de la propensión), conocer los sesgos debido a elementos no observables, examinar lo que se podría haber hecho con una encuesta básica o de referencia (dobles diferencias), usar variables instrumentales, probar las diversas metodologías, incorporar información recogida en el terreno y planificar trabajo futuro.

#### **Descripción del programa hipotético, PROSCOL**

En el programa PROSCOL se identifican las familias calificadas para participar usando diversas variables sustitutivas de la pobreza, entre ellas el número de personas en el hogar, la educación del jefe de hogar y diversos atributos de la vivienda. PROSCOL paga una cantidad fija por niño en edad escolar a todos los hogares seleccionados, a condición de que los niños asistan al 85% de sus clases escolares. Esto debe verificarse mediante una nota de la escuela. Los hogares deben mantener a sus hijos en la escuela hasta los 18 años.

Este programa se introdujo hace 12 meses, está financiado por el Banco Mundial y funciona a partir del Ministerio de Desarrollo Social. En un esfuerzo por evaluar el impacto de PROSCOL en la pobreza y así ayudar a determinar si se debe ampliar el programa hasta abarcar el resto del país o abandonarse, el Banco Mundial ha solicitado que el Ministerio de Finanzas evalúe sus efectos. La solicitud se hizo al Ministerio de Finanzas, para así asegurar una evaluación independiente y para ayudar a crear capacidades para este tipo de evaluaciones en una unidad central del gobierno, una unidad cercana a la entidad en que se

---

\* Este capítulo se basa en gran medida en un documento de información básica de Martin Ravallion, *The Mystery of the Vanishing Benefits: Ms. Speedy Analyst's Introduction to Evaluation*, Documento de trabajo de estudio de políticas N° 2153, 1999.

realizan las asignaciones presupuestarias.

### **Identificación de las preguntas que se abordarán en la evaluación del impacto**

El primer paso para el analista del Ministerio de Finanzas asignado a la tarea de la evaluación de PROSCOL es aclarar los objetivos del proyecto que se deberán considerar al evaluar sus impactos. El proyecto tiene dos objetivos: uno, las transferencias de efectivo, destinadas a reducir la pobreza actual y dos, al insistir en que los destinatarios de la transferencia mantengan a sus niños en la escuela, intentar reducir la futura pobreza aumentando los niveles de educación entre la actual población de niños pobres. Por consiguiente, se necesitarán dos elementos de información sobre el programa para evaluar sus impactos. Primero ¿las transferencias de efectivo llegan principalmente a familias de bajos ingresos? Y segundo, ¿en qué medida el programa está aumentando las tasas de matrícula escolar?

### **Evaluación de los datos disponibles**

Para realizar la evaluación, el analista cuenta con dos recursos principales. El primero es un informe basado en entrevistas cualitativas a los administradores de programas y grupos representativos de participantes. Sin embargo, no está claro si los entrevistados eran representativos de los participantes en PROSCOL o cuán pobres eran con respecto a los que no se seleccionaron para el programa y no fueron entrevistados. El informe señala que los niños asistían a la escuela, pero es posible que también habrían ido a la escuela si no hubiera existido el programa. Aunque este informe es un punto de partida importante, no indica al analista el nivel de pobreza de los participantes en PROSCOL y cómo el programa afecta la escolaridad. La segunda fuente es una encuesta domiciliaria independiente a nivel nacional realizada por la Oficina de Estadísticas del país, llamada Encuesta Sobre las Condiciones de Vida (ENCOVI). La ENCOVI incluía una muestra aleatoria de 10.000 hogares y consultaba acerca de los ingresos del hogar por fuente, empleo, gastos, estado de salud, grados de instrucción, atributos demográficos y otras características de la familia. La encuesta incluía una pregunta sobre si el hogar muestreado había participado en PROSCOL y una partida para el dinero recibido de PROSCOL en el listado de las fuentes de ingresos.

### **Un primer examen de los datos**

Luego el analista se consigue los datos sin elaborar de la ENCOVI, con el fin de evaluar quién se beneficia con el programa. Utiliza un paquete de software estadístico, como SPSS o SAS, para generar una tabulación cruzada de la cantidad promedio recibida de PROSCOL por los deciles de hogares, donde los deciles se crean clasificando todos los hogares de la muestra según su ingreso por persona. Para calcular esto último, el analista decide restar todos los dineros recibidos de PROSCOL como una buena medida de los ingresos en ausencia del programa, y así identificar quién se benefició con respecto a su ingreso anterior a la intervención.

La tabulación cruzada sugiere que las transferencias de efectivo del programa están orientadas adecuadamente hacia los pobres. Según la línea de pobreza oficial del país, aproximadamente el 30% de la población del noroeste es pobre. Los cálculos del cuadro indican que el 30% más pobre de la muestra de la encuesta recibe el 70% de las transferencias de PROSCOL. A primera vista, esto parece ser un resultado positivo.

La pregunta siguiente es acerca del impacto en la escolaridad. Éste se examina a través de una tabulación cruzada de las tasas de matrícula escolar promedio de diversos grupos de edades para familias dentro de PROSCOL versus familias fuera de PROSCOL. La tabulación casi no muestra diferencia alguna entre ambos grupos; la tasa de matrícula promedio para niños entre 6 y 18 años es de alrededor del 80% en los dos casos. Luego el analista calcula los años promedio de escolaridad para cada edad y los resultados se representan por separado para las familias dentro de PROSCOL y las familias fuera de éste. Si bien las dos cifras no son idénticas, son muy cercanas. En esta etapa, el analista se pregunta si en realidad el programa no ha influido en la escolaridad o si el enfoque es incorrecto.

### Comprensión de los sesgos

Con esta incertidumbre, el analista luego busca información de un estadístico superior para que averigüe por qué los resultados sugieren que los niños del PROSCOL no tienen más posibilidades de estar en la escuela que los niños fuera del PROSCOL. El estadístico plantea la hipótesis de que los resultados pueden tener un sesgo importante. Para evaluar el impacto del programa, es necesario saber qué habría sucedido sin el programa. Sin embargo, el analista no ha tomado en consideración este punto; en su lugar, las familias fuera del PROSCOL se usan como grupo de comparación para inferir lo que habría sucedido con la escolaridad de los participantes del PROSCOL si no hubiera existido el programa.

En otras palabras,  $P_i$  representa la participación en PROSCOL del  $i$ -ésimo niño. Esto puede tener dos valores posibles, a saber,  $P_i = 1$  si el niño participa en PROSCOL y  $P_i = 0$  si no lo hace. Si el  $i$ -ésimo niño no participa, su nivel de escolaridad es  $S_{0i}$ , lo que representa la escolaridad  $S$  del niño  $i$  cuando  $P = 0$ . Si el niño participa, su escolaridad es  $S_{1i}$ . Su ganancia en la escolaridad debido a PROSCOL es  $S_{1i} - S_{0i}$ . Luego, la ganancia para el  $i$ -ésimo niño que participe ( $P = 1$ ) es

$$G_i = S_{1i} - S_{0i} \mid P_i = 1.$$

El símbolo  $\mid$  significa “dado que” o “bajo la condición de que” y es necesario para dejar en claro que el cálculo es la ganancia para un niño que realmente participó. Si se desea saber la ganancia promedio, ésta simplemente es la media de todas las  $G_i$ , con lo que se obtiene la ganancia media en escolaridad entre todos los que participaron en PROSCOL. Siempre que esta media se calcule en forma correcta (usando las ponderaciones de muestras adecuadas de la encuesta), proporcionará una estimación insesgada de la verdadera ganancia media. Esta última es el “valor esperado” de  $G$  y puede representarse como

$$G = E(S_{Ii} - S_{0i} \mid P_i = 1).$$

Esta es otra forma de decir “media”. Sin embargo, no necesita ser exactamente igual a la media calculada de los datos de la muestra, dado que habrá algún error de muestreo. En la literatura sobre evaluación,  $E(S_{Ii} - S_{0i} \mid P_i = 1)$  a veces se denomina el “efecto tratamiento” o “efecto tratamiento promedio de los tratados”. En este caso, PROSCOL es el tratamiento.

