

Metodología cuantitativa de la evaluación del impacto

Ponce, Carmen; Escobal, Javier

Postprint / Postprint

Sammelwerksbeitrag / collection article

Empfohlene Zitierung / Suggested Citation:

Ponce, C., & Escobal, J. (2016). Metodología cuantitativa de la evaluación del impacto. In J. Escobal, & C. Ponce (Eds.), *Combinando protección social con generación de oportunidades económicas: una evaluación de los avances del programa Haku Wiñay* (pp. 35-65). Lima: GRADE Group for the Analysis of Development. <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-51455-4>

Nutzungsbedingungen:

Dieser Text wird unter einer CC BY-NC Lizenz (Namensnennung-Nicht-kommerziell) zur Verfügung gestellt. Nähere Auskünfte zu den CC-Lizenzen finden Sie hier: <https://creativecommons.org/licenses/by-nc/4.0/deed.de>

Terms of use:

This document is made available under a CC BY-NC Licence (Attribution-NonCommercial). For more information see: <https://creativecommons.org/licenses/by-nc/4.0>

3. METODOLOGÍA CUANTITATIVA DE LA EVALUACIÓN DE IMPACTO

Carmen Ponce y Javier Escobal

El componente cuantitativo de este estudio evalúa el impacto inicial del proyecto Haku Wiñay en un conjunto de áreas: la calidad de vida en el hogar —prácticas saludables, seguridad alimentaria, manejo de residuos sólidos, entre otros—, el sistema de producción familiar —tecnologías de producción agrícola y forestal, manejo de animales menores, entre otros—, la generación de ingresos autónomos —actividades independientes de tipo agropecuario, servicios, comercio, artesanías, entre otros—, y las capacidades de uso y manejo de instrumentos financieros.

En particular, el estudio se centra en identificar el impacto inicial que el proyecto ha tenido en hogares usuarios del programa Juntos. Cabe notar que, en la medida en que todos los hogares estudiados han sido intervenidos por Juntos, cuando mencionamos genéricamente el programa/tratamiento/intervención, nos referimos al proyecto Haku Wiñay. Con el fin de obtener estimadores consistentes (insesgados) del efecto del programa sobre la población intervenida, se aplicó un diseño experimental y estimadores de doble diferencia. En esta sección presentamos el marco metodológico utilizado, así como una breve explicación del diseño experimental implementado.

En segundo lugar, se discuten potenciales problemas derivados del tamaño pequeño de la muestra, de imperfecciones en la asignación programada del tratamiento, de la heterogeneidad de la intervención y de la pérdida de hogares entre la primera y segunda visita (*attrition*).

3.1. Estrategia metodológica para la estimación de impactos

Como se mencionó, el objetivo del análisis cuantitativo de este estudio es identificar los impactos iniciales del proyecto Haku Wiñay en hogares que, además, han sido usuarios del programa Juntos. Para ello, se analizó un conjunto de «variables resultado» de interés en las que el programa podría incidir, ya sea intencionalmente —como objetivo de Haku Wiñay— o como efecto colateral de sus actividades. Estas variables resultado incluyen indicadores asociados a seguridad alimentaria del hogar, aplicación de prácticas saludables en la vivienda, nivel y grado de diversificación de ingresos del hogar, conocimiento de servicios financieros, expectativas de los hogares y percepciones sobre el futuro, entre otros.

La pregunta que se busca responder es cuánto ha mejorado —o se ha deteriorado—, *por efecto del programa*, una variable resultado específica. Para ello, comparamos el desempeño observado de esa variable resultado con el desempeño que habría tenido si este mismo hogar no hubiera sido intervenido por el programa. El reto metodológico radica en que no conocemos este segundo escenario, en el cual el hogar, que en realidad sí fue intervenido, no lo hubiera sido.

Como es usual en la literatura sobre evaluación de impacto, usamos el concepto de resultados potenciales, desarrollado por Rubin (1974), para referirnos a cada escenario alternativo. Definamos Y_t^d como la variable resultado de un hogar en el escenario potencial d . Así, Y_t^P es la variable resultado en el periodo t en el escenario en el cual el hogar participa en el programa; mientras Y_t^{NP} es la variable resultado en ese mismo periodo si se asume un escenario diferente, en el cual el hogar no participa en el programa. t denota el periodo de referencia, 0 para el periodo previo a la intervención y 1 para el periodo posterior al inicio de la intervención. El efecto que buscamos estimar es el efecto promedio del tratamiento sobre la población que participó en el programa, denominado usualmente ATET, siglas en inglés de *average treatment effect*.¹³

13 Como señala Lechner (2010), la ecuación (1) no asume forma funcional específica para la variable resultado ni descarta potenciales heterogeneidades en el impacto estimado. También es importante notar

$$D = E [Y_i^P - Y_i^{NP} | T] \quad (1)$$

T indica que el grupo de hogares al que se refiere este valor esperado es el de aquellos que fueron tratados o intervenidos por el proyecto Haku Wiñay. Como se mencionó, el reto metodológico radica en que el resultado potencial Y_i^{NP} es desconocido para el grupo de hogares tratados (T).

Como señala la literatura de evaluación de impacto, necesitamos construir un grupo control, también llamado contrafactual. El grupo control es un conjunto de hogares que no participan en el programa y cuyas variables resultado muestran —en promedio— el mismo desempeño que hubieran mostrado las variables resultado del grupo tratado en un escenario potencial de no intervención. Si C representa al grupo control ideal, la ecuación (1) podría ser estimada a partir de la información del grupo control en el escenario de no intervención (escenario real):

$$\hat{D} = \hat{E} [Y_i^P | T] - \hat{E} [Y_i^{NP} | C] \quad (2)$$

Sin embargo, como señalan Duflo, Glennerster y Kremer (2008), construir un grupo control adecuado es un reto complejo. Para hacer evidente el sesgo que generaría utilizar un grupo control inadecuado, sumamos y restamos el valor desconocido de $E [Y_i^{NP} | T]$ a (2):

$$= E [Y_i^P | T] - E [Y_i^{NP} | C] + E [Y_i^{NP} | T] - E [Y_i^{NP} | T]$$

Reagrupando,

$$= \underbrace{E [Y_i^P | T] - E [Y_i^{NP} | T]}_{\text{ATET}} + \underbrace{E [Y_i^{NP} | T] - E [Y_i^{NP} | C]}_{\text{Sesgo de selección}}$$

que (1) representa el efecto agregado del programa sobre la variable resultado —en contraposición al efecto marginal, para lo cual requeriríamos la función de producción de cada variable resultado— (Duflo, Glennerster y Kremer 2008).

