



Grupo de Análisis para el Desarrollo

Informe final

Evaluación de Impacto del Programa “Formación de líderes empresariales en el departamento de Huancavelica”

Miguel Jaramillo Baanante

Sandro Parodi Sifuentes

Febrero, 2005

1. Introducción

La acumulación de capital humano es un aspecto clave para cualquier estrategia de reducción de la pobreza y promoción del desarrollo. La inversión temprana en nutrición, salud y educación provee la base para la inserción productiva y el desarrollo personal. No obstante, el proceso es permanente, particularmente en el desarrollo de habilidades para insertarse y desempeñarse en el mercado laboral. En esta dirección, las políticas y programas de capacitación juegan un rol clave, especial pero no únicamente al inicio de la vida laboral. El Fondo Nacional de Capacitación Laboral y de Promoción del Empleo (FONDOEMPLEO), viene financiando desde 1998, proyectos de instituciones públicas y privadas destinados a desarrollar competencias laborales y empresariales, mediante capacitación y asistencia técnica que promueva oportunidades de empleo sostenibles. Parte de las políticas de FONDOEMPLEO abarca la verificación y cuantificación de los efectos de los proyectos que patrocina y de esta manera puede contar con información para recuperar lecciones para futuras intervenciones. En este contexto, FONDOEMPLEO solicitó a GRADE la evaluación de impacto del programa orientado a jóvenes emprendedores aplicado por el Colectivo Integral de Desarrollo en el departamento de Huancavelica.

El Colectivo Integral de Desarrollo (CID) viene ejecutando proyectos orientados a la formación y promoción de emprendedores y creadores de negocios entre la población joven desde 1992. Entre junio de 1999 y noviembre de 2001 ejecutó, con recursos de FONDOEMPLEO, el “Programa de Calificación de Jóvenes Emprendedores de Microempresas” en ocho ciudades del país. En la actualidad, el CID tiene a su cargo el proyecto “Formación de Líderes Empresariales en el Departamento de Huancavelica”, el cual se viene implementando en las provincias de Huancavelica, Tayacaja y Angaraes del departamento de Huancavelica. El presente documento expone los resultados de la evaluación realizada de los impactos de este programa. Específicamente, los objetivos que persigue la evaluación son:

- Para los que inician el programa con un negocio: identificar los impactos en ingresos y empleo de los beneficiarios del programa

- Para los que inician el programa sin un negocio: identificar si el programa tiene efectos en la creación de negocios, además de impactos en sus ingresos.
- Probar un modelo de selección aleatoria de participantes a fin de obtener una medida cuantitativa rigurosa del efecto del Programa.

Los resultados centrales del análisis arrojan que el Programa tiene un impacto positivo tanto sobre la tenencia de negocios como sobre los ingresos de los participantes. En el primer caso, el Programa incrementa la probabilidad de tener un negocio en 8 puntos porcentuales. En cuanto a la segunda variable de resultado, la participación en el Programa incrementa los ingresos mensuales en 7.7 puntos porcentuales.

El texto se organiza en cinco secciones adicionales a esta introducción. La segunda sección describe las principales características del Programa. La tercera discute aspectos conceptuales referidos al problema de la evaluación de programas y, dadas las características del programa “Formación de líderes empresariales”, argumenta a favor de la conveniencia de aplicar un diseño experimental. La sección cuatro aborda aspectos operativos y metodológicos de la estrategia de evaluación elegida, describe como se desarrolló el proceso de selección del grupo de tratamiento y de control y, presenta comparaciones de algunas variables de interés entre ambos grupos. La sección finaliza con el detalle y resultados del recojo de información para la evaluación de impacto. La quinta sección presenta y discute los resultados de la evaluación. La sexta sección concluye.

2. Descripción del Proyecto

El proyecto ejecutado por el Colectivo Integral de Desarrollo (CID) en Huancavelica se denominó “Formación de Líderes Empresariales” y estuvo orientado a apoyar a jóvenes con voluntad de iniciar un negocio o fortalecer uno recién iniciado. Sus principales objetivos fueron: (1) incrementar la capacidad de crear negocios sostenibles y (2) generar empleo e ingresos en los emprendedores y creadores de negocios de Huancavelica a través del desarrollo de sus competencias empresariales. Su población objetivo estuvo conformada por personas menores de 35 años *creadores de negocios*, esto es, aquellos que tuvieran negocios en

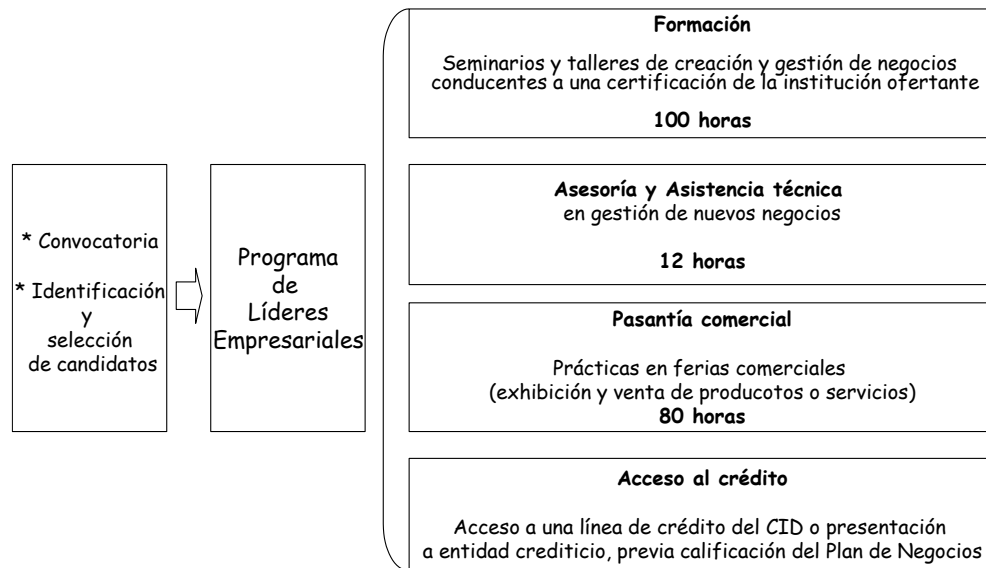
funcionamiento (principalmente microempresas) y con una antigüedad menor a un año, o personas *emprendedoras*, con ideas de negocios a punto de ser implementadas o que ya estuvieran desarrollando alguna actividad generadora de ingresos pero como una ocupación secundaria.