El estadístico señala al analista que éste no ha calculado  $G$ , sino más bien la diferencia en la escolaridad media entre los niños de familias beneficiadas por PROSCOL y los de familias fuera de PROSCOL. Esta es la estimación de muestra de

$$D = E(S_{Ii} \mid P = 1) - E(S_{0i} \mid P = 0)$$

Hay una simple identidad que vincula  $D$  y  $G$ , a saber:

$$D = G + B.$$

Este término “ $B$ ” es el sesgo de la estimación y está dado por

$$B = E(S_{0i} \mid P_i = 1) - E(S_{0i} \mid P_i = 0)$$

En otras palabras, el sesgo es la diferencia esperada en escolaridad sin PROSCOL entre los niños que de hecho participaron en el programa y aquellos que no lo hicieron. Se podría corregir este sesgo si se conociera  $E(S_{0i} \mid P_i = 1)$ , pero ni siquiera es posible obtener una estimación de muestra de éste. No se puede observar cómo habría sido la escolaridad de los niños que realmente participaron en PROSCOL si no hubieran participado; estos son datos faltantes, también denominados una media “contrafactual”.

Este sesgo es muy importante. A falta del programa, los padres de PROSCOL probablemente habrían enviado a sus hijos a la escuela en menor proporción que otros padres. Si hubiera sido así, habrá un sesgo en el cálculo. Volviendo a las preguntas de evaluación originales, lo que interesa es la escolaridad adicional debido a PROSCOL. Supuestamente esto sólo afecta a las familias que efectivamente participan. Es decir, es necesario saber cuánta escolaridad menos se podría haber esperado sin el programa. Si no hay sesgo, la escolaridad adicional generada por el programa es la diferencia en la escolaridad media entre los que participaron y los que no lo hicieron. Así, el sesgo se genera si hay una diferencia en la escolaridad media entre los padres dentro de PROSCOL y los padres fuera de PROSCOL en ausencia del programa.

Para eliminar el sesgo, el mejor método sería asignar el programa en forma aleatoria. Entonces los participantes y no participantes tendrían la misma escolaridad esperada sin el programa, esto es,  $E(S_{0i} \mid P_i = 1) = E(S_{0i} \mid P_i = 0)$ . La escolaridad de las familias no participantes, por consiguiente, revelaría en forma correcta el escenario contrafactual, es decir, la escolaridad que se habría observado para los participantes si no hubieran tenido acceso al programa. Por cierto, la asignación aleatoria será igual a la distribución total, no sólo a la media. Aún existiría un sesgo debido al error de muestreo, pero para muestras lo suficientemente grandes se puede suponer con seguridad que cualquier diferencia

estadística significativa en la distribución de la escolaridad entre los participantes y no participantes es atribuible al programa.

Dentro del actual diseño del programa, es claro que la participación no es aleatoria. Por cierto, sería una seria crítica para PROSCOL descubrir que sí lo era. El mismo hecho de su orientación intencional hacia las familias pobres (que supuestamente tienen menos posibilidad de enviar a sus hijos a la escuela) crearía un sesgo.

Esto plantea la siguiente pregunta: si PROSCOL funciona bien, se debería esperar que los participantes tuvieran peor escolaridad sin el proceso. Luego  $E(S_{0i} | P_i = 1) < E(S_{0i} | P_i = 0)$ . y el cálculo original de los analistas subestimaría la ganancia del programa. Se podrían encontrar muy pocas o ninguna ganancia, aun cuando el programa estuviera funcionando en forma adecuada.

Ahora el analista examina si la magnitud de este sesgo podría ser considerable. Supóngase que las familias pobres envían a sus hijos a trabajar en lugar de la escuela; como son pobres y no pueden endeudarse fácilmente, necesitan dinero adicional ahora. Las familias que no son pobres envían a sus niños a la escuela. El programa selecciona a familias pobres, las que entonces envían a sus niños a la escuela. Se observa una diferencia insignificante en la escolaridad media entre las familias de PROSCOL y las familias fuera de éste; en efecto,  $E(S_{1i} | P_i = 1) = E(S_{0i} | P_i = 0)$  en valor esperado. Pero el impacto del programa es positivo y está dado por  $E(S_{0i} | P_i = 0) - E(S_{0i} | P_i = 1)$ . El no haber considerado la focalización intencional del programa hacia los pobres, bien podría haber producido una subestimación considerable de los beneficios de PROSCOL cuando el analista comparó la escolaridad media entre familias dentro de PROSCOL y familias fuera de éste.

### **Adquisición de conocimientos sobre el ingreso no percibido**

A continuación, el analista muestra a otro colega, un economista del Ministerio de Finanzas, los resultados de su tabulación cruzada de cantidades recibidas de PROSCOL versus el ingreso. El economista plantea un aspecto importante: las ganancias que los pobres obtuvieron de PROSCOL claramente fueron sobreestimados, porque se ha ignorado el ingreso no percibido o sacrificado. Los niños deben ir a la escuela si la familia desea recibir la transferencia de PROSCOL; por consiguiente no podrán trabajar en el negocio de la familia ni en el mercado laboral. Por ejemplo, los niños entre 15 y 18 años pueden ganar dos tercios o más del salario de un adulto en el sector agrícola y de la construcción. Las familias que participan en PROSCOL perderán este ingreso proveniente del trabajo de sus hijos. Este ingreso no percibido se debe tomar en cuenta al calcular las ganancias de ingreso netas que produce el programa. Y se debe restar esta ganancia de ingreso neta, no la transferencia bruta, para calcular el ingreso previo a la intervención. Esto también tendrá importancia al determinar qué tan pobre habría sido la familia en ausencia de la transferencia de PROSCOL. En consecuencia, el cuadro actual podría sobrestimar en forma considerable los beneficios que tiene el programa para los pobres.

El analista se pregunta por qué debe descontar el ingreso no percibido del trabajo de los niños, dado que menos trabajo infantil es algo positivo. El economista explica que debe examinar los beneficios de un menor trabajo infantil, del cual el principal beneficio es la

escolaridad adicional y, por ende, futuros ingresos más altos para las familias actualmente pobres. El analista ha generado cuadros que reflejan las dos formas principales en que PROSCOL reduce la pobreza: aumentando los actuales ingresos de los pobres y aumentando sus ingresos futuros. El impacto en el trabajo infantil es importante para ambos casos, pero en sentidos opuestos; por consiguiente, PROSCOL representa una relación inversa de compensación.

Esto pone de relieve por qué es importante obtener una buena estimación del impacto en la escolaridad; sólo entonces será posible determinar el ingreso no percibido. Por ejemplo, es posible que el tiempo adicional en la escuela sea en desmedro del tiempo no trabajado.

Con respecto a la segunda tabulación cruzada, el aspecto principal planteado por el economista es que no se consideran todas las demás determinantes de la escolaridad, aparte de la participación en PROSCOL. Él sugiere ejecutar una regresión de los años de escolaridad en un conjunto de variables de control y también si la familia del niño estaba o no cubierta por PROSCOL. Para el  $i$ -ésimo niño de la muestra, sea

$$S_i = a + bP_i + cX_i + \varepsilon_i .$$

Aquí  $a$ ,  $b$  y  $c$  son parámetros;  $X$  representa las variables de control, como edad del niño, educación del padre y de la madre, tamaño y composición demográfica del hogar y características de la escuela; y  $\varepsilon$  es un valor residual que incluye otras determinantes de la escolaridad y errores de medición. El valor estimado de  $b$  proporciona el impacto de PROSCOL en la escolaridad.

Cabe observar que si la familia del  $i$ -ésimo niño participa en PROSCOL, entonces  $P = 1$  y su escolaridad será  $a + b + cX_i + \varepsilon_i$ . Si no participa, luego  $P = 0$  y su escolaridad será  $a + cX_i + \varepsilon_i$ . La diferencia entre ambos es la ganancia en la escolaridad debido al programa, que es exactamente  $b$ .