Como se observa en esta última ecuación, cuando el grupo control no representa adecuadamente al grupo tratado en el escenario de no intervención, se genera un sesgo en la estimación; es decir, al menos una parte del efecto estimado no es atribuible al programa. Este sesgo aparece cuando las variables resultado del escenario de no intervención dependen de si el hogar forma parte del grupo tratado (T) o del grupo control (C). Este sesgo de selección se puede deber a características observables o no observables, que afecten tanto la probabilidad de ser parte de T o C como el valor de las variables resultado. Por ejemplo, el sesgo podría llevar a una sobreestimación de los impactos del programa sobre los ingresos por actividad independiente, si se debiera a que los hogares tratados son —en promedio— menos adversos al riesgo, o tienen mayor acceso a redes extraterritoriales que podrían favorecer conexiones a mercados más rentables. Por el contrario, el sesgo podría llevar a una subestimación de los impactos del programa sobre la actividad agropecuaria si se debiera a que las tierras de los hogares control son, en promedio, más fértiles o tienen mayor acceso al agua.

El supuesto fundamental que permite afirmar que el estimador ATET es insesgado es el denominado *supuesto de exogeneidad del tratamiento*. Este supuesto exige que no existan diferencias sistemáticas en factores observables o no observables que expliquen tanto la pertenencia a T o a C (selección) como el valor de la variable resultado. Como señalan Duflo, Glennerster y Kremer (2008), no es posible medir el sesgo de selección —pues no se conoce el resultado potencial de los hogares tratados en un escenario de no intervención $E[Y_i^{NP} | T]$ —, pero sí es posible eliminarlo si se asigna de manera aleatoria la intervención entre hogares elegibles, conformando grupos de hogares tratados y controles comparables. En una muestra suficientemente grande, en la cual el supuesto de exogeneidad del tratamiento se cumple, (2) converge a (1).¹⁴

14 Como se detalla posteriormente, en el marco de esta evaluación de impacto se implementó una asignación aleatoria del programa a nivel de centros poblados. Más adelante se describen los retos asociados a esta implementación y los ajustes econométricos realizados para minimizar posibles sesgos derivados del tamaño pequeño de la muestra.

En una versión menos estricta, el supuesto de exogeneidad o de independencia condicionada en características observables asume que el sesgo de selección se debe enteramente a la diferencia en características observables —y, por tanto, en principio, puede ser eliminado en el proceso de estimación—. Sin embargo, es importante resaltar que esta versión del supuesto de independencia condicional asume que en no observables no hay diferencias que generen sesgos en la estimación (2). Si bien el diseño aleatorio de la muestra permite verificar el balance de un conjunto de características observables que podrían generar el sesgo de selección mencionado,¹⁵ en muestras pequeñas es posible que subsistan diferencias sistemáticas en características no observables —y por ello no detectables— que, a su vez, influyan en las variables resultado potenciales, contribuyendo a la aparición del sesgo de selección. Para enfrentar este problema, se diseñó una encuesta que recogió información de antes y después de que se implementara el programa, en el 2013 y el 2015, respectivamente. De esta manera, es posible hacer una estimación de diferencias en diferencias, que tiene la ventaja de reducir sustancialmente potenciales sesgos de selección al eliminar sesgos producidos por diferencias en características —observables y no observables— invariables en el tiempo o que varían en plazos mayores que los de este estudio. El estimador de diferencias en diferencias del efecto promedio del tratamiento sobre la población intervenida, ATET, es el siguiente:

$$DD = [\hat{E} [Y_t^P | T] - \hat{E} [Y_0^{NP} | T]] - [\hat{E} [Y_t^{NP} | C] - \hat{E} [Y_0^{NP} | C]] \quad (3)$$

Como señalan Duflo, Glennerster y Kremer (2008), (3) permite obtener un estimador sin sesgos siempre que se cumpla el supuesto de tendencias paralelas.¹⁶ Según este supuesto, si el programa no se hubiera implementado, ambos grupos de hogares —controles y tratados— habrían

15 Como se explica después, aunque una muestra aleatoria pequeña puede mostrar desbalances entre el grupo tratado y el grupo control en algunas características observables, es posible estimar pesos que permitan corregir estos desbalances, de manera similar a como se corrigen en las encuestas de hogares con diseños complejos, como las que realiza el INEI.

16 $\hat{E} [Y_0^{NP} | T]$ es estimable debido a que en $t = 0$ (2013) el programa todavía no se había implementado, por lo que Y_0^{NP} es la variable resultado observada de los hogares del grupo tratado en ese año.

experimentado el mismo cambio en sus variables resultado promedio. Esto se ve con mayor claridad a continuación. Si el programa no se hubiera implementado, los hogares del grupo tratado habrían experimentado el siguiente cambio en sus variables resultado:¹⁷

$$E [Y_t^{NP} | T] - E [Y_0^{NP} | T]$$

Sumando y restando $E [Y_t^{NP} | T]$ en (3), obtenemos:

$$\begin{aligned} \widehat{DD} = & \underbrace{[\hat{E} [Y_t^P | T] - \hat{E} [Y_t^{NP} | T]]}_{\text{ATET}} + \\ & \underbrace{[\hat{E} [Y_t^{NP} | T] - \hat{E} [Y_0^{NP} | T]] - [\hat{E} [Y_t^{NP} | C] - \hat{E} [Y_0^{NP} | C]]}_{\text{Sesgo}} \end{aligned}$$

Como es evidente en esta ecuación, si las tendencias en un escenario de no intervención fueran las mismas —en promedio— entre los hogares del grupo tratado y del grupo control, el sesgo sería nulo y el efecto estimado sería, efectivamente, el impacto promedio del programa sobre el grupo tratado: ATET, (1). Así, el estimador DD relaja parcialmente el supuesto de exogeneidad, al eliminar el efecto que tendrían las características del hogar invariantes en el tiempo sobre las variables resultado.