Los potenciales beneficiarios fueron captados por el CID a través de convocatorias abiertas y públicas. Los interesados pasaron luego por un proceso de selección y los que resultaron beneficiados pudieron acceder al denominado “Programa Regular” que contaba con los siguientes componentes:

- Un proceso de formación compuesta por seminarios y talleres de creación y gestión de empresas,
- Asesorías y asistencia técnica con una duración conjunta de 12 horas,
- Una pasantía comercial donde los participantes participan en ferias comerciales de la zona o en eventos organizados por el propio CID,
- Acceso al crédito o presentación a entidad crediticia al final del proceso, previa calificación de un Plan de Negocios.

En el Diagrama 1 se puede observar los componentes del Programa de manera esquemática.

Diagrama 1. Componentes del Programa Formación de Líderes Empresariales



Fuente: Colectivo Integral de Desarrollo

Adicionalmente, existen los denominados programas especiales, que básicamente contienen un tratamiento de capacitación y asesoría similar al que vienen recibiendo los del programa regular, pero que se imparten en un menor tiempo y de manera más intensiva. El CID a mediados de febrero de 2004, durante su quinta convocatoria, implementó un programa especial. Los participantes de este programa recibieron un módulo de capacitación compuesto por seminarios de 60 horas de duración y talleres de 40 horas. Este módulo se impartió entre la tercera semana de febrero y la segunda de marzo de 2004. Posteriormente, el programa asignó a cada beneficiario un asesor a cargo de brindar orientación en temas de gestión empresarial, así como asistencia técnica específica para el tipo de negocio que el participante estuviera desarrollando. Los beneficiarios tuvieron derecho a esta asesoría personalizada durante seis meses, desde marzo hasta agosto de 2004. Cabe mencionar que la evaluación de impacto se desarrolló en el contexto de este programa especial. Más adelante, en la sección metodológica, se explica en detalle la implementación de la evaluación experimental.

El Programa Líderes Empresariales llevó a cabo seis convocatorias en las que pudieron participar un total de 690 individuos. Descontando a los desertores del Programa, el grupo de participantes activos llega a 587. Por otro lado, hasta la fecha, 151 de estos jóvenes

completaron el programa y recibieron la constancia, mientras que 255 lograron el certificado. Un grupo de participantes todavía se encuentra culminando su capacitación. Asimismo, se realizaron 7 ferias y exhibiciones en las que 124 participantes tuvieron la oportunidad de ofrecer sus productos. Respecto al acceso al crédito, se puede mencionar que a un total de 72 beneficiarios se les otorgó un crédito.

3. Aspectos conceptuales

Durante los últimos años, la literatura sobre el diseño de modelos metodológicos para la implementación de evaluaciones de proyectos sociales ha tenido un fuerte desarrollo. Cada vez más los diseñadores de política tienen clara la importancia de evaluar los proyectos sociales, en especial en los países en vías de desarrollo, en los cuales los recursos son escasos y es primordial asignarlos de la manera más eficiente posible. Esto ha motivado e impulsado un fuerte desarrollo de la investigación en este campo.¹ En esencia, toda evaluación de proyectos sociales intenta responder tres preguntas básicas, a saber: ¿Cómo se ha venido implementando el programa? ¿Cuáles han sido los efectos del programa? y de haber efectos ¿A qué costo se han conseguido esos efectos? Para responder la primera interrogante típicamente se aplican instrumentos de monitoreo donde se intenta verificar que las actividades del programa se hayan desarrollado de acuerdo a lo planificado. La segunda pregunta viene a ser la más compleja y se aborda mediante las denominadas evaluaciones de impacto, cuyas características detallaremos más adelante. Finalmente, la tercera interrogante intenta medir la eficiencia económica de los costos de un programa versus sus beneficios y permite a las autoridades responsables comparar alternativas de programas.²

¹ En Estados Unidos la evaluación de programas sociales tiene larga data. Más recientemente, en algunos países, como México, por ejemplo, se ha establecido la obligatoriedad de los programas sociales implementados con fondos públicos de contar con planes de evaluación de sus resultados previamente al inicio de su implementación. Recientemente, el Congreso alemán (*Bundestag*) paralelamente a la aprobación de un paquete de reformas del mercado laboral solicitó a entidades académicas el desarrollo de un concepto de evaluación que permita medir empíricamente de manera confiable sus efectos dos y tres años después de su implementación (Fertig y Kluve 2004).

² Baker, Judy (2000). "Evaluación del Impacto de los Proyectos de Desarrollo en la Pobreza: Manual para Profesionales". Washington, D. C.: Banco Mundial.

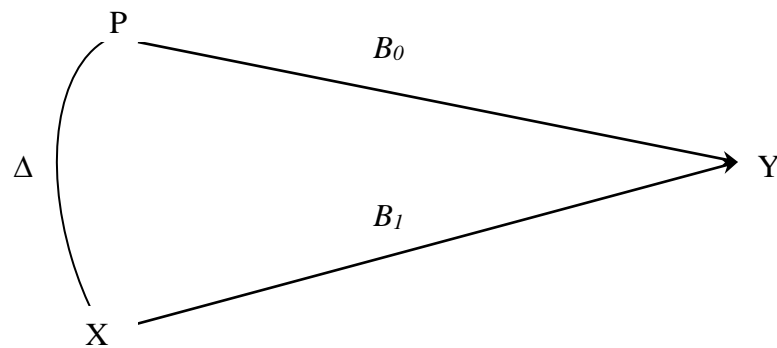
Una evaluación de impacto trata de determinar cambios en el bienestar de los individuos que puedan ser atribuibles a un proyecto, un programa o una política en particular. Mediante las evaluaciones de impacto se recogen lecciones que sirven para retroalimentar programas o proyectos en curso o ya finalizados, y con esto mejorar la efectividad de intervenciones actuales o futuras. Este tipo de evaluaciones se justifican debido a que brindan información valiosa a los tomadores de decisiones con respecto a la necesidad de expandir, modificar, o incluso eliminar una política o intervención en particular. De manera formal, lo que pretende verificar una evaluación de impacto es la hipótesis de la existencia de una relación causa-efecto entre la intervención (causa) y el logro de los objetivos (efectos) de un determinado programa. La pregunta no es sencilla de responder, debido a que no es posible conocer que hubiese pasado con cada miembro de la población beneficiaria de no haber participado en el programa. Aunque uno no puede observar tal situación, el evaluador intenta aproximarla a través de la construcción de un contrafactual, definido como una situación hipotética que intenta representar los niveles de bienestar alcanzados por los individuos en ausencia de una política o de un programa.