### **Incorporación de variables de control**

Como se sugirió, el analista luego ejecuta una regresión con y sin las variables de control. Cuando la ejecuta sin ellas, los resultados muestran que el valor estimado de  $b$  no es significativamente distinto de cero (usando la prueba  $t$  estándar que proporciona el paquete estadístico). Estos resultados parecen ser muy similares a los resultados finales, considerando la diferencia en las medias entre los participantes y los no participantes, lo que sugiere que PROSCOL no está produciendo ningún impacto en la escolaridad. Sin embargo, cuando se incluyen diversas variables de control en la regresión, hay un coeficiente positivo y significativo en la participación en PROSCOL. El cálculo muestra que a la edad de 18 años, el programa ha agregado dos años a la escolaridad.

El analista se pregunta, ¿por qué estas variables de control afectan tanto? Además, ¿se están usando las variables de control correctas? Luego visita a su ex profesor de econometría y le muestra sus regresiones. El primer aspecto relacionado con la regresión de escolaridad en  $P$  y  $X$  que plantea el profesor es que ésta no permite que el impacto del programa varíe con  $X$ ;

el impacto es igual para todo el mundo, lo que no parece ser muy probable. Es más probable que los padres con mayor escolaridad envíen a sus hijos a la escuela, de modo que para ellos los beneficios de PROSCOL serán menores. Para permitir que los beneficios varíen con  $X$ , sea la escolaridad media de los no participantes  $a_0 + c_0X_i$  y la de los participantes,  $a_1 + c_1X_i$ , de modo que el nivel observado de escolaridad es

$$S_i = (a_1 + c_1X_i + \varepsilon_{1i})P_i + (a_0 + c_0X_i + \varepsilon_{0i})(1 - P_i),$$

donde  $\varepsilon_0$  y  $\varepsilon_1$  son errores aleatorios, cada uno con medias de cero y no correlacionados con  $X$ . Para estimar este modelo, a la regresión ya ejecutada se debe agregar un término adicional para los efectos de interacción entre la participación en el programa y las características observadas. Por lo tanto, la regresión ampliada es

$$S_i = a_0 + (a_1 - a_0)P_i + c_0X_i + (c_1 - c_0)P_i X_i + \varepsilon_i,$$

donde  $\varepsilon_i = \varepsilon_{1i}P_i + \varepsilon_{0i}(1 - P_i)$ . Luego  $(a_1 - a_0) + (c_1 - c_0)X$  es el impacto medio del programa en cualquier valor dado de  $X$ . Si se usa la media  $X$  en la muestra de participantes, ésta arrojará la ganancia media del programa.

### Comprensión de la importancia de la exogeneidad

Otro punto que planteó el profesor de econometría es la forma en que se estimó la regresión. Al usar el comando de regresión en el paquete estadístico, los mínimos cuadrados ordinarios (MCO), hay un problema porque las estimaciones de MCO de los parámetros estarán sesgadas incluso en muestras grandes, a menos que las variables del lado derecho sean exógenas. Exogeneidad significa que las variables del lado derecho estén determinadas en forma independiente de las opciones de escolaridad y que de este modo no estén correlacionadas con el término de error en la regresión de escolaridad. Dado que la participación en el programa se orientó intencionalmente hacia grupos específicos, la participación en PROSCOL no es exógena. Esto puede afectar el cálculo del impacto del programa de la siguiente forma: La ecuación de los años de escolaridad es

$$S_i = a + bP_i + cX_i + \varepsilon_i.$$

Se empleó el valor de  $a + b + cX_i + \varepsilon_i$  como la estimación de la escolaridad del  $i$ -ésimo hogar cuando participa en PROSCOL, mientras que se usó  $a + cX_i + \varepsilon_i$  para estimar la escolaridad si éste no participa. Así, la diferencia,  $b$ , es la ganancia que arroja el programa. Sin embargo, al efectuar este cálculo, el supuesto implícito es que  $\varepsilon_i$  era igual en cualquier caso. En otras palabras, el supuesto era que  $\varepsilon_i$  era independiente de  $P$ , lo que afectaría el cálculo del impacto del programa.

Esto destaca el sesgo que genera la asignación no aleatoria del programa, lo que también podría estar afectando la estimación basada en el modelo de regresión sugerido anteriormente por el economista ( $S_i = a + bP_i + cX_i + \varepsilon_i$ ). Sin embargo, esto no significa que los resultados sean completamente incorrectos.

El profesor de econometría aclara esto con una ecuación explícita para  $P$ , a saber,

$$P_i = d + eZ_i + v_i,$$

donde  $Z$  representa diversas variables que incluyen todos los “sustitutos de pobreza” usados para la focalización de PROSCOL. También habrá algún término de error puramente aleatorio que influye en la participación; éstos son sustitutos de la pobreza no representados en los datos; y también habrá errores en la selección de participantes que se incluyen en este término  $v$ . Esta ecuación es lineal, aunque  $P$  sólo puede tomar dos valores posibles, 0 y 1. Los valores predichos entre cero y uno son aceptables, pero un modelo lineal no puede eliminar la posibilidad de valores predichos negativos o de valores superiores a uno. Existen modelos no lineales que pueden tratar este problema, pero para simplificar el análisis, será más fácil confinar la atención en los modelos lineales.

Existe un caso especial en que la anterior regresión MCO de  $S$  sobre  $P$  y  $X$  dará una estimación insesgada de  $b$ . Esto es cuando  $X$  incluye todas las variables en  $Z$  que también influyen en la escolaridad, y el término de error  $v$  no está correlacionado con el término de error  $\varepsilon$  en la regresión para la escolaridad. En la literatura sobre evaluación, esto a veces se denomina “selección sobre observables”.

Supóngase que las variables de control  $X$  en la regresión anterior para la escolaridad incluyen todas las variables observadas  $Z$  que influyen en la participación  $P$  y  $v$  no está correlacionado con  $\varepsilon$  (de modo que las variables no observadas que afectan la asignación del programa no influyen en la escolaridad, dependiendo de  $X$ ). Luego esto ha eliminado cualquier posibilidad de que  $P$  esté correlacionada con  $\varepsilon$ . Ésta ahora será exógena en la regresión para la escolaridad. En otras palabras, la idea clave de la selección sobre observables es que hay algún  $X$  observable, de modo que el sesgo desaparece dependiendo de  $X$ .

La adición de las variables de control a la regresión de escolaridad sobre la participación en PROSCOL produjo una gran diferencia, porque  $X$  debe incluir variables que estaban entre los sustitutos de pobreza usados para la focalización, o que estaban correlacionados con éstos, y son variables que también influyen en la escolaridad. Sin embargo, esto sólo funciona si los supuestos son válidos. Se deben tener en cuenta dos problemas. En primer lugar, el método anterior falla si no hay determinantes no observadas de la participación, es decir, si el término de error  $v$  tienen varianza cero y todas las determinantes de la participación también afectan la escolaridad. Entonces no hay una variación independiente en la participación en el programa que permita identificar su impacto en la escolaridad; es perfectamente posible predecir  $P$  a partir de  $X$  y entonces no estimará la regresión. Es poco probable que este problema surja a menudo, dado que casi siempre existen determinantes no observadas de la asignación de un programa.

El segundo problema es más común y en este caso, más preocupante. El término de error  $\varepsilon$  en la regresión de la escolaridad probablemente contiene variables que no se encuentran en la encuesta, pero bien podrían influir en la participación en el programa, es decir, podrían estar correlacionadas con el término  $v$  en la ecuación de participación. Si esto sucede, entonces  $E(\varepsilon|X,P) \neq 0$  y los métodos de regresión ordinaria todavía estarán sesgados cuando

se estimen las regresiones para la escolaridad. Así, el punto clave es la magnitud de la correlación entre el término de error de la ecuación para la participación y el correspondiente a la ecuación de la escolaridad.

### **Búsqueda de mejores formas para crear un grupo de comparación: correspondencia de puntuación de la propensión**

Contando con más información del profesor, el analista aprende que existen mejores formas de crear un grupo de comparación. El objetivo es comparar los niveles de escolaridad según las características observadas. Si los grupos de la muestra se dividen en grupos de familias con valores iguales o similares de  $X$ , se comparará la media condicional para las familias dentro de PROSCOL y las familias fuera de PROSCOL. Si, en ausencia del programa, la escolaridad es independiente de la participación, dado  $X$ , la comparación generará una estimación insesgada del impacto de PROSCOL. Esto a veces se denomina “independencia condicional” y es el supuesto clave formulado por todos los métodos de grupos de comparación.