Adicionalmente al supuesto asociado al sesgo de selección —cuya plausibilidad es razonable cuando se aleatoriza la asignación del tratamiento entre controles y tratados—, se requiere un segundo supuesto para estimar de manera consistente el efecto tratamiento promedio sobre los tratados (ATET) de la ecuación (1). Este segundo supuesto es el que sostiene que la variable resultado de un hogar, en cualquiera de los escenarios potenciales, no se ve afectada por la condición de participante o no participante de otro hogar —usualmente, este supuesto es denominado SUTVA, siglas en inglés de *stable unit treatment value*—. Este supuesto implica que la

¹⁷ Denotamos el escenario de no intervención en $t = 1$ (2015) en color gris, para enfatizar que se trata de un escenario potencial.

intervención no genera externalidades que afecten las variables resultado de otros hogares. Si este supuesto no se cumple —debido a que el grupo control ha sido afectado al menos parcialmente por el tratamiento, vía entrenamiento por parte de un hogar tratado, por ejemplo—, el valor observado no representa la variable resultado de los tratados en el escenario potencial de no intervención. Al respecto, Lechner (2010) señala que, en una investigación previa en la que se analizó el fenómeno de externalidades en el tratamiento, Miguel y Kremer (2004) sugirieron que, cuando es posible implementar diseños experimentales, es más adecuado aleatorizar grupos que aleatorizar individuos —u hogares—. Esta es una de las razones por las que el diseño de este estudio aleatorizó centros poblados.

Lechner (2010: 178) señala que, en el marco del estimador de diferencias en diferencias (3), un supuesto adicional se desprende del SUTVA: los hogares tratados no modifican su comportamiento, durante el periodo de preintervención, ante la expectativa de que se implemente el programa; es decir, $\hat{E} [Y_0^{NP} | T]$ coincide con los valores poblacionales $E [Y_0^{NP} | T]$ y $E [Y_0^{NP} | C]$. Por ello, es importante que los diseños experimentales restrinjan el anuncio de la implementación de un programa hasta después de haber realizado la encuesta de línea de base.

3.2. Diseño experimental y selección de la muestra

Esta evaluación se sustenta en un diseño experimental coordinado conjuntamente con FONCODES y la Dirección de Evaluación del MIDIS. El conjunto de centros poblados que constituyen el marco muestral —utilizado para configurar la muestra de hogares que iban a ser encuestados— fue seleccionado en dos etapas. En primer lugar, se identificó el conjunto de centros poblados considerados elegibles por el programa y cuya intervención se planeaba iniciar en el periodo 2013-2016. En segundo lugar, se establecieron criterios adicionales que facilitarían la identificación de los impactos del programa —se eliminó centros poblados en zonas de influencia de la gran minería, para minimizar externalidades asociadas que podrían sesgar los

impactos identificados— y maximizaran la probabilidad de encontrar a los hogares de interés —hogares elegibles para el programa que, a su vez, fueran usuarios del programa Juntos en febrero del 2013—. Sobre la base de estos criterios se configuró el marco muestral de la encuesta, compuesto por los 230 centros poblados que cumplían con tales criterios en los departamentos de Cajamarca, Huánuco y Huancavelica.

Con estos centros poblados se construyeron parejas de centros poblados, tomando como criterio de emparejamiento la similitud de las características asociadas a los ámbitos productivo, económico y social. Luego, se asignó aleatoriamente un centro poblado de cada pareja al grupo de tratados, mientras que el otro fue asignado al grupo control. Para evitar potenciales externalidades o efecto contagio entre uno y otro centro poblado,

Tabla 3.1
Muestra de hogares en las encuestas de línea de base y revisita

Departamento	Provincia	Distrito	Centros poblados encuestados	Asignación al tratamiento	Número de hogares encuestados en ambas operaciones de campo**
Huancavelica	Acobamba	Anta	6	Tratado	66
		Rosario	2	Control	23
		Andabamba	4	Control	47
Huánuco	Panao	Umari	6	Tratado	79
		Panao	6	Control	76
Cajamarca	Chota	Chalamarca	6*	Tratado	62*
		Conchán	3	Control	38
		Paccha	3	Control	37
Total	3	8	36*		428

* Un centro poblado fue eliminado del análisis debido a que, a pesar de haber sido asignado al grupo de tratados, no aparece en el listado de FONCODES de centros poblados donde operó el programa durante este periodo, y los 11 hogares encuestados que residen allí reportaron no haber participado en el programa ni tener vecinos que lo hubieran hecho. El centro poblado pareja que se le había asignado fue incorporado a otra pareja de centros poblados, como centro poblado control.

** Las implicancias de la deserción de algunos hogares de este estudio (*attrition*) son analizada en la siguiente subsección.

se cuidó de que aquellos que conformaban cada pareja pertenecieran a distintos distritos.

Cabe resaltar que la asignación aleatoria al tratamiento implicó, en realidad, la asignación al grupo de centros poblados que serían intervenidos en el 2013, mientras que los centros poblados del grupo control serían intervenidos en el 2016, cuando esta evaluación ya hubiera finalizado.

La tabla 3.1 muestra la distribución de hogares visitados en febrero del 2013 en la encuesta de línea de base, y revisitados entre agosto y setiembre del 2015, después de la implementación del programa en los centros poblados tratados.