El contrafactual se simula comparando las variables que el programa intenta afectar en dos grupos de individuos. El grupo de tratamiento, conformado por aquellos individuos que participaron en el programa y, el grupo de control, por individuos que presentan características similares a los beneficiarios pero que no participaron. La clave del éxito en toda evaluación de impacto descansa en la construcción de un grupo de control capaz de simular lo que hubiera pasado con los individuos del grupo de tratamiento en caso de no haber participado en el programa. Es común que en este proceso aparezca lo que se considera el problema metodológico fundamental de toda evaluación: el sesgo de selección, que se origina como consecuencia de las diferencias, observables y no observables, entre el grupo de tratamiento y el grupo de comparación que se haya construido.

El sesgo de selección

El Diagrama 2, presentado a continuación, ilustra un modelo que permite visualizar dicho problema, así como sus consecuencias para las evaluaciones.³ Este modelo define Y como la variable que se espera el programa impacte, pudiendo ser, por ejemplo, los niveles de ingresos o la identificación de si el individuo tiene o no un negocio. La variable P determina si los individuos pertenecen o no al grupo de tratamiento y el vector X recoge información sobre las características de los individuos, referidas al sexo, grado de instrucción, edad, motivación, habilidad, etc.

Diagrama 2. Sesgo de selección



El impacto del programa cuando el grupo de tratamiento y el de control son estadísticamente equivalentes, esto es, cuando se cumple la condición de validez interna, está representado por la línea que va de P a Y , donde B_0 recoge el signo y el tamaño del efecto. La línea que va de X a Y representa el efecto del vector de características observables y no observables de los individuos que afectan las variable de resultado y, como en el caso anterior, el coeficiente B_1 resume el signo y el efecto causal que tienen estas variables sobre la variable de interés. Sin embargo, es posible que exista una relación entre el status del individuo con relación al programa, P , y las variables de características individuales, X . El tamaño y el signo de esta relación se recogen a través de Δ y viene a representar la diferencia media que existe en las variables individuales X entre el grupo participante en el programa y

³ Bloom, Howard, et al (2002). "Can Nonexperimental Comparison Group Methods Match the Findings from a Random Assignment Evaluation of Mandatory Welfare-to-Work Programs?" MDRC Working Papers on Research Methodology.

el grupo de control. En otras palabras, Δ da una idea de la magnitud del sesgo de selección y tiene importantes implicancias en la estimación del efecto del programa. Si Δ es diferente de cero, el efecto estimado ($B_0 + \Delta$) es más o menos que el impacto verdadero del programa y los resultados son sesgados.

En la práctica estos sesgos o diferencias entre el grupo de tratamiento y de control tienen lugar debido a diversas razones. Por un lado, los individuos que deciden participar en el programa presentan características particulares en comparación con los que quedan fuera, tanto observables (sexo, educación, etc.) como no observables (motivación, autoestima, ambición, etc.), esto se conoce como autoselección. Por otro lado, el mismo programa puede tener mecanismos de reclutamiento o escrutinio que niegan el acceso a individuos con ciertas características (*staff selection*). Por último, el evaluador puede escoger un grupo de comparación que no es el apropiado.

La literatura privilegia dos métodos para asegurar la construcción de un grupo de comparación adecuado. En primer lugar se tiene el “diseño experimental” que, partiendo de un conjunto de individuos que han mostrado interés en el programa y que cumplen con los requisitos, los asigna de manera aleatoria a dos sub-conjuntos. Por un lado, un grupo de individuos que recibirán el programa-el grupo de tratamiento- y por otro lado, un grupo de individuos a los cuales se les niega o difiere la participación y sirven como grupo de comparación. El segundo método es el “no experimental” “cuasi-experimental” o “econométrico” que, con apoyo en metodologías estadísticas y el uso de modelos de comportamiento, construye un grupo de control utilizando una variedad de fuentes de datos, muchas veces pre-existentes, con la finalidad de comparar las diferencias que hubieran con el grupo de tratamiento en las variables que el programa quiere afectar.

Durante las últimas dos décadas ha tenido lugar un amplio debate entre renombrados académicos, en el intento de determinar la superioridad de alguno de los dos métodos. Aunque en la actualidad dicho debate mantiene vigencia, existe consenso en que el diseño experimental es el enfoque que brinda a los evaluadores los resultados más rigurosos y que el diseño no experimental, a pesar de los muchos avances metodológicos logrados recientemente, en contadas ocasiones logra igualar los resultados del diseño experimental.

Uno de los primeros en comparar estos dos enfoques fue La Londe (1986). Tomando como base los beneficiarios de una evaluación experimental, la National Supported Work Demonstration (NSW), construyó grupos de comparación no experimental utilizando información de encuestas existentes con la finalidad de estimar el impacto del programa y detectar diferencias en relación a los resultados del experimento. Luego de utilizar diferentes estimadores usados comúnmente en las evaluaciones no experimentales, el autor encontró una gran variedad de efectos, los cuales diferían sustancialmente del resultado del experimento. Fraker y Maynard (1987) llegan resultados similares que La Londe y llaman la atención sobre los riesgos de usar métodos no experimentales para estimar el impacto de programas sociales. Sin embargo, Heckman y Hox (1989), principales defensores de los métodos no experimentales, argumentan que es posible estimar impactos precisos de los programas cuando se tiene cuidado en escoger, mediante un *test* que logran desarrollar, especificaciones adecuadas y la información ha sido cuidadosamente recogida y analizada.

Recientemente, Bloom et al. (2002), en un intento de echar luces sobre el tema, estiman una serie de resultados aplicando los más recientes procedimientos de construcción de grupos de control así como metodologías de estimación no experimental, siguiendo un riguroso protocolo pre-establecido. Una vez estimados los efectos, son comparados con los resultados de diseños experimentales aplicados en los mismos programas. Entre sus principales hallazgos, encuentran que el mejor método no experimental, entendido como aquél que permite construir el grupo de control que arroja los resultados de impacto del programa más parecidos a aquéllos del diseño experimental, o que minimiza los sesgos, es aquél que empata individuos de la misma localidad, además de las variables que comúnmente se utilizan. Asimismo, encuentran que los estimadores obtenidos con diferencias de medias simples o con el método de Mínimos Cuadrados Ordinarios (MCO), especialmente cuando se usan con grupos balanceados usando el puntaje de propensión (*propensity score*), son iguales o mejores que otros métodos más complejos (modelos censurados, modelos de selección, etc.). Sin embargo, agregan que aún con el mejor método no experimental no se puede acercar con precisión a la estimación mediante métodos experimentales.