Así, una mejor forma de seleccionar un grupo de comparación, dados los datos existentes, es usar, para cada participante, un no participante con las mismas características observadas como control. Sin embargo, esto podría resultar muy difícil, porque el conjunto de datos podría tener muchas de esas variables. Podría no haber nadie entre los no participantes con exactamente los mismos valores de todas las características observadas para cada uno de los participantes en PROSCOL.

Un enfoque estadístico de correspondencia de puntuación de la propensión proporciona técnicas para simplificar enormemente el problema. En lugar de intentar asegurar que el control pareado para cada participante tenga exactamente el mismo valor de  $X$ , se puede lograr el mismo resultado realizando una correspondencia sobre el valor predicho de  $P$ , dado  $X$ , que se denomina la puntuación de la propensión de  $X$ . Rosenbaum y Rubin (1983) demuestran que si (en este caso) la escolaridad sin PROSCOL es independiente de la participación dado  $X$ , los participantes también son independientes de la participación dada la puntuación de la propensión de  $X$ . Puesto que la puntuación de la propensión es sólo un número, es mucho más fácil controlarla que  $X$ , que podría ser diversas variables. Y sin embargo, la correspondencia de puntuación de la propensión es suficiente como para eliminar el sesgo, siempre que haya independencia condicional dado  $X$ .

En otras palabras, primero se ejecuta una regresión de  $P$  sobre  $X$  para obtener el valor predicho de  $P$  para cada valor posible de  $X$ , que luego se estima para toda la muestra. Para cada participante, se debe encontrar el no participante con el valor más cercano a esta probabilidad predicha. Luego la diferencia en la escolaridad es la ganancia estimada del programa para ese participante.

Se puede sacar la media de todas estas diferencias para estimar el impacto. O bien, sacar la media para diferentes grupos de ingreso. Sin embargo, la forma en que se estima el modelo de la participación requiere cierta prudencia. Un modelo lineal podría proporcionar probabilidades predichas irregulares, mayor a uno, o negativas. Es mejor usar el comando

LOGIT del paquete estadístico. Este supone que el término de error  $v$  de la ecuación de participación tiene una distribución logística y estima los parámetros en forma consecutiva con el supuesto de los métodos de probabilidad máxima. Esto se basa en los principios de la estimación de probabilidad máxima de los modelos de respuesta binaria.

Otro aspecto que hay que tener en cuenta es que desde el principio se deben excluir algunos de los no participantes como posibles pares. En efecto, algunos resultados recientes dentro de la literatura sobre econometría indican que una fuente importante de sesgo en las evaluaciones es el no comparar los participantes y los controles a valores comunes de variables de pareo (véase Heckman y otros, 1998).

La intuición indica que el grupo de comparación debe ser lo más similar posible al grupo de tratamiento en términos de los elementos observables, como lo resume la puntuación de la propensión. Se podría concluir que parte de la muestra de no participantes tiene una puntuación de la propensión más baja que cualquiera de los correspondientes a la muestra de tratamiento. Esto a veces se denomina “falta de soporte común”. Al crear el grupo de comparación, se deben eliminar esas observaciones del conjunto de no participantes, para garantizar que sólo se compararán las ganancias dentro del mismo margen de puntuaciones de la propensión. También se deben excluir los no participantes para los cuales la probabilidad de participar es cero. Se recomienda reducir una pequeña proporción de la muestra, digamos un 2%, del extremo superior e inferior de la distribución de no participantes en términos de puntuaciones de la propensión. Una vez identificados los participantes y asimismo los no participantes en una región de correspondencia común, se recomienda sacar un promedio de (por ejemplo) los cinco o más cercanos en términos de diferencia absoluta en puntuaciones de la propensión (Recuadro 3.1).

### **Recuadro 3.1 Pasos en la correspondencia de puntuación de la propensión**

El objetivo de la correspondencia es encontrar el grupo de comparación más cercano de una muestra de no participantes a la muestra de participantes en el programa. Lo “más cercano” se mide en términos de características observables. Si sólo hay una o dos de esas características, la correspondencia será más fácil. Pero es habitual que existan muchas eventuales características. Los principales pasos para la correspondencia basada en las puntuaciones de la propensión son las siguientes:

**Paso 1:** Se debe contar con una encuesta de una muestra representativa de no participantes calificados y una para los participantes. Mientras mayor sea la muestra de no participantes calificados, mejor, puesto que facilitará una buena correspondencia. Si las dos muestras provienen de diferentes encuestas, deberían ser encuestas muy comparables (mismo cuestionario, mismos entrevistadores o capacitación de los entrevistadores, mismo período de encuesta, etc.).

**Paso 2:** Reunir las dos muestras y estimar un modelo logit de la participación en el programa como función de todas las variables en los datos que probablemente determinen la participación.

**Paso 3:** Crear los valores predichos de la probabilidad de participar a partir de la regresión logit; éstos se denominan “puntuaciones de la propensión”. Se obtendrá una puntuación de la propensión para cada participante y no participante de la muestra.

**Paso 4:** Es posible que desde el principio se deba excluir a algunos de la muestra de no participantes, por tener una puntuación de la propensión fuera del margen (en general demasiado baja) con respecto al de la muestra de tratamiento. El margen de puntuaciones de la propensión estimado para el grupo de tratamiento debe ser muy parecido al de la submuestra conservada para los no participantes. Se podría tener que restringir las posibles correspondencias en otras formas, según el escenario. Por ejemplo, es posible que se decida permitir sólo correspondencias con la misma área geográfica, para asegurar que las correspondencias provengan del mismo entorno económico.

**Paso 5:** Para cada individuo de la muestra de tratamiento, ahora se necesita encontrar la observación en la muestra de no participantes que se acerque más en su puntuación de la propensión, medida por la diferencia absoluta en puntuaciones. Esto se denomina el “vecino más cercano”. Se pueden encontrar, por ejemplo, los cinco vecinos más cercanos.

**Paso 6:** Calcular el valor de la media del indicador de resultados (o cada uno de los indicadores, si hay más de uno) para los cinco vecinos más cercanos. La diferencia entre esa media y el valor real para la observación tratada es la estimación de la ganancia debido al programa para esa observación.

**Paso 7:** Calcular la media de las ganancias de estos individuos para obtener la ganancia promedio total. Esto se puede estratificar por medio de alguna variable de interés, como el ingreso, en la muestra de no participantes.

Esta es la forma más simple de correspondencia de puntuación de la propensión. En la práctica pueden surgir complicaciones. Por ejemplo, si se realiza un segundo muestreo de los participantes, se pueden emplear métodos de muestreo basados en la selección para corregir por esto (Manski y Lerman, 1977); como alternativa, se puede utilizar la razón de disparidad ( $p/(1 - p)$ , donde  $p$  es la puntuación de la propensión) para la correspondencia. En lugar de depender del vecino más cercano, se podrían utilizar todos los no participantes como posibles correspondencias, pero ponderarlos de manera diferente, según su cercanía (Heckman y otros, 1998).

Luego se deben incluir todas las variables del conjunto de datos que son o podrían ser sustitutos de los indicadores de pobreza utilizados al seleccionar los participantes en PROSCOL. Nuevamente,  $X$  debe incluir las variables en  $Z$ . Sin embargo, esto pone en evidencia una debilidad de la correspondencia de puntuación de la propensión. Con la correspondencia, un  $X$  diferente generará una estimación diferente del impacto. Con la aleatorización, el experimento ideal, los resultados no dependen de las variables  $X$  que se seleccionen. La aleatorización tampoco requiere especificar un modelo para la participación, ya sea un modelo logit u otro. En el Recuadro 3.1 se resumen los pasos para realizar la correspondencia de puntuación de la propensión.