Cabe señalar que las estimaciones de los efectos del programa presentadas en la siguiente sección incorporaron ajustes en los errores estándar para dar cuenta del diseño muestral de la encuesta. Para ello, se tomó en cuenta que el muestreo al interior de cada departamento se hizo de manera independiente, que la primera etapa de muestreo seleccionó parejas de centros poblados —tratado y control— y que este muestreo fue hecho sin reemplazo, por lo que se incorporó un factor de corrección por población finita.¹⁸ El factor de corrección fue calculado como la proporción de centros poblados tratados seleccionados para la encuesta con relación al número de centros poblados elegibles para el programa en cada departamento. Este factor de corrección calculado para cada centro poblado tratado fue asignado al centro poblado control identificado como pareja en el diseño original.

3.3. Potenciales problemas en la implementación del diseño experimental

Como señalan Duflo, Glennerster y Kremer (2008), los diseños experimentales usualmente enfrentan un conjunto de problemas en el proceso de implementación. Este estudio no es una excepción. Los autores sugieren

18 Se usa el factor de corrección por población finita en contextos de muestreo sin reemplazo, como el de este estudio. Este factor de corrección da cuenta de la reducción en la varianza, como producto de un muestreo sin reemplazo en términos relativos a uno con reemplazo realizado sobre la misma población.

algunas estrategias para recuperar la consistencia interna del estimador de impacto. Aquí presentamos los principales problemas enfrentados en el marco de este proceso de evaluación, así como las soluciones aplicadas.

3.3.1. Potencial sesgo por tamaño pequeño de la muestra

En principio, el diseño experimental de la intervención asegura el balance entre hogares tratados y hogares control, en la medida en que se selecciona aleatoriamente el grupo de centros poblados intervenidos en el 2013 y el grupo de centros poblados control que serán intervenidos después del 2015. En consecuencia, el diseño muestral de la encuesta debería recoger submuestras balanceadas de hogares control y hogares tratados. Sin embargo, debido a que la muestra seleccionada es pequeña y la evaluación se hace a nivel de hogares, se detectaron algunos desbalances en características observables que podrían generar sesgos en la evaluación de impacto sobre determinadas variables resultado. Estos desbalances fueron ajustados para eliminar potenciales problemas de sesgos de selección.

Como es conocido, la literatura de evaluación de impacto ha desarrollado un conjunto de metodologías que buscan asegurar el cumplimiento del supuesto de independencia condicional del tratamiento en contextos de evaluación a) en los cuales el diseño experimental no es perfecto —y no permite asegurar el balance perfecto entre los tratados y los controles— o b) en los que, en ausencia de diseños experimentales, se recurre a experimentos naturales o diseños cuasiexperimentales. Estas metodologías incluyen el emparejamiento individual por características específicas, el emparejamiento utilizando la propensión a ser tratado (*propensity score matching*), el emparejamiento genético (*genetic matching*), el diseño de regresión discontinua (*discontinuity regression design*), y el balance entrópico, entre otras (Imbens y Wooldridge 2009; Hainmueller 2012; Duflo, Glennerster y Kremer 2008). En este estudio utilizamos esta última metodología.

Hainmueller (2012) propone, desde la ciencia política, una metodología de estimación de pesos para el grupo control que permite igualar determinados momentos de la distribución —media, varianza, kurtosis— de las

características observables de los dos grupos —tratados y controles—. ¹⁹ De esta manera, se busca asegurar la ortogonalidad entre el tratamiento y tales características. Hainmueller (2012: 31) señala que el balance entrópico puede ser entendido como una suerte de generalización del método convencional de estimación de ponderadores asociados a la propensión de un hogar a ser tratado (*propensity score*). Mientras este último método procede en dos etapas —en la primera etapa, estima pesos individuales a partir de una regresión logística de la probabilidad de ser tratado; y en la segunda, evalúa el balance entre las dos muestras (ponderadas por los pesos) para asegurar que estos igualan la distribución de las características de los hogares (covariables)—, el balance entrópico se concentra desde el inicio en el problema de balance de las características de los hogares.

El balance entrópico planteado por Hainmueller (2012) consiste en estimar el efecto tratamiento (3) a partir de la diferencia entre el promedio ponderado de variables resultado de controles y tratados —en vez de la diferencia entre promedios simples—. Los ponderadores o pesos utilizados para ese promedio ponderado se estiman únicamente para los hogares control —los hogares tratados mantienen un peso de 1—, siguiendo el siguiente esquema de optimización (Hainmueller 2012: 30):

$$\min_{w_i} H(w) = \sum_{i/C} w_i \log(w_i / q_i)$$

Sujeto a

$$\sum_{i/C} w_i c_{ri} (X_i) = m_r, \text{ con } r \in 1, \dots, R \quad (4)$$

$$\sum_{i/C} w_i = 1$$

$w_i \geq 0$ para todo i que pertenece al grupo control

¹⁹ Estos pesos pueden usarse para ajustar el diseño muestral a encuestas que buscan ser representativas de determinados parámetros poblacionales conocidos, construyendo pesos que minimizan la distancia entre el parámetro muestral y el parámetro poblacional.

Cabe notar que, si bien la métrica de distancia utilizada en la función objetivo puede ser otra, Hainmueller (2012) propone usar la de divergencia entrópica que Kullback planteó en 1959. El objetivo, entonces, es minimizar la distancia entre el vector de pesos estimados w_i y el de pesos base q_i ($q_i \geq 0$) para todos los hogares del grupo control, y $\sum_{i \in C} q_i = 1$, $q_i = 1/n_C$, donde n_C representa el número de hogares control). La primera restricción es la asociada al balance de los momentos de la distribución entre tratados y controles, $c_{ri}(X_i) = X_{ij}^r$, donde r representa el momento de la distribución de la variable X_j que se busca balancear. Por ejemplo, si se busca balancear el primer momento de la distribución (la media) de una sola variable X_j , $\sum_{i \in C} w_i c_{ri}(X_i) = \sum_{i \in C} w_i X_{ij}$ y $m_r = \sum_{i \in T} X_{ij} / n_T$, donde n_T es el número de hogares tratados y m_r representa aquí el promedio de la variable X_j en el grupo de hogares tratados. Así, el método permite balancear las muestras en un conjunto de características manteniendo la mayor cercanía posible de los pesos a los originalmente prevalentes en la encuesta, para conservar, en la medida de lo posible, la eficiencia en la estimación de los impactos del programa (Hainmueller 2012: 31).