El diseño experimental y la solución del sesgo de selección

La principal fortaleza del diseño experimental es que al determinar al azar el grupo de tratamiento y el grupo de control elimina el sesgo de selección. Siguiendo a Heckman y Smith (1995) se puede demostrar formalmente como el diseño experimental logra solucionar el problema del sesgo de selección. Supongamos que una persona puede estar en un estado de tratamiento, denotado por “1”, ó un estado de no tratamiento, denotado por “0” y que existe una variable resultado, denotada por Y_1 y Y_0 , asociada a cada estado. Si uno quisiera estimar el efecto del tratamiento bastaría con estimar la diferencia, D , entre ambos estados, es decir $Y_1 - Y_0$. Por otro lado, dado que muchos de los programas de capacitación no se implementan sobre toda una población sino que se concentran en algunos segmentos, es necesario distinguir entre los individuos de la población que participaron del tratamiento, $d = 1$, de aquellos que no participaron, $d = 0$. Así, el impacto esperado del tratamiento se puede definir de la siguiente manera: $E(D/d=1) = E(Y_1 - Y_0/d=1)$. En dicha ecuación $E(Y_1/d=1)$ es conocido, pero en cambio, $E(Y_0/d=1)$, que recoge lo que hubiera sucedido con el grupo de tratamiento de no haber participado, es imposible de estimar. El evaluador puede construir un grupo de control, $E(Y_0/d=0)$, para usarlo como *proxy* de $E(Y_0/d=1)$ y así estimar el impacto a través de la siguiente ecuación:

$$E(Y_1/d=1) - E(Y_0/d=0) \dots \dots \dots (1)$$

Sin embargo, como mencionamos líneas arriba, el sesgo de selección tendrá lugar de existir diferencias entre el grupo de control $E(Y_0/d=0)$ y el grupo de tratamiento en el estado de no participación $E(Y_0/d=1)$. Al sumar y restar $E(Y_0/d=1)$ a la ecuación se puede tener una idea más clara de la naturaleza del sesgo.

$$\underbrace{E(Y_1/d=1) - E(Y_0/d=1)}_D + \underbrace{E(Y_0/d=1) - E(Y_0/d=0)}_\Delta \dots \dots \dots (2)$$

El primer término viene a ser el impacto del tratamiento, mientras que el segundo la magnitud del sesgo de selección discutido líneas arriba. El diseño experimental resuelve este problema generando un grupo de control experimental compuesto por personas que tenían

interés en participar pero que aleatoriamente se les negó el acceso al tratamiento. Así, la media esperada del grupo de control experimental estima el contrafactual deseado, a saber, $E(Y_0/d=1)$. De manera formal, supongamos que $d^*=1$ representa las personas con interés en participar en el tratamiento y $d^*=0$, a las demás personas. La aleatorización solamente se aplica al grupo $d^*=1$ y denota como $r=1$ al grupo que salió sorteado para recibir tratamiento y como $r=0$ al grupo de control. De esta manera, es posible estimar el impacto medio esperado del programa libre del sesgo de selección mediante la siguiente ecuación:

$$E(Y_1 - Y_0/d=1) = E(Y_1 / r=1 y d^*=1) - E(Y_0 / r=0 y d^*=1).....(3)$$

Donde $E(Y_1 / r=1 y d^*=1)$ representa el resultado en la variable de interés del estudio para el grupo que recibió el tratamiento, mientras que $E(Y_0 / r=0 y d^*=1)$ mide el resultado del grupo de control experimental que recoge el resultado en la variable de interés que hubiera alcanzado el grupo de tratamiento de no haber participado en el programa.

Tomando en consideración los argumentos presentados, este estudio implementó un diseño experimental a fin de obtener los estimados más confiables de los efectos que el Programa haya podido lograr en sus beneficiarios en el departamento de Huancavelica.

4. Implementación del diseño experimental

Un primer paso para la ejecución del diseño experimental consistió en la construcción de una línea de base que asegurara un recojo de información adecuada y pertinente para los fines de la evaluación. En esta dirección, el equipo de evaluación llevó a cabo una revisión completa del sistema de información de programa del CID, así como de los alcances de la información recolectada. Del análisis realizado, los consultores alcanzaron una serie de recomendaciones con relación a las fichas de recojo de información que el programa utilizaba. Estas recomendaciones fueron implementadas por el Programa.

En esta etapa también se definió más precisamente el proceso para la implementación del experimento. Así, se estableció que la implementación se realizaría en el contexto de la

quinta convocatoria del Programa. Para la segunda semana de febrero de 2004 el CID, conjuntamente con CARITAS, logró reunir un total de 196 individuos que cumplían los requisitos y estaban interesados en participar del programa, ($d^* = 1$). Por la misma fecha, personal de GRADE viajó a la ciudad de Huancavelica a fin de realizar la asignación aleatoria para participar en el Programa en presencia de representantes de ambas instituciones. Mediante la utilización de un software estadístico se procedió a constituir el grupo de tratamiento, es decir, ($r = 1 / d^* = 1$) con el 60% de los individuos interesados. El restante 40% vino a constituir el grupo de control experimental, ($r = 0 / d^* = 1$). Cabe señalar que, con el fin de evitar malestar entre los individuos asignados al grupo de control, el Programa se comprometió a asegurarles a cada uno su condición de beneficiario para la siguiente convocatoria, mediante la entrega de una credencial.⁴

Las características relevantes de los individuos asignados a cada grupo se presentan en la Tabla 1. Así mismo, la tercera columna muestra el resultado del test de diferencia de medias entre ambos grupos. Como se puede comprobar mediante este procedimiento, los valores de las características observables en ambos grupos son estadísticamente similares (ninguno de los valores ‘t’ resulta significativo), con lo que se valida la construcción de un contrafactual adecuado.