### **Adquisición de conocimientos sobre sesgos debido a elementos no observables**

Incluso después de formar el grupo de comparación, el analista no puede estar seguro de que esto le brinde una estimación mucho mejor del impacto del programa. Los métodos antes descritos sólo eliminarán el sesgo si existe independencia condicional, de modo que las determinantes no observables de la escolaridad (no incluidas en el conjunto de variables de control  $X$ ) no estén correlacionadas con la asignación del programa. Hay dos fuentes distintas de sesgo, aquella causada por diferencias en los elementos observables y aquella causada por diferencias en los elementos no observables; estas últimas a menudo se denominan “sesgo de selección”. En el Recuadro 3.2 se explica en detalle esta diferencia.

Regresando a la última ecuación del profesor, ésta muestra que se mantendrá la independencia condicional si  $P$  es exógeno, para que entonces  $E(\varepsilon_i | X_i, P_i) = 0$ . Sin embargo, la asignación endógena del programa debido a la focalización intencional basada en elementos no observables todavía dejará un sesgo. Esto a veces se denomina selección sobre observables. Así, las condiciones requeridas para justificar el método planteado anteriormente por el economista no son menos restrictivas que las necesarias para justificar una versión del primer método basado en comparar familias dentro de PROSCOL con familias fuera de éste para hogares con valores similares de  $X$ . Ambos se basan en la convicción de que estos elementos no observables no influyen en conjunto en la escolaridad y la participación el programa, dado  $X$ .

En forma intuitiva, se podría pensar que una cuidadosa correspondencia reduce el sesgo, pero no necesariamente es así. La correspondencia elimina parte del sesgo en la primera estimación simplista del impacto de PROSCOL. Pero se mantiene el sesgo generado por cualquier elemento no observable difícil. Sin embargo, estas dos fuentes de sesgo se podrían compensar: una podría ser positiva y la otra negativa. Heckman y otros (1998) hacen ver esto. Así, la estimación de correspondencia bien podría tener más sesgo que la estimación simplista. No se puede saber a priori qué tanto favorece incluso un grupo de comparación bien elegido, ya que se trata de un aspecto empírico.

## **Revisión de lo que se podría haber hecho con una encuesta básica: Estimaciones de doble diferencia**

Luego el analista consulta si habría otro método aparte de la aleatorización que sea sólido para estos elementos no observables dificultosos. Esto requeriría datos básicos o de referencia para los participantes y los no participantes recopilados antes de que se implementara el programa PROSCOL. La idea es recopilar datos sobre los resultados y sus determinantes antes y después de introducir el programa, junto con datos de un grupo de comparación sin tratamiento y del grupo de tratamiento. Luego basta con sólo restar la diferencia entre la escolaridad de los participantes y el grupo de comparación antes de introducir el programa, de la diferencia después del programa.

### **Recuadro 3.2 Fuentes de sesgo en estimaciones simplistas del impacto de PROSCOL**

El sesgo descrito por el estadístico es la diferencia esperada en la escolaridad sin PROSCOL entre las familias seleccionadas para el programa y aquellas que no fueron elegidas. Se puede desglosar en dos fuentes de sesgo:

- Sesgo debido a diferencias en las características observables. Esto se puede originar de dos formas. En primer lugar, podría no haber soporte común. El “soporte” es el conjunto de valores de las variables de control para las que se observan los resultados y la participación en el programa. Si el soporte difiere entre la muestra de tratamiento y el grupo de comparación, esto sesgará los resultados. En efecto, no se están comparando elementos semejantes. En segundo lugar, incluso con un soporte común, la distribución de las características observables podría ser diferente dentro del área de soporte común; efectivamente, los datos del grupo de comparación están ponderados en forma incorrecta. Una selección cuidadosa del grupo de comparación puede eliminar esta fuente de sesgo.
- Sesgo debido a diferencias en las características no observables. El término sesgo de selección a veces se limita únicamente a este componente (aunque algunos autores usan ese término para el sesgo total en una evaluación no experimental). Esta fuente de sesgo surge cuando, para valores dados de  $X$ , existe una relación sistemática entre la participación en el programa y los resultados en ausencia del programa. En otras palabras, hay variables no observables que en conjunto influyen en la escolaridad y la participación en el programa dadas las variables observadas en los datos.

No hay nada que garantice que estas dos fuentes de sesgo funcionarán en la misma dirección. De modo que eliminar alguna de ellas no significa que se reducirá el sesgo total en el valor absoluto. Esa es una pregunta empírica. En uno de los pocos estudios que abordan esta pregunta, el verdadero impacto, medido por un experimento bien diseñado, se comparó con diversas estimaciones no experimentales (Heckman y otros, 1998). El sesgo en la estimación simplista fue enorme, pero una cuidadosa correspondencia del grupo de comparación basada en los observables redujo en forma considerable el sesgo.

Esto se llama la estimación de “doble diferencia” o “diferencia en las diferencias”. Abordará las variables no observadas dificultosas, siempre que no varíen en el tiempo.

Se puede explicar agregando subíndices a la ecuación anterior, de modo que se introduce la escolaridad después del programa:

$$S_{ia} = a + bP_i + cX_{ia} + \varepsilon_{ia} .$$

Antes del programa, en la encuesta básica, el rendimiento escolar en su lugar es

$$S_{ib} = a + c X_{ib} + \varepsilon_{ib} .$$

(Por cierto,  $P = 0$  antes de introducir el programa). Los términos de error incluyen un impacto adicional constante en el tiempo, de modo que se pueden representar como

$$\varepsilon_{it} = \eta_i + \mu_{it} \text{ (para } t = a, b),$$

donde  $\eta_i$  es el impacto constante en el tiempo, que se permite que esté correlacionado con  $P_i$ , y  $\mu_{it}$  es un error de innovación, que no está correlacionado con  $P_i$  (o  $X_i$ ).

Aquí la idea esencial es usar los datos básicos para revelar los elementos no observables problemáticos. Cabe notar que puesto que la encuesta básica corresponde a los mismos hogares que se tienen ahora, el  $i$ -ésimo hogar de la ecuación para  $S_{ia}$  es el mismo que el  $i$ -ésimo hogar de la ecuación para  $S_{ib}$ . Luego se puede sacar la diferencia entre la ecuación “después” y la ecuación “antes”:

$$S_{ia} - S_{ib} = bP_i + c(X_{ia} - X_{ib}) + \mu_{ia} - \mu_{ib}$$

Ahora es posible hacer una regresión del cambio en la escolaridad y la participación en el programa y los cambios en  $X$ . MCO proporcionará una estimación insesgada del impacto del programa. Se han eliminado los elementos no observables, aquellos que están correlacionados con la participación en el programa.

Dado esto, si la asignación del programa se basara sólo en variables (tanto observadas como no observadas) que se conocieran en el momento de la encuesta básica, sería razonable suponer que los  $\eta$  no cambian entre las dos encuestas. Esto se mantendría siempre que los no observables problemáticos fueran constantes en el tiempo. Los cambios en la escolaridad en el tiempo para el grupo de comparación revelarán lo que habría sucedido al grupo de tratamiento sin el programa.

Para ello se necesitaría conocer bien el programa y poder sincronizar las encuestas de evaluación, de modo de coordinarlas con el programa. De lo contrario, inevitablemente serían los cambios no observados después de la encuesta básica los que influyen en quién se ve favorecido con el programa. Esto crearía ceros que cambiarían entre las dos encuestas.

Esta última ecuación se puede interpretar como que las características del niño y del hogar en  $X$  no son pertinentes para el cambio en la escolaridad si estas características no varían en el tiempo. Pero el beneficio de la escolaridad podría depender de la educación de los padres (y no sólo de cualquier cambio en su educación) y posiblemente del lugar donde se encuentra el hogar, porque esto determinará el acceso a las escuelas. También puede haber situaciones en las cuales los cambios en el tiempo en el indicador de resultados están influidos por las condiciones iniciales. Luego sólo se necesitará controlar por las diferencias en las condiciones iniciales. Esto se puede hacer agregando simplemente  $X_a$  y  $X_b$  por separado a la regresión, de modo que la regresión adopte la forma.

$$S_{ia} - S_{ib} = b P_i + c_a X_{ia} + c_b X_{ib} + \mu_{ia} - \mu_{ib}.$$

Incluso si algunas (o todas) las variables de  $X$  no varían en el tiempo, aún se puede permitir que  $X$  afecte los cambios en el tiempo en la escolaridad.