Hainmueller resalta la ventaja operativa de este método frente al de estimación de ponderadores o pesos asociados a la propensión de un hogar a ser tratado (*propensity score*), en primer lugar, porque ajusta exactamente los momentos de la distribución de un conjunto de características. En ese sentido, no requiere la verificación iterativa del balance entre tratados y controles por bloques que la metodología de *propensity score matching* sí exige en el proceso de búsqueda de balance entre las características de las tratados y controles, búsqueda que, en muchos casos, es poco exitosa (Hainmueller 2012, Hainmueller y Xu 2014). En segundo lugar, el autor muestra, a partir de simulaciones, que el método sería más exitoso que métodos alternativos en reducir problemas de dependencia de los resultados a especificaciones alternativas del modelo (*model dependence*).

No obstante las ventajas operativas de este método, es importante notar que la identificación de las características cuyo balance es prioritario para eliminar problemas potenciales de sesgo de selección en la evaluación de impactos queda en manos del investigador. En el caso de estudios basados

en una muestra pequeña, el conjunto de características que pueden ser balanceadas es reducido, por lo que la identificación de variables clave puede constituir un reto complejo. Por ello, es importante estudiar la significancia estadística de las diferencias sin y con pesos para el conjunto de variables que no pueden ser incorporadas en la estimación de balance entrópico. Afortunadamente, en este estudio contamos con un diseño experimental previo, por lo que el desbalance entre tratados y controles no es muy grande.

Para efectos de este estudio, seleccionamos un conjunto de variables que podrían influir en el nivel y en los cambios de las variables resultado de la evaluación, tanto en variables asociadas a actividades económicas como a indicadores de seguridad alimentaria y prácticas cotidianas de manejo de la vivienda. Las variables seleccionadas incluyen años de educación formal y sexo del jefe de hogar. Además, incluimos el departamento de residencia del hogar, que da cuenta de diferencias en características regionales como, por ejemplo, el hecho de que, en Cajamarca —mayoritariamente hispanohablante—, la lengua materna del jefe del hogar o de su cónyuge no tiende a reflejar las diferencias en el acceso a servicios o exposición a procesos de exclusión, lo que sí sucede en otros departamentos del Perú, como Huánuco y Huancavelica. Adicionalmente, se aseguró el balance de algunas variables de contexto que dan cuenta del tipo de dinámica económica y tecnologías predominantes en el centro poblado del hogar antes de la intervención. Para ello, se tomó información proveniente del Censo Agropecuario 2012 a nivel de centro poblado, incluyendo el valor promedio del ganado en las unidades agropecuarias, así como la proporción de unidades agropecuarias a) en las que se contrata a trabajadores remunerados, b) que cuentan con algún tipo de riego en por lo menos una de sus parcelas, y c) cuyo conductor o conductora reporta que la actividad agropecuaria le permite generar suficientes ingresos para la familia. En la tabla 3.2 se muestran las diferencias entre características promedio de los hogares tratados y control antes y después de utilizar los ponderadores resultantes del balance entrópico.²⁰ Cabe señalar que la restricción de

20 La estimación se hizo en Stata utilizando el programa *ebalance* desarrollado por Haimnueller y Xu (2013) para Stata.

Tabla 3.2
Diferencia entre las características
de centros poblados tratados y controles

Características previas a la intervención	Muestra sin pesos			Muestra con pesos		
	Hogares en centros poblados controles	Hogares en centros poblados tratados	Sig ^{1/}	Hogares en centros poblados controles	Hogares en centros poblados tratados	Sig ^{1/}
Características de los hogares de la muestra						
Hogares en Cajamarca (%)	33,9	30,0		30,0	30,0	
Hogares en Huánuco (%)	34,4	38,2		38,1	38,2	
Años de educación del jefe de hogar	4,6	4,0	*	4,0	4,0	
Número de miembros del hogar	4,8	4,9		4,6	4,9	*
Hogares en cuyo centro poblado opera alguna ONG (%)	27,2	64,4	***	64,6	64,4	
Hogares cuyo jefe de hogar se dedica a la actividad agrícola (%)	56,9	65,9	***	65,9	65,9	
Hogares cuya jefa de hogar es mujer (%)	12,2	19,8	**	19,8	19,8	
Edad del jefe de hogar	41,5	42,0		40,9	42,0	
Indicadores de contexto del CENAGRO 2012 (agregados a nivel de centro poblado)						
Hogares que cuentan con teléfono (%)	18,2	20,3	*	11,5	20,2	*
Hogares con algún miembro que migra para obtener otros ingresos (%)	32,9	23,4	***	28,9	23,4	
Unidades agropecuarias en las que se ha contratado a trabajadores remunerados (%)	59,4	48,0	***	48,0	47,9	
Unidades agropecuarias que cuentan con algún tipo de riego (%)	18,6	19,6		19,5	19,6	
Promedio del valor del ganado (a precios de 1994)	1029	952	**	952	952	
Hogares que reportan que la actividad agropecuaria les produce ingresos suficientes (%)	20,3	25,7	**	25,7	25,7	
Número de hogares	221	207		221	207	

^{1/} Significancia estadística de la diferencia. ***p<0,01, **p<0,05, *p<0,1.

balance (4) introducida en la estimación se concentró en la igualdad de medias entre tratados y controles.

Como se observa en la tabla 3.2, el balance entrópico permite balancear no solo las características incorporadas en la restricción (4), sino además otras variables que consideramos que también pueden afectar a las variables resultado, como, por ejemplo, la presencia de alguna ONG en el centro poblado —que podría afectar el tipo de apoyo que reciben los hogares y, eventualmente, el acceso a determinadas tecnologías que el programa ofrece—, la participación del jefe de hogar en actividades agrícolas como fuente de generación de ingresos; y en variables contextuales, la migración estacional como estrategia de los hogares del centro poblado para generar ingresos. Cabe señalar que, en el proceso de mejora del balance multivariado, dos características mantienen sus diferencias entre tratados y controles: el número de miembros del hogar, que presenta una diferencia pequeña (4,6 versus 4,9), y el acceso promedio a teléfono en el centro poblado en el 2012 (variable que proviene del Censo Agropecuario). Sin embargo, las ganancias en balance entre controles y tratados en los frentes productivo, de diversificación de ingresos, de características demográficas y de capital humano son sustantivas.