Tabla 1. Diferencias entre ambos grupos luego de asignación aleatoria

Características	Control	Tratamiento	T-test
	%	%	
Creador de negocio	56.4	56.8	0.960
Sexo masculino	89.7	90.7	0.830
Casado o conviviente	75.6	75.4	0.973
Máximo nivel educativo primaria	34.6	35.6	0.889
Máximo nivel educativo secundaria	56.4	54.2	0.766
Máximo nivel educativo superior	6.4	4.2	0.501
Edad promedio	28.7	29.0	0.726
Total de observaciones	78	118	

⁴ El inicio de dicha convocatoria (noviembre de 2004) fue inmediatamente posterior al recojo de la información para la evaluación, por lo que no ha habido posibilidad de “contaminación”.

Con relación a las características de los participantes, se puede decir que poco más de la mitad ya cuenta con una actividad independiente a la que dedican la mayor parte de su tiempo. Son mayoritariamente hombres, nueve de cada diez aproximadamente y sus edades promedio alcanzan 28 años en el grupo de control y 29 años en el grupo de tratamiento. Además, tres de cada cuatro individuos se encuentran casados o conviviendo con una pareja. En cuanto al nivel educativo, como era de esperarse, muy pocos han alcanzado un grado universitario. Asimismo, más de la tercera parte solamente ha logrado el nivel educativo primario. El logro educativo en el que se concentra la mayoría es el nivel secundario con poco más de la mitad de los individuos en ambos grupos.

El programa especial comenzó sus actividades a finales de febrero con un seminario de capacitación intensiva, de una semana de duración. Cabe mencionar, sin embargo, que en este tipo de evaluaciones, es común que algunos de los individuos que resultan seleccionados renuncien a su condición de beneficiarios y dejen de recibir el tratamiento, lo que se conoce en la literatura como la deserción ó el *dropout*. En el contexto de esta evaluación, como se muestra en la Tabla 2, de los 118 individuos pertenecientes al grupo de tratamiento, 42 no asistieron nunca al programa. Asimismo, tres individuos que resultaron asignados al grupo de control se pasaron al grupo de tratamiento (tránsfugas) recibiendo la totalidad de la capacitación. Así, tomando en consideración estos hechos, al final se tiene un grupo de tratamiento conformado por 76 individuos y un grupo de control con 75 individuos. El impacto del programa se estimó comparando las medias de las variables de resultado de estos dos grupos.

Tabla 2. Estadísticas luego del programa especial

Categoría	Frecuencia	%
Control	75	38.3
Tratamiento	76	38.8
Desertores (<i>dropouts</i>)	42	21.4
Tránsfugas	3	1.5
Total	196	100

Estos cambios podrían o no introducir sesgos en los estimadores, dependiendo de si la deserción es aleatoria con relación a los resultados potenciales. Puesto que esto no es obvio, como paso previo para la estimación del impacto del Programa, es indispensable comprobar la validez interna del diseño experimental, luego de los cambios en la composición original de los grupos. La presencia de desertores plantea un problema para la evaluación, ya que el estimador, $E(Y_1) - E(Y_0)$, aproxima el impacto medio sobre los asignados al tratamiento más no el impacto medio del tratamiento mismo, lo que se conoce como “tratamiento completo sobre los completamente tratados” (*full treatment on the fully treated*) y que se puede definir de la siguiente manera:⁵

$$D_p = E(Y_1/k=0) - E(Y_0/k=0) \dots \dots \dots (4).$$

Donde k toma valor “1” para los desertores, personas del grupo de tratamiento que no recibieron el tratamiento, y “0” para los miembros del grupo de tratamiento que si recibieron el programa. En la fórmula presentada, $E(Y_1/k=0)$, denota la media esperada de la variable resultado para el grupo de individuos que recibieron el tratamiento y $E(Y_0/k=0)$ denota la media esperada de la variable resultado para los individuos del grupo de control que hubieran recibido el tratamiento de haber sido seleccionados para participar del programa. Sin embargo, el estimador del impacto del programa, D_p , presenta una seria dificultad para ser identificado, debido a que $E(Y_0/k=0)$, es desconocido por el evaluador.

Mallar et al. (1980), Smith et al. (1984) y Bloom (1984), entre otros, resuelven el problema al plantear un supuesto de identificación. En él, se asume que la media esperada de la variable resultado de los desertores del grupo control, es la misma que la media esperada de los desertores del grupo de tratamiento. Formalmente, asumen que $E(Y_1/k=1) = E(Y_0/k=1)$ y luego de alguna manipulación algebraica llegan al siguiente término:⁶

⁵ Heckman, James, J. Smith and C. Taber (1998). “Accounting for Dropouts in Evaluations of Social Program”. *The Review of Economics and Statistics*, Vol. 80, N° 1, pp. 1-14.

⁶ Asumiendo la presencia de desertores, la media esperada en la variable resultado para el grupo de tratamiento se puede expresar de la siguiente manera: $E(Y_1) = t_1 E(Y_1/k=1) + (1-t_1) E(Y_0/k=0)$, donde t representa la probabilidad de abandonar el programa. La misma expresión se puede plantear para el caso del grupo de control: $E(Y_0) = t_0 E(Y_0/k=1) + (1-t_0) E(Y_0/k=0)$. Con estos datos y asumiendo que la asignación aleatoria produce que $t_1 = t_0 = t$, es posible identificar la ecuación 5. Para mayor detalle ver Heckman et al. (1998), op. cit.

$$D_p = \frac{E(Y_1) - E(Y_0)}{1-t} \dots\dots\dots(5)$$

donde t representa la fracción de miembros del grupo de tratamiento que no recibieron el tratamiento, los desertores. El estimador D_p escrito de esa forma, corrige el efecto de los desertores y transforma el estimador del impacto sobre aquéllos asignados al programa en un estimador del impacto medio de la participación completa en el programa. Esto al asumir un impacto nulo del programa sobre los desertores.⁷

A continuación, en la Tabla 3 se presentan las características más relevantes de los individuos asignados a cada grupo sin tomar en cuenta los desertores ni los trásfugas. Asimismo, la tercera columna reporta los resultados del *test* de igualdad de medias. Como se puede observar, a pesar de los desertores las características del grupo de tratamiento son estadísticamente similares a las del grupo de control.