El método de correspondencia de puntuación de la propensión antes analizado puede ayudar a asegurar que el grupo de comparación sea similar al grupo de tratamiento antes de realizar la doble diferencia. En un interesante estudio de un programa de empleos estadounidense, se concluyó que el no asegurar que las comparaciones se hubieran efectuado en un área de apoyo común fue una fuente importante de sesgo en la estimación de doble diferencia en comparación con un grupo de control aleatorizado. Sin embargo, dentro del área de apoyo común, el sesgo dependiente de  $X$  no varió mucho en el tiempo. Por lo tanto, es razonable tomar la doble diferencia después de realizar la correspondencia (véase Heckman y otros (1998)).

En la práctica, sin embargo, puede ser difícil hacer un seguimiento de los hogares de las encuestas. Podría no ser fácil encontrar a todos los hogares que se incluyeron originalmente en la encuesta básica. Algunas personas de la encuesta básica podrían no desear ser entrevistadas de nuevo o podrían haberse trasladado a una ubicación desconocida.

Si los retiros de la muestra son exclusivamente aleatorios, la encuesta de seguimiento todavía será representativa de la misma población de la encuesta básica. Sin embargo, si existe alguna tendencia sistemática de las personas con ciertas características a abandonar la muestra, esto sí representaría un problema. Esto se llama “sesgo de desgaste o de disminución normal”. Por ejemplo, PROSCOL podría ayudar a algunas familias pobres a trasladarse a una vivienda mejor. Además, incluso cuando la selección de los participantes sólo se base en información disponible en la fecha básica de referencia o alrededor de ella (el efecto  $0_i$  constante en el tiempo), los participantes seleccionados bien podrían retirarse en forma voluntaria sobre la base de cambios después de esa fecha. Ese desgaste del grupo de tratamiento claramente sesgará una estimación de doble diferencia del impacto del programa. En el Recuadro 3.3 se resumen los pasos para realizar una estimación de doble diferencia.

### Recuadro 3.3. Ejecutando una doble diferencia

El método de doble diferencia supone comparar un grupo de tratamiento con un grupo de comparación (que idealmente se determinaría por medio del método de correspondencia del Recuadro 3.2) antes y después de la intervención. Los pasos principales son los siguientes:

**Paso 1:** Se necesita una encuesta básica antes de establecer la intervención y la encuesta debe cubrir tanto a participantes como a no participantes. Si no se sabe quién participará, se debe realizar una estimación bien informada. Se deberá conversar con los administradores del programa.

**Paso 2:** Luego se requerirán una o más encuestas de seguimiento después del establecimiento del programa. Éstas deben ser muy comparables a las encuestas básicas (en términos del cuestionario, la entrevista, etc.). Idealmente, las encuestas de seguimiento deben incluir las mismas observaciones muestreadas que la encuesta básica. Si no es posible, deben corresponder a los mismos grupos geográficos o estratos con respecto a alguna otra variable.

**Paso 3:** Hay que calcular la diferencia media entre los valores después y antes del indicador de resultados para cada uno de los grupos de tratamiento y comparación.

**Paso 4:** Se debe calcular la diferencia entre estas dos diferencias medias. Esa es la estimación del impacto del programa.

Esta es la versión más simple de la doble diferencia. También se podría necesitar controlar por las diferencias en las condiciones iniciales exógenas o los cambios en las variables exógenas, posiblemente dando cabida a los efectos de interacción con el programa (de modo que la ganancia producida por la intervención del programa sea alguna función de las variables observables). Un modelo de regresión adecuado puede admitir estas variaciones.

### Uso de variables instrumentales

Dado que no hay encuesta básica de los mismos hogares para emplear el método de doble diferencia, el profesor recomienda otra metodología para obtener una estimación que sea sólida para los elementos no observables problemáticos: una “variable instrumental”.

Una variable instrumental es la clásica solución para el problema de un regresor endógeno. Una variable instrumental es una fuente observable de variación exógena de la participación en el programa. En otras palabras, está correlacionada con  $P$ , pero aún no está en la

regresión para la escolaridad y no está correlacionada con el término de error en la ecuación de escolaridad,  $\varepsilon$ . De modo que se debe tener al menos una variable en  $Z$  que no esté en  $X$  y que no esté correlacionada con  $\varepsilon$ . Luego, la estimación de variables instrumentales del impacto del programa se obtiene reemplazando  $P$  por su valor predicho dependiente de  $Z$ . Puesto que este valor predicho depende sólo de  $Z$  (que es exógeno) y  $Z$  está correlacionado con  $\varepsilon$ , ahora es razonable aplicar los MCO a esta nueva regresión.

Puesto que los valores predichos sólo dependen de la variación exógena debido a la variable instrumental y a las demás variables exógenas, los no observables ya no son problemáticos porque estarán correlacionados con el término de error en la regresión de escolaridad. Esto también sugiere otra forma más eficiente de tratar el problema. Cabe recordar que la fuente de sesgo en la estimación anterior del impacto del programa era la correlación entre el término de error de la ecuación de escolaridad y el correspondiente a la ecuación de participación. Esto es lo que crea la correlación entre la participación y el término de error en la ecuación de la escolaridad. Así, una forma natural de deshacerse del problema cuando se tiene una variable instrumental es agregar los valores residuales de la ecuación de participación de la primera etapa a la ecuación de escolaridad, pero manteniendo la participación real en la regresión de escolaridad. Sin embargo, puesto que ahora se ha agregado el valor estimado del término de error de la ecuación de participación a la regresión de escolaridad, es posible tratar la participación como exógena y ejecutar los MCO. Esto sólo funciona si hay un instrumento válido. Si no lo hay, la regresión no realizará la estimación, porque en un modelo lineal el valor residual de la participación será perfectamente predecible a partir de la participación real y  $X$ .

Una variable instrumental también puede ayudar si hay un error de medición apreciable en los datos de participación en el programa, otra posible fuente de sesgo. Un error de medición significa que existe la posibilidad de que la participación en el programa varíe más de lo que realmente hace. Naturalmente, esta sobrestimación en la variación de  $P$  produce una subestimación de su coeficiente  $b$ . Esto se denomina sesgo de atenuación, porque este sesgo atenúa el coeficiente de regresión estimado.

Aunque una variable instrumental puede ser extremadamente útil, en la práctica es necesario ser prudente. Cuando la participación real sólo se sustituye por su valor predicho y se ejecutan los MCO, esto no entregará los errores estándar correctos: la computadora no sabrá que se deben usar parámetros previamente estimados para obtener los valores predichos. Se requiere una corrección a los errores estándar de MCO, aunque existen paquetes estadísticos que permiten hacerlo fácilmente, al menos para los modelos lineales.

Sin embargo, si hubiera una variable dependiente que sólo pudiera adoptar dos posibles valores, por ejemplo, en la escuela o no en la escuela, se debería usar un modelo de respuesta binario no lineal, como logit o probit. El principio para probar la exogeneidad de la participación en el programa es similar en este caso. Existe un estudio por Rivers y Vuong (1988) en que se analiza el problema para estos modelos; en Blundell y Smith (1993) se proporciona una visión general útil de diversos modelos no lineales en los cuales existe un regresor endógeno.

## Pruebas de las metodologías

Cuando el analista comienza a pensar sobre la identificación de una variable instrumental, comprende que no es un proceso directo. Todas las posibilidades en que ha pensado también se podrían incluir junto a las variables en  $X$ . El problema es encontrar una “restricción de exclusión” válida que justifique colocar alguna variable en la ecuación de participación, pero no en la ecuación de escolaridad.

El analista decide intentar el método de correspondencia de puntuación de la propensión. El modelo logit de participación parece bastante sensible y sugiere que PROSCOL está bien focalizado. Casi todas las variables de las que se podría esperar una asociación con la pobreza tienen coeficientes positivos y significativos. Luego el analista realiza la correspondencia de puntuación de la propensión. En una comparación de las tasas de matrícula escolar medias, los resultados muestran que los niños del grupo de comparación pareado tenían una tasa de matrícula del 60%, en comparación con el 80% para las familias del PROSCOL.