3.3.2. Imperfecciones en la asignación del tratamiento a nivel de hogares (de ATET a ITT)

Como se mencionó en el capítulo anterior, la participación en el proyecto Haku Wiñay es voluntaria, e involucra la contribución en tiempo, trabajo y, en algunos casos, activos por parte del hogar tratado. Por ello, como mencionamos al inicio de este capítulo, el supuesto de exogeneidad del tratamiento sería poco plausible si no se incorporara un diseño aleatorio en el proceso de asignación del tratamiento. Sin embargo, incluso en este escenario de diseño experimental, algunos hogares que pertenecen a centros poblados tratados y que, por tanto, podrían participar en el programa, pueden decidir no hacerlo. Como señalan Duflo, Glennerster y Kremer (2008), este es el caso en muchas evaluaciones de impacto con diseño

experimental, dado que los investigadores tienen poco control sobre lo que potenciales tratados y potenciales controles deciden hacer.

Debido a que el diseño de la evaluación aleatorizó la intervención a nivel de centros poblados, la tasa de potenciales controles que participa en el programa es 0. Sin embargo, existen potenciales tratados que decidieron no participar en el programa. En nuestra muestra, 86% de los hogares potencialmente tratados fueron efectivamente intervenidos por el programa; es decir, 29 de los 207 hogares de la muestra de centros poblados tratados no participaron en el programa. Aunque no es un problema masivo, es necesario tomar en cuenta las implicancias que esto puede tener en la interpretación del estimador (3).

En este caso, en el cual algunos hogares del grupo de centros poblados tratados no participan en el programa, la diferencia entre las variables resultado promedio de los hogares que residen en los centros poblados tratados y aquellos que residen en los centros poblados controles —descrita en (3)— no estima el efecto promedio sobre los hogares efectivamente tratados. En este contexto, la ecuación (3) permite obtener el estimador denominado *intención de tratamiento* (ITT), que es el efecto esperado sobre la población que *se planea* intervenir/beneficiar/atender —al interior de la cual tenemos hogares que decidirán participar y hogares que decidirán no participar—. Siguiendo a Duflo, Glennerster y Kremer (2008), sobre la base del estimador ITT es posible utilizar una aproximación de variables instrumentales para estimar el efecto promedio del programa sobre la población efectivamente intervenida (ATET).²¹ El instrumento que utilizamos es la pertenencia a un centro poblado aleatoriamente seleccionado para ser tratado.

Siguiendo a Duflo, Glennerster y Kremer (2008), dos supuestos son necesarios para utilizar como instrumento la asignación aleatoria de centros

21 En principio, en este contexto de participación imperfecta o parcial, el estimador que se deriva de ITT usando como instrumento la asignación aleatoria de centros poblados sería el denominado LATE (por sus siglas en inglés, *local average treatment effect*). LATE representa el efecto promedio del programa sobre la población efectivamente tratada —que podría incluir, en una evaluación como esta, a hogares de centros poblados tratados que efectivamente participen y a hogares de centros poblados control que también participen—. Sin embargo, debido a que no existe ningún hogar de centros poblados control que participe en el programa, el estimador LATE es, en realidad, el estimador ATET.

poblados para estimar el efecto local del tratamiento sobre la población efectivamente intervenida. En primer lugar, se requiere el supuesto de independencia, que implica que la comparación entre variables resultado expuestas a diferentes valores del instrumento identifica el impacto causal del instrumento. Siguiendo a Duflo, Glennerster y Kremer (2008), esto se cumple por construcción, dado que el instrumento es asignado en forma aleatoria. Adicionalmente, las variables resultado potenciales no deben ser afectadas directamente por el instrumento. Cabe recordar que este supuesto ha sido discutido en este capítulo, dado que el instrumento al que nos referimos es, justamente, la asignación aleatoria del programa entre centros poblados tratados y controles; en el acápite anterior se muestra el ajuste hecho para cumplir con el supuesto de independencia condicional en observables vía la metodología de balance entrópico. El segundo supuesto requerido es el de monotonicidad, que exige que la probabilidad de participar en el programa sea igual o mayor entre hogares expuestos al instrumento que entre hogares no expuestos al instrumento. Este supuesto se cumple en la medida en que la probabilidad de participar en el programa para aquellos que no residen en centros poblados tratados es nula, mientras que la probabilidad de participación entre hogares que pertenecen a centros poblados tratados es mayor que 0, y bastante alta, dada la tasa de participación de 86% en centros poblados tratados.

Cabe notar que este paso del estimador ITT al estimador ATET, utilizando como instrumento la asignación aleatoria de centros poblados, puede generar sesgos importantes si existen externalidades en el tratamiento que afecten a las variables resultado de los hogares no participantes que residen en centros poblados tratados. Para indagar sobre posibles externalidades, en la encuesta a los hogares que decidieron no participar se preguntó si conocían a usuarios del programa. El resultado fue que 22 de los 29 hogares que no participaron reportaron conocer a algún beneficiario del programa; 16 de ellos dijeron que estos eran familiares o parientes suyos. Estos reportes sugieren que el efecto contagio podría ser alto, dada la cercanía de estos usuarios a los no participantes. Sin embargo, como se

muestra en la tabla 3.3, cuando se preguntó por cada uno de los rubros de intervención del programa, solo entre 0 y 4 personas reportaron haber recibido un beneficio indirecto de parte de un beneficiario del programa en cada uno de esos rubros —capacitación o activos—. Por ello, se consideró que el estimador ATET obtenido a partir del ITT, usando como instrumento la asignación aleatoria del tratamiento entre centros poblados, es consistente y, en el peor de los casos, tendría un sesgo muy pequeño.