Tabla 3. Diferencias entre ambos grupos descontando desertores y trásfugas

Características	Control	Tratamiento	T-test
	%	%	
Creador de negocio	58.7	57.9	0.924
Sexo masculino	90.7	90.8	0.979
Casado o conviviente	77.3	72.4	0.485
Máximo nivel educativo primaria	36.0	36.8	0.915
Máximo nivel educativo secundaria	54.7	51.3	0.682
Máximo nivel educativo superior	6.7	3.9	0.459
Edad promedio	28.9	30.1	0.277
Total de observaciones	75	76	

⁷ Heckman et al. (1998) argumentan que este último supuesto es difícil de sostener en contextos de desertores con tratamiento parcial. En nuestro caso, los desertores tiene tratamiento nulo pues a pesar de haber sido notificados nunca recibieron ninguno de los componentes del programa.

Levantamiento de información ex post

El programa evaluado se ejecutó desde fines de febrero de 2004 hasta el mes de junio del mismo año. El trabajo de campo se planificó para la última semana de octubre, cuatro meses después de finalizado el programa, esto con la idea de darles tiempo a los participantes de aplicar lo aprendido y evaluar la sostenibilidad de las ideas de negocio que hubieran podido implementar durante el tratamiento. Dicho trabajo de campo contempló entrevistar no sólo el grupo de tratamiento y control sino también a los desertores.

Los resultados del trabajo de campo se reportan en la siguiente Tabla 4. En total se logró entrevistar a 149 individuos, distribuidos en 53 controles, 64 beneficiarios, 30 desertores y 2 tráfugas. Esto viene a representar porcentajes superiores al 70% de la muestra original que se planteó entrevistar como meta.

Tabla 4. Resultados del trabajo de campo

Categoría	Total a entrevistar	Total entrevistados	% de logro	Total de Atrición
Controles	75	53	70.7	22
Beneficiarios	76	64	84.2	12
Desertores	42	30	71.4	12
Tráfugas	3	2	66.7	1
Total	196	149		47

Al igual que la deserción, la atrición plantea problemas metodológicos que podrían sesgar la estimación del efecto del programa. Una manera de determinar la importancia de este problema consiste en comparar las características entre los individuos del grupo de tratamiento y control para los que se tiene información *ex post* y ver si la similitud inicial en dichas variables todavía se mantiene. Tal como se pudo observar en la Tabla 5, que excluye los desertores ($k=1$) y los individuos de los que no se tiene información *ex post*, (atrición=1), el grupo de tratamiento y el grupo de control presenta características estadísticamente similares para todas las variables, salvo en el caso del estado civil. En este caso, sin embargo, la significancia estadística no es muy fuerte, por un lado, y por otro, análisis anteriores han mostrado que esta variable no influye significativamente sobre las variables de resultado.⁸

⁸ Ver Jaramillo y Parodi, (2003).

Tabla 5. Características de los grupos de análisis luego del trabajo de campo

Características	Control	Tratamiento	T-test
	%	%	
Creador de negocio	62.3	57.8	0.628
Sexo masculino	94.3	92.2	0.650
Casado o conviviente	84.9	71.9	0.093*
Máximo nivel educativo primaria	39.6	39.1	0.951
Máximo nivel educativo secundaria	52.8	51.6	0.893
Máximo nivel educativo superior	5.7	1.6	0.228
Edad promedio	29.2	30.1	0.520
Total de observaciones	53	64	

Nota: *significativo al 90%, se rechaza la hipótesis nula de igualdad de medias
El grupo de beneficiarios no considera los desertores, ni los trásfugas.

5. Resultados de la evaluación

Tal como se mencionó anteriormente la presente evaluación experimental tiene como objetivos específicos determinar los efectos del programa sobre: (1) la creación de negocios y (2) los ingresos de los participantes. A continuación se presentan las estimaciones realizadas con el grupo de tratamiento de 64 individuos y el grupo de control de 53 individuos.

Efectos sobre la tenencia de negocios

En primer lugar es importante mencionar que el programa se desarrolló en un contexto netamente rural. En este tipo de contexto es poco común encontrar individuos en condición de desempleados o inactivos. Por otro lado, entre los activos, la presencia de actividades independientes con la finalidad de agenciarse recursos económicos es altamente probable. Asimismo, los individuos, sin distinguir a que grupo pertenezcan, dedican gran parte de su tiempo a actividades agrícolas y al comercio de cereales y granos. Tanto los productos que cosechan como los que compran son usualmente comercializados en ferias dominicales. Por otro lado, muchos individuos se dedican a la crianza y/o engorde de animales (cuyes, vacas, toros, etc.) que también son vendidos en las ferias dominicales. Por último, cabe anotar que

estas actividades generalmente se entrecruzan, pudiendo un joven, por ejemplo, durante los días de la semana estar trabajando en sus tierras y alimentando sus animales y en los días de feria dedicarse a actividades comerciales.

Tomando en cuenta estos aspectos, para efectos de la evaluación se ha considerado una definición amplia del concepto de negocio. Así, un individuo cuenta con un negocio si desarrolla cualquier actividad independiente que le genere ingresos, aunque sea su actividad secundaria. Dado que la selección aleatoria del experimento iguala las condiciones iniciales entre el grupo de tratamiento y de control, se puede asumir que las diferencias que se encuentren en el estimador de creación de negocios al utilizar esta definición son atribuibles al programa.

La Tabla 6 presenta el efecto del programa sobre la variable tenencia de negocios. Como se puede observar existe una clara diferencia a favor del grupo de tratamiento, donde el 85.9% de los individuos desempeñan actividades independiente, mientras que en el grupo de control lo hace el 73.6% de los individuos. Esta diferencia alcanza el 12.4%. Sin embargo, estos resultados se obtienen de comparar el grupo de tratamiento que completó el programa con el grupo de control y no incorpora a los individuos que formaban parte del grupo de tratamiento pero que abandonaron el programa, los desertores.

Tabla 6. Resultados de tenencia de negocio

Grupo	Tiene negocio	
	Frecuencia	%
Control $E(Y_0)$	39	73.6
Tratamiento $E(Y_1/k=0)$	55	85.9
Diferencia $E(Y_1/k=0) - (Y_0)$		12.4

Como se explicó antes, para estimar el impacto del programa se debe tomar en cuenta a los desertores del programa. Siguiendo la metodología propuesta por Heckman et al. (1998) el impacto del programa D_p se obtiene de la diferencia en la variable resultado entre el grupo de tratamiento original-que incluye los desertores- y el grupo de control. En la Tabla 7 se

presentan los resultados de las estimaciones. El 78.7% de los individuos del grupo de tratamiento original, $E(Y_1)$, conformado por aquellos que terminaron el programa y los desertores tienen un negocio. Al hacer la diferencia con el grupo de control se encuentra una diferencia a favor del programa de 5% y después de corregir este estimado por la tasa de desertores se encuentra que el impacto del programa es de 8%. Esto quiere decir que el hecho de participar en el Programa incrementa las probabilidades de contar con un negocio en ocho puntos porcentuales.