Para responder al problema del ingreso no percibido, el análisis recurre a una encuesta existente del trabajo infantil en que se consultaba sobre los ingresos (en este país en desarrollo, existe una prohibición oficial de que los niños menores de 16 años trabajen, pero al gobierno le ha resultado difícil hacerla cumplir; no obstante, los salarios infantiles son un problema delicado). Según esta encuesta, se pueden determinar los ingresos que habría recibido un niño si no hubiera asistido a la escuela.

Luego es posible restar el monto del ingreso no percibido del pago en efectivo de PROSCOL a los participantes y calcular así la transferencia neta de ingresos. Al restar esta transferencia neta del ingreso total, será posible calcular de dónde provienen los participantes en PROSCOL en la distribución de ingresos previa a la intervención. No son tan pobres como se pensaba al principio (si se ignora el ingreso no percibido), pero siguen siendo pobres; por ejemplo, dos tercios de ellos están bajo la línea de pobreza oficial del país.

Habiendo calculado la ganancia de ingreso neta de todos los participantes, ahora es posible calcular la tasa de pobreza con y sin PROSCOL. En términos simples, la tasa de pobreza posterior a la intervención (con el programa) es la proporción de la población que vive en hogares con un ingreso por persona bajo la línea de pobreza, donde “ingreso” es el ingreso observado (incluidos los ingresos de transferencia bruta de PROSCOL). Esto se puede calcular directamente a partir de la encuesta domiciliaria. Al sustraer la ganancia de ingreso neta (la transferencia de efectivo de PROSCOL menos los ingresos no percibidos del trabajo infantil) atribuido a PROSCOL de todos los ingresos observados, los resultados muestran una nueva distribución de los ingresos anteriores a la intervención. Entonces, la tasa de pobreza sin el programa es la proporción de personas que viven en hogares pobres, según esta nueva distribución. El análisis concluye que la tasa de pobreza observada del 32% en el noroeste habría sido un 36% si no hubiera existido PROSCOL. El programa permite que ahora un 4% de la población se libere de la pobreza. Las ganancias en escolaridad significan que también habrá ventajas pecuniarias y no pecuniarias para los

pobres en el futuro. En el proceso de medición de la pobreza, el analista recuerda haber aprendido que la proporción de personas bajo la línea de pobreza sólo es una medida básica, porque no indica nada sobre los cambios bajo la línea (véase el Recuadro 3.4). Luego calcula el índice de la brecha de pobreza y el índice de la brecha de pobreza al cuadrado y los resultados sugieren que éstos también han descendido como resultado del PROSCOL.

### **Recuadro 3.4 Mediciones de pobreza**

La medida más simple y más común de la pobreza es el índice de recuento de la pobreza. En este caso, es la proporción de la población que vive en hogares con un ingreso por persona bajo la línea de pobreza (en otros países es una medida basada en el consumo, lo que tiene ciertas ventajas; para un análisis y referencias, véase Ravallion, 1994).

El índice de recuento de la pobreza no indica nada sobre la distribución del ingreso bajo la línea de pobreza: una persona pobre puede quedar en peor situación, pero el índice de recuento de la pobreza no variará ni reflejará los progresos entre los pobres, a menos que crucen la línea de pobreza.

Una alternativa para el índice de recuento de la pobreza, utilizada en forma generalizada, es el índice de brecha de pobreza (PG). La brecha de pobreza para cada hogar es la diferencia entre la línea de pobreza y el ingreso del hogar; para aquellos sobre la línea de pobreza, la brecha es cero. Cuando la brecha de pobreza se normaliza por línea de pobreza y se calcula su media en todos los hogares (pobres o no), se obtiene el índice de brecha de pobreza.

El índice de brecha de pobreza indicará el nivel en que el programa ha influido en la profundidad de la pobreza, pero no reflejará ningún cambio en la distribución entre los pobres causado por el programa. Por ejemplo, si el programa supone un pequeño avance para una persona pobre que se encuentra sobre el ingreso medio de los pobres, a expensas de una pérdida equivalente de alguien bajo esa media, el índice de brecha de pobreza no variará.

Existen diversas mediciones “sensibles a la distribución” que reflejarán esos cambios en la distribución entre los pobres. Una de estas mediciones es la “brecha de pobreza al cuadrado” (Foster, Greer y Thorbecke, 1984). Se calcula de la misma forma que la brecha de pobreza, excepto que cada brecha de pobreza como proporción de la línea de pobreza se eleva al cuadrado antes de sacar la media (nuevamente entre pobres y no pobres). Otro ejemplo de una medición de pobreza sensible a la distribución es el índice Watts. Es la media del logaritmo de la razón entre la línea de pobreza y el ingreso, donde la razón se fija en uno para los no pobres. Atkinson (1987) describe otros ejemplos en sus estudios.

En este cálculo, el analista también reconoce que hay cierta incertidumbre acerca de la línea de pobreza del país. Para probar los resultados, repite el cálculo sobre una amplia variedad de líneas de pobreza y concluye que en una línea de pobreza para la cual el 50% de la población es pobre (según los ingresos observados después de la intervención), la proporción habría sido del 52% sin PROSCOL. En una línea de pobreza en que el 15% no es cubierto por el programa, la proporción habrá sido del 19% sin éste. Mediante la repetición de estos cálculos en toda la gama de ingresos, se trazaron todas las “curvas de incidencia de la pobreza”, con el programa y sin él. Esto también se denomina “función de distribución acumulativa” (véase el Recuadro 3.5).

### **Recuadro 3.5 Comparación de la pobreza con el programa y sin él**

Con la ayuda de los métodos descritos en el texto principal y los recuadros anteriores, se obtiene una estimación de los beneficios para cada hogar. En las evaluaciones más simples es sólo un número. Pero es mejor dejar que varíe con las características del hogar. Luego se puede resumir esta información en forma de curvas de incidencia de la pobreza (PIC) con el programa y si él.

**Paso 1:** Ya debería existir el ingreso posterior a la intervención (u otro indicador de bienestar) para cada hogar de toda la muestra (que comprenda participantes y no participantes); éstos son datos. También se sabe el número de personas que vive en cada hogar. Y, por cierto, se conoce el número total de personas de la muestra ( $N$ ; o podría ser el tamaño estimado de la población, si se han usado tasas de muestreo inversas para “agotar” la observación de cada muestra).

**Paso 2:** Se puede representar esta información en forma de una curva PIC. Esto proporciona (en el eje vertical) el porcentaje de la población que vive en hogares con un ingreso menor o igual al valor del eje horizontal. Para trazar este gráfico, se puede comenzar con el hogar más pobre, marcar su ingreso en el eje horizontal y luego contar en el eje vertical 100 veces el número de personas de ese hogar dividido por  $N$ . El siguiente punto es la proporción que vive en los dos hogares más pobres, y así sucesivamente. Con esto se obtiene la PIC posterior a la intervención.

**Paso 3:** Ahora hay que calcular la distribución de ingreso previo a la intervención. Para obtenerla, se resta la ganancia estimada para cada hogar de su ingreso posterior a la intervención. Luego hay una lista de ingresos posteriores a la intervención, uno para cada hogar de la muestra. Repetir entonces el Paso 2. Se obtiene la PIC previa a la intervención.

Si se considera cualquier nivel de ingreso dado del eje horizontal como una línea de pobreza, la diferencia entre las dos PIC en ese punto da el impacto en el índice de recuento de la pobreza para esa línea de pobreza (Recuadro 3.4). Alternativamente, al examinar el eje horizontal se obtiene el aumento de ingreso en ese percentil. Si ninguna de las ganancias es negativa, la PIC posterior a la intervención debe quedar bajo la respectiva curva antes de la intervención. La pobreza habrá disminuido, sin importar qué línea de pobreza se esté utilizando. Por cierto, esto también se mantiene para una amplia clase de medidas de pobreza; véase Atkinson (1987). Si algunas ganancias son negativas, las PIC se intersectarán. En ese caso, la comparación de la pobreza es ambigua; la respuesta dependerá de las líneas de pobreza y medidas de pobreza que se usen (un análisis más detallado se encuentra en Ravallion, 1994). Se podrían aplicar restricciones a priori a la gama de líneas de pobreza admisibles. Por ejemplo, se podría confiar en que la línea de pobreza no supere cierto valor máximo y si se produce la intersección sobre ese valor, la

comparación de la pobreza sería inequívoca. Si el punto de intersección (y podría haber más de uno) está bajo la línea de pobreza máxima admisible, sólo sería posible una comparación de pobreza sólida para un conjunto restringido de medidas de pobreza. Para verificar el grado de restricción que debe tener el conjunto, se pueden calcular las curvas de profundidad de la pobreza (PDC). Es posible obtenerlas formando simplemente la suma acumulativa hasta cada punto de la PIC (de modo que el segundo punto de la PDC sea el primer punto de la PIC más el segundo punto, y así sucesivamente).