Tabla 3.3
Beneficios indirectos del programa entre hogares de centros poblados tratados que no participaron en este

Tipo de apoyo recibido por parte de usuarios de Haku Wiñay	Número de hogares	%
Riego tecnificado	1	3
Huerto de hortalizas	4	14
Parcela de pastos	2	7
Agroforestería	0	0
Producción de abonos y plaguicidas orgánicos	2	7
Crianza de animales menores	4	14
Instalación o uso de cocina mejorada	4	14
Prácticas de higiene en el hogar	4	14
Manejo de residuos sólidos	4	14
Educación financiera	0	0
Ayuda en algún otro negocio	0	0
Total de hogares que no recibieron beneficios directos del programa	29	100

3.3.3. Intensidad del tratamiento (heterogeneidad de la intervención)

Imbens y Wooldridge (2009: 71) señalan que las intervenciones implementadas por programas sociales son usualmente heterogéneas, diseñadas con cierto grado de flexibilidad con el fin de que puedan adecuarse a las necesidades o demandas de la población tratada; así, en algunos casos, se interviene con mayor intensidad a unos participantes que a otros, y en

otros, se aplican tratamientos secuenciales. Como se señaló en el capítulo 2, el proyecto Haku Wiñay es un programa de intervenciones heterogéneas. Cuenta con varios componentes, en el marco de los cuales se desarrollan actividades de capacitación, asesoría técnica y/o transferencia de activos con el objetivo de mejorar las condiciones de vida y producción de hogares en situación de pobreza. Por un lado, hay un menú de tecnologías entre las cuales se privilegia algunas, de acuerdo con las preferencias y necesidades de los productores; por otro, las actividades de capacitación y transferencia de activos en el módulo de vivienda saludable buscan complementar y mejorar las condiciones de vida de los participantes del programa. Mientras el componente de servicios financieros es más homogéneo, el de concursos de negocios inclusivos es el que beneficia de manera más heterogénea a los participantes, en la medida en que no todos participan; y entre aquellos que participan, solo algunos ganan.

El riesgo fundamental cuando se evalúa el impacto de programas heterogéneos está asociado al supuesto de exogeneidad o independencia condicional en características observables. Si el tipo de tratamiento implementado depende de las características del hogar tratado, es fundamental que la estrategia de estimación controle por estas características en la búsqueda del grupo control ideal, para mantener la consistencia interna del estimador. Sin embargo, a diferencia del caso de tratamientos homogéneos, la literatura econométrica sobre estimadores de impacto de tratamientos heterogéneos es todavía limitada.²²

Dada la heterogeneidad del paquete de actividades de los hogares tratados,²³ y dado que la restricción común al paquete fue el monto máximo permitido por hogar, hicimos una estimación del equivalente monetario transferido a cada hogar. Sobre la base de esta escala monetaria de intensidad del tratamiento recibido por cada hogar, clasificamos a la población tratada en tres terciles. El cálculo del equivalente monetario fue realizado

22 Para mayor detalle sobre los avances en este frente, véase Imbens y Wooldridge (2009).

23 Para efectos de una metodología cuantitativa de evaluación de impactos, la heterogeneidad del tratamiento puede ser resumida de dos maneras: a partir de una variable discreta con múltiples valores, o a partir de una variable continua que homogenice la diversidad de tratamiento en una escala comparable.

combinando los reportes de los hogares encuestados sobre las transferencias recibidas de activos, capacitación y servicios con información proporcionada por FONCODES sobre costos promedio (aproximados) por actividad.

Sobre la base de esta estimación monetaria de intensidad del tratamiento, se exploró el impacto del programa en sus versiones de menor intensidad y de mayor intensidad. Adicionalmente, se exploró el impacto del programa en el grupo que ganó concursos de negocios rurales inclusivos. En la siguiente sección se discuten estos resultados, junto con los resultados para la muestra agregada. Cabe señalar que, con el fin de minimizar el riesgo de incurrir en sesgo de selección en la estimación del ATET de cada subgrupo, replicamos el ejercicio de balance de características observables vía la estimación de pesos entrópicos para cada subgrupo. Es decir, estimamos pesos muestrales que permitieran balancear la muestra de hogares control con cada subgrupo de hogares tratados: tercio de menor intensidad, tercio de mayor intensidad, hogares ganadores de concursos de negocios rurales inclusivos.²⁴ Como se observa en la tabla 3.4, el balance entrópico permite mejorar la comparabilidad de las submuestras en todas las características de contexto, excepto en la distribución de hogares residentes en Cajamarca. Asimismo, permite balancear características demográficas y de actividad productiva del hogar, excepto por el número de sus miembros. En la medida en que la evaluación se hace en un horizonte de dos años, se espera que el efecto de estas características se diluya en el estimador de diferencias en diferencias.

Es importante señalar que estos resultados para subpoblaciones específicas deben ser tomados con cautela, en la medida en que el diseño aleatorio de asignación del tratamiento no consideró esta heterogeneidad en el tipo de tratamiento aplicado a cada hogar —o disponible para cada centro poblado—. En ese sentido, estos resultados constituyen un análisis exploratorio de las heterogeneidades del tratamiento en el marco de un programa complejo y con componentes liderados por demanda.