Tabla 7. Impacto del tratamiento sobre la tenencia de negocio

Categoría		Tiene negocio	
		Frecuencia	%
Controles	$E(Y_0)$	39	73.6
Tratados	$E(Y_1/k = 0)$	55	85.9
Desertores	$E(Y_1/k = 1)$	19	63.3
Tratados + Desertores	$E(Y_1)$	74	78.7
Diferencia	$E(Y_1) - E(Y_0)$		5.1
Impacto	D_p		8.0

Nota: el impacto del programa, D_p , se determina corrigiendo la diferencia $E(Y_1) - E(Y_0)$ con la tasa de deserción del Programa (35.59%), siguiendo la siguiente fórmula:

$$D_p = \frac{E(Y_1) - E(Y_0)}{1 - t}$$

Efectos sobre los niveles de ingreso mensual

Se ha utilizado una definición de ingreso que incorpore el hecho de que los ingresos de las personas que participan en la evaluación muestran fluctuaciones a lo largo del año. Por ejemplo, en algunos meses del año es la época de siembra y los ingresos del individuo por concepto de esta actividad es nulo. Precisamente, el trabajo de campo se realizó en octubre de 2004, que coincide con la época de siembra. En este sentido, se ha procurado aproximar una medida del ingreso habitual que mensualmente perciben los individuos. Así, los ingresos que obtienen de sus cosechas fueron divididos en partes iguales para cada mes del año.⁹ En el caso de los individuos que desempeñan actividades comerciales, de manufactura u otra actividad no agrícola se tomó el dato mensual reportado en la encuesta.

⁹ Este procedimiento sólo se aplicó a los que reportaron en la encuesta un monto de ingresos global de sus ingresos agrícolas. En la mayoría de los casos, sin embargo, fueron los mismos agricultores quienes estimaron sus ingresos mensuales.

La Tabla 8 presenta el efecto del Programa sobre la variable del ingreso mensual habitual. Nuevamente, se encuentra una diferencia en el estimador a favor del grupo de tratamiento. Comparando los que completaron el programa $E(Y_1/k = 0)$ con el grupo de control se encuentra una diferencia de 44.7 soles. Sin embargo, el impacto del programa, D_p , se estima en 24.8 soles, esto después de corregir por la deserción. En términos porcentuales, se puede afirmar que el hecho de participar en el programa del CID incrementa los ingresos en 7.7%.¹⁰

Tabla 8. Impacto del tratamiento sobre los ingresos

Categoría	Ingresos mensuales		Impacto en %
	Frecuencia	S/.	
Controles $E(Y_0)$	48	321.0	
Tratados $E(Y_1/k = 0)$	59	365.8	
Desertores $E(Y_1/k = 1)$	28	276.4	
Tratados + Desertores $E(Y_1)$	87	337.0	
Diferencia $E(Y_1)-E(Y_0)$		16.0	
Impacto D_p		24.8	7.7%

Nota: el impacto del programa, D_p , se determina con corrigiendo la diferencia $E(Y_1)-E(Y_0)$ con la tasa de deserción del Programa (35.59%), siguiendo la siguiente formula:

$$D_p = \frac{E(Y_1) - E(Y_0)}{1 - t}$$

6. Reflexiones finales

Los resultados muestran impactos positivos y de nivel razonable para los dos variables de resultado. Aquí quisiéramos añadir algunos puntos adicionales sobre la lectura de los resultados, por un lado, algunas lecciones metodológicas del estudio y recomendaciones para Fondoempleo y el Programa mismo. El primer punto se refiere a la validez externa del diseño. Con esto nos referimos a la pregunta de ¿a qué población se puede extrapolar los resultados? Una primera cuestión en esta dirección es que los resultados en ningún caso significan un endose de, en general, políticas orientadas a capacitar para el autoempleo. Los resultados se refieren a un tipo de intervención específica en un contexto específico. Así, lo

¹⁰ Es importante mencionar que se separaron del análisis 12 observaciones, pues presentaban información inconsistente. De estas 7 pertenecen al grupo de beneficiarios y 5 al grupo de control. Los dos tráfugas también fueron separados del análisis.

que se puede afirmar es que el Programa Líderes Empresariales en Huancavelica tiene un impacto positivo sobre los resultados de creación de negocios e ingresos de sus beneficiarios. Debe subrayarse que el ámbito de acción del Programa es el rural, por tanto el análisis no asegura que los resultados sean los mismos en un contexto urbano, por ejemplo. Tampoco se puede asegurar que de aplicarse el mismo programa a escala nacional se obtendrían los mismos resultados. Esto no sólo por el tema de ámbito rural versus urbano, sino además porque puede haber un efecto escala importante. Dicho esto, se debe decir también que los resultados son alentadores, puesto que el Programa ha obtenido resultados en una zona particularmente pobre.

En segundo lugar, una lección metodológica se relaciona al recojo de información de las variables de resultado. En cuanto a la creación de negocios, sería importante detallar más el cuestionario en cuanto a la asignación del tiempo de trabajo de los individuos en el análisis. Esto porque la mayoría tienen alguna actividad (principal o secundaria) independiente y a veces es difícil definir cuál es la ocupación principal. Típicamente señalan a la agricultura, pero con frecuencia no es esta actividad la que les provee la porción mayor de sus ingresos. En cuanto al rubro de ingresos, el módulo aplicado es sumamente modesto para recoger información en un contexto rural y de multiplicidad de actividades. La sugerencia es completar el módulo considerando ingresos a lo largo de todo el año, puesto que la estacionalidad es un tema muy relevante en este tipo de contextos. Así mismo, debería detallarse las diferentes fuentes, que pueden ser variables en diferentes momentos del año.