Si las PDC no se intersectan, el impacto del programa en la pobreza es inequívoco, siempre y cuando la atención quede restringida al índice de brecha de pobreza o a cualquiera de las medidas de pobreza sensibles a la distribución descritas en el recuadro 3.4. Si las PDC se intersectan, se pueden calcular las “curvas de severidad de la pobreza” con el programa y sin él, formando las sumas acumulativas bajo las PDC. Si estas no se intersectan sobre la gama de líneas de pobreza admisibles, el impacto en cualquiera de las medidas de pobreza sensibles a la distribución del Recuadro 3.4 es inequívoco.

### **Incorporación de información del terreno**

En la implementación de cada programa, existen conocimientos de los beneficiarios y administradores del programa que podrían estar o no reflejados en los datos del programa. Por ejemplo, en este caso la percepción de los que trabajan en terreno es que la mayoría de las familias de PROSCOL son pobres y que el programa desde luego sirve de ayuda. Cuando el analista trata esto con un sociólogo que trabaja en el programa, percibe cierta incertidumbre en la realidad del ingreso no percibido y el tema del trabajo. El sociólogo comenta que en el terreno se observan muchos niños de familias pobres que trabajan y asisten a la escuela y que algunos de los niños más pequeños que no asisten a la escuela

parecen no estar trabajando. El analista comprende que esto requiere la verificación de si existe alguna diferencia en la cantidad de trabajo realizada por los niños dentro de PROSCOL con respecto al realizado por un grupo de comparación pareado. Sin embargo, estos datos no están disponibles en la encuesta domiciliaria, aunque sería posible presentar los resultados con la deducción del ingreso no percibido y sin ella.

El sociólogo también ha observado que para que una familia pobre ingrese a PROSCOL, tiene mucha importancia el área de la junta escolar (SBA, *school-board area*) donde ésta viva. Todas las SBA obtienen una asignación en PROSCOL desde el centro, incluso aquellas que tienen muy pocas familias pobres. Si se es pobre, pero se habita dentro de una SBA próspera, es más probable obtener ayuda de PROSCOL que si se vive en una SBA pobre. Por consiguiente, lo que realmente importa es la pobreza relativa —relativa a las demás familias del área en que se vive—, la que influye mucho más que el nivel de vida absoluto.

Esta asignación influiría en la participación en PROSCOL, pero no es dable esperar que tuviera importancia para la asistencia a la escuela. Esta última dependería más del nivel de vida absoluto, las circunstancias familiares y las características de la escuela. Así, la asignación presupuestaria de PROSCOL en las SBA se puede usar como variable instrumental para eliminar el sesgo en las estimaciones del impacto del programa.

Hay información acerca de la SBA a la que pertenece cada hogar en la encuesta domiciliaria, las normas aplicadas por el centro en la asignación de fondos de PROSCOL a las SBA y el monto que el centro ha asignado a cada SBA. Las asignaciones se basan en el número de niños en edad escolar, con un “factor de ajuste” correspondiente al grado de pobreza estimado para la SBA. Sin embargo, la norma es algo vaga.

El analista intenta tomar en cuenta estos puntos y vuelve a ejecutar la regresión para la escolaridad, pero reemplaza la participación real en PROSCOL por su valor predicho (la puntuación de la propensión) de la regresión para la participación, que ahora incluye la asignación presupuestaria a la SBA. Es útil tener de antemano tantas características como sea posible de las escuelas en la regresión correspondiente a la asistencia. Aunque las características de la escuela oficialmente parecen no influir en la forma en que se asignan los recursos de PROSCOL, cualquier característica escolar omitida que influya en forma simultánea en las asignaciones de PROSCOL por SBA y en cada resultado de escolaridad dejará un sesgo en las estimaciones de variables instrumentales del analista. Aunque siempre existe la posibilidad de sesgo, con abundantes variables de control geográfica este método debería ofrecer al menos un comparador creíble para la estimación de correspondencia.

De los resultados se desprende que la asignación de presupuesto a la SBA efectivamente tiene un coeficiente positivo significativo en la regresión logit para la participación en PROSCOL. Ahora la participación en PROSCOL (predicha) es significativa en una regresión para la matrícula escolar, en la cual se incluyen las mismas variables de la regresión logit, excepto la asignación de presupuesto de la SBA. El coeficiente implica que la tasa de matrícula es 15 puntos porcentuales más alta para los participantes en PROSCOL

de lo que habría sido de otro modo. También se ejecuta una regresión para los años de escolaridad, para niños y niñas por separado. Para niños y niñas de 18 años, los resultados indican que habrían abandonado la escuela casi dos años antes si no hubiera sido por PROSCOL. Sin embargo, esta regresión plantea preguntas sobre si se están utilizando los errores estándares correctos y si se deben usar modelos lineales.

### **Planificación del trabajo futuro**

Por último, el analista está preparado para presentar un informe con los resultados de las evaluaciones. Estos resultados indican que PROSCOL está funcionando bastante bien y debido a ello, las autoridades responsables se muestran interesadas en ampliar el programa. El analista tiene algunas observaciones importantes que hacer, las que han surgido del proceso por el que ha pasado para llevar a cabo la evaluación:

- La evaluación del impacto puede ser mucho más difícil de lo previsto al principio;
- Es posible obtener una gama de estimaciones, preocupante por su extensión, según las especificaciones de la metodología usada;
- Es apropiado usar métodos alternativos en las situaciones, por cierto frecuentes, de contar con datos poco adecuados, aunque cada método tiene dificultades, y
- Se debe ser ecléctico con respecto a los datos.

Además de las lecciones que ha aprendido el analista, tiene algunas recomendaciones clave para el futuro trabajo de evaluación de PROSCOL. En primer lugar, sería conveniente excluir aleatoriamente algunas familias calificadas dentro de PROSCOL en el resto del país y luego realizar una encuesta de seguimiento de los participantes reales y los excluidos en forma aleatoria de la participación. Esto entregaría una estimación más precisa de los beneficios. Sin embargo, sería políticamente delicado excluir a algunos. No obstante, si el programa no cuenta con suficientes recursos para cubrir todo el país en una sola etapa y se deberá optar por quienes serán seleccionados primero, por cierto sería preferible efectuar esa selección en forma aleatoria entre los participantes calificados. Alternativamente, sería posible seleccionar las escuelas o las áreas de juntas escolares en forma aleatoria en la primera etapa. Sin duda, esto haría de la elección de las escuelas o áreas de junta escolar una buena variable instrumental para la asignación de cada programa.

En segundo lugar, si esto no es factible, se recomienda realizar una encuesta básica en las áreas en las cuales es probable que haya una elevada concentración de participantes en PROSCOL antes que comience el programa en el sur. Esto se podría realizar simultáneamente con la siguiente etapa de la encuesta nacional que se utilizó para evaluar el programa PROSCOL. También sería conveniente agregar algunas preguntas a la encuesta, por ejemplo, si los niños realizan algún trabajo remunerado.

Y en tercer lugar, sería útil incluir trabajo cualitativo, para ayudar a plantear las hipótesis

que se deben probar y evaluar la credibilidad de los supuestos clave establecidos en el análisis cuantitativo.

### **Nota**

1. En Heckman, Lalonde y Smith (1999) y Abadie, Angrist e Imbens (1998) se encuentra un análisis sobre los efectos del tratamiento en los cuartiles.