Una recomendación tanto para el Programa como para Fondoempleo se refiere a la importancia de hacer un seguimiento en el tiempo de los impactos de la intervención. En esta dirección, debe tomarse en cuenta que se ha evaluado el impacto tan sólo cuatro meses luego de concluida la participación en el Programa. Una hipótesis a investigar es si los impactos sobre los ingresos son más altos cuando ha transcurrido un mayor tiempo y se han podido poner en práctica un mayor número de herramientas de gestión. La sugerencia es aprovechar que ya se tienen contruidos los grupos de manera aleatoria para realizar una nueva medición un año después de egresados los participantes del Programa. Si bien el grupo de control ya habrá pasado por el Programa, esta característica precisamente permitiría

medir el impacto adicional logrado por un mayor despliegue de las habilidades y capacidades aprendidas en el Programa.

Finalmente, reconocidas las ventajas de la aplicación de diseños experimentales, Fondoempleo haría bien en propiciar un mayor uso de este tipo de metodologías, donde sea posible hacerlo, en los programas que financia. Una posibilidad es, como en el caso de esta evaluación, la realización de experiencias “especiales”, tipo pilotos, que guardando todas las características del Programa, permitan una asignación aleatoria de beneficiarios y controles.

REFERENCIAS

- A. Woodbury, Stephen and Robert G. Spiegelman (1987). "Bonuses to Workers and Employers to Reduce Unemployment: Randomized Trial in Illinois". *The American Economic Review*, Vol. 77. N° 4, pp. 513-530.
- Baker, Judy (2000). "Evaluación del Impacto de los Proyectos de Desarrollo en la Pobreza: Manual para Profesionales". Washington, D. C.: Banco Mundial.
- Bloom, Howard, et al (2002). "Can Nonexperimental Comparison Group Methods Match the Findings from a Random Assignment Evaluation of Mandatory Welfare-to-Work Programs?" MDRC Working Papers on Research Methodology.
- Bloom, Howard (1984). "Accounting for No-Shows in Experimental Evaluation Designs". *Evaluation Review*, vol.8, n. 2.
- Boruch, Robert F. (1997). "Randomized Experiment for Planning and Evaluation". Applied Social Research Methods Series, volume 44. Thousand Oaks: SAGE Publications.
- Burtless, Gary (1995). "The Case for Randomized Field Trials in Economic and Policy Research". *The Journal of Economic Perspectives*, Vol. 9, N° 2, pp.63-84.
- Coady, David, X. Dai and L. Wang (2001). "Community Programs and Women's Participation: The Chinese Experience". Washington, D. C.: International Food Policy Research Institute.
- Fertig Michael y Jochen Kluve (2004). "A conceptual Framework for the evaluation of Comprehensive Labor Market Policy Reforms in Germany". IZA, discussion paper N° 1099.

- Fraker, Thomas M. and Rebecca A. Maynard (1987). "The Adequacy of Comparison Group Designs for Evaluations of Employment Related Programs," *The Journal of Human Resources*, Vol. 22, No. 2 (Spring): 194-227.
- Galasso, Emanuela, M. Ravallion and A. Salvia (2001). "Assisting the Transition from Workfare to Work: Argentina's Proempleo Experiment". Washington, D. C.: Banco Mundial.
- Galiani, Sebastian y E. Schargrodsky (2004). "Effects of Land Titling". Working Paper version preliminar.
- Gertler, Paul, J. Newman and L. Rawlings (1994). "Using Randomized Control Designs in Evaluating Social Sector Programs in Developing Countries". *The World Bank Research Observer*, Vol. 9, No. 2, pp. 181-201.
- Heckman, James, H. Ichimura, J. Smith and P. Todd (1999). "Characterizing Selections Bias Using Experimental Data". *Econometrica*, Vol. 66, N° 5, pp 1017-1098.
- Heckman, James, J. Smith and C. Taber (1998). "Accounting for Desertoers in Evaluations of Social Program". *The Review of Economics and Statistics*, Vol. 80, N° 1, pp. 1-14.
- Heckman, James, J. Smith and N. Clements (1997). "Making the Most Out of Programme Evaluations and Social Experiments: Accounting for Heterogeneity in Programme Impacts". *The Review of Economic Studies*, Vol. 64, N° 4, pp. 487-535.
- Heckman, James and J. Smith (1997). "The Sensitivity of Experimental Impact Estimates: Evidence from the National JTPA Study". NBER working paper 6105.
- _____ (1995). "Assessing the Case for Social Experiments". *The Journal of Economic Perspectives*, Vol. 9, N° 2, pp. 85-110.

- Heckman, James H. and V. Joseph Hotz (1989) "Choosing Among Alternative Nonexperimental Methods for Estimating the Impact of Social Programs: The Case of Manpower Training," *Journal of the American Statistical Association*, Vol. 84, No. 408, pp. 862-874.
- IFPRI (International Food Policy Research Institute) (2002). "Sistema de Evaluación de la Fase Piloto de la Red de Protección Social de Nicaragua," Washington, D.C.: Informe Final.
- IFPRI (2001a). "Sistema de Evaluación de la Fase Piloto de la Red de Protección Social de Nicaragua," Washington, D.C.: Informe Final.
- IFPRI (2001b). "Sistema de Evaluación de la Fase Piloto de la Red de Protección Social de Nicaragua: Línea de Base 2000". Washington, D.C.: Informe Final.
- Jaramillo Miguel y S. Parodi (2003). "Jóvenes Emprendedores: Evaluación de Programas de Promoción". Instituto APOYO.
- Kima, Jooseop, H. Aldermanb and P. Orazema (1998). "Can Private Schools Subsidies Increase Schooling for the Poor? The Quetta Urban Fellowship Program". Washington, D. C.: Banco Mundial.
- La Londe, Robert J. (1986) "Evaluating the Econometric Evaluations of Training Programs with Experimental Data," *The American Economic Review*, Vol. 76, No. 4, pp. 604-620.
- Mallar et al (1980). "The Short Term Economic Impact of the Job Corps Program". En Ernst Stromsdorfer y G. Farkas (eds), *Evaluation Studies Review Annual*, Vol. 5.
- Pradhan, Menno, L. Rawlings, and G. Ridder (1998). "The Bolivian Social Investment Fund: An analysis of Baseline Data for Impact Evaluation". Washington, D. C.: Banco Mundial.

Prennushi, Giovanna, G. Rubio y K. Subbarao (2000). “Seguimiento y Evaluación”. Washington, D. C.: Banco Mundial. Mimeo.

Yap, Yoon-Tien, G. Sedlacek Peter and F. Orazem (2001). “Limiting Child Labor Through Behavior-Based Income Transfers: An Experimental Evaluation of the PETI Program in Rural Brazil”. Washington, D. C.: Banco Mundial.