24 En el anexo se presentan las tablas completas asociadas a la tabla 3.4.

Tabla 3.4
Significancia estadística de la diferencia entre controles y tratados
antes y después de aplicar ponderadores de balance entrópico

Características previas a la intervención	Tercio menos intervenido		Tercio más intervenido		Ganadores de concursos de negocios rurales inclusivos	
	Sin pesos	Con pesos	Sin pesos	Con pesos	Sin pesos	Con pesos
Características de los hogares de la muestra						
Hogares en Cajamarca		*			**	***
Hogares en Huánuco						
Años de educación del jefe de hogar						
Número de miembros del hogar		**		**		**
Hogares en cuyo centro poblado opera alguna ONG	***	*	***		***	
Hogares cuyo jefe de hogar se dedica a la actividad agrícola	***					
Hogares cuya jefa de hogar es mujer	**					
Edad del jefe de hogar						
Indicadores de contexto del CENAGRO 2012 (agregados a nivel de centro poblado)						
Hogares que cuentan con teléfono						
Hogares con algún miembro que migra para obtener otros ingresos	***		***	*	***	
Unidades agropecuarias que contratan a trabajadores remunerados	***		**		*	
Unidades agropecuarias con algún tipo de riego						
Promedio del valor del ganado (a precios de 1994)			*			
Hogares que reportan que la actividad agropecuaria les produce ingresos suficientes	**					
Número de hogares tratados	63		59			53

3.3.4. Sesgo por deserción (attrition)

Como se mencionó, la muestra de la revisita incluyó a 447 hogares distribuidos tal como se expone en la tabla 3.5. Luego de la revisita, se logró ubicar y encuestar a los representantes de 428 hogares; no se pudo entrevistar a 19 hogares que ya no vivían en los centros poblados incorporados en la muestra.

Tabla 3.5
Distribución de la muestra para la revisita

Departamento	Grupo control	Grupo tratamiento	Total
Cajamarca	78	65	143
Huancavelica	72	72	144
Huánuco	78	82	160
Total	228	219	447

Los 19 hogares que no pudieron ser revisitados representan el 4,3% de la muestra. De ellos, 11 pertenecen a centros poblados del grupo de tratamiento, y 7, a centros poblados del grupo de control. A pesar de que el número absoluto de hogares no ubicados es pequeño, cabe preguntarse si la omisión de ellos podría generar algún sesgo de selección. Para explorar este tema, en la tabla 3.6 mostramos las características más importantes de los hogares, divididos entre hogares entrevistados y hogares no entrevistados. Tal como se observa, las diferencias entre ambos grupos, en la línea de base, son pequeñas y estadísticamente no significativas. Solo para el caso de la extensión total de la superficie agrícola las diferencias son estadísticamente significativas y muestran que los no revisitados cuentan con mayor extensión agrícola. Sin embargo, en las demás características del jefe de hogar, de su vivienda y de sus parcelas no existe ningún patrón sistemático que valga la pena resaltar.

Con el fin de analizar la relevancia conjunta de los indicadores mencionados en la tabla 3.7 se estima un *probit* para evaluar el perfil de quienes no fueron entrevistados. La tabla 3.7 muestra que solo una de las variables

Tabla 3.6
Diferencia de medias de variables estructurales,
medidas en la línea de base, entre hogares entrevistados
y hogares no entrevistados en la revisita

Variables explicativas	Medias		Diferencia
	Hogares entrevistados	Hogares no entrevistados	
Jefa de hogar mujer (%)	15,9	10,5	-5,4
Número de años de educación del jefe de hogar	4,3	5,1	0,8
Número de miembros del hogar	4,8	4,4	-0,4
Edad del jefe de hogar	41,8	39,2	-2,6
El jefe de hogar o su cónyuge tienen el quechua como lengua materna (%)	61,9	57,9	4,0
Alumbrado por electricidad (%)	84,1	89,5	-5,4
Número de parcelas	2,7	2,0	-0,7
Superficie total (hectáreas)	1,1	1,8	0,7 **
Superficie que se encuentra bajo riego (%)	10,3	8,4	-1,9
Recibió asistencia técnica en cualquier tema (%)	23,8	10,5	-13,3
Recibió alguna capacitación (%)	53,3	42,1	-11,2
Hogares cuyo jefe se dedica a la actividad agrícola (%)	61,3	70,3	9,0
Hogares que reportan que la actividad agropecuaria produce ingresos suficientes (CENAGRO 2012) (%)	22,9	24,5	1,6
Unidades agropecuarias en las que se usa tractor (CENAGRO 2012) (%)	9,6	15,1	5,4

Nota: *** $p < 0,01$, ** $p < 0,05$, * $p < 0,1$.

—superficie total— es estadísticamente significativa al 99% y otras tres variables —número de parcelas, acceso a asistencia técnica y capacitación— son marginalmente significativas. A pesar de ello, la prueba de significancia conjunta de todas las variables incluidas muestra que no se puede rechazar la hipótesis de que la pérdida de hogares es, simplemente, aleatoria (prueba F con un valor de 0,26). De hecho, el estimado del grado de ajuste del modelo sugiere que el 85% de la varianza en la variable que registra a quienes no fueron entrevistados se puede atribuir a factores no directamente observables.

Tabla 3.7
Probit-perfil de quienes no fueron entrevistados en la revisita (*attrition*)

Variables explicativas	Coeficiente	
Jefa de hogar mujer	-0,49 (0,43)	
Número de años de educación del jefe de hogar	0,05 (0,05)	
Número de miembros del hogar	-0,12 (0,10)	
Edad del jefe de hogar	-0,01 (0,01)	
El jefe de hogar o su cónyuge tienen el quechua como lengua materna	-0,21 (0,31)	
Alumbrado por electricidad	0,05 (0,36)	
Número de parcelas	-0,19 (0,11)	*
Superficie total (hectáreas)	0,21 (0,07)	***
Superficie que se encuentra bajo riego	-0,02 (0,45)	
Recibió asistencia técnica en cualquier tema	-0,69 (0,40)	*
Recibió alguna capacitación	-0,45 (0,27)	*
Hogares cuyo jefe de hogar se dedica a la actividad agrícola	0,59 (0,52)	
Unidades agropecuarias que reportan que la actividad agropecuaria les produce ingresos suficientes (CENAGRO 2012)	0,21 (0,55)	
Unidades agropecuarias en las que se usa tractor (CENAGRO 2012)	0,75 (0,67)	
Constante	-1,00 (0,91)	
Pseudo R ₂	0,15	
F-test	0,26	

Nota: Los valores entre paréntesis son los errores estándar. *** p<0,01, ** p<0,05, * p<0,1

Aunque los resultados no muestran ningún desbalance importante en los atributos de quienes no pudieron ser entrevistados, podría darse el caso de que las diferencias en no observables generen en los impactos algún sesgo atribuible a esta pérdida de 4,3% de la muestra. Sin embargo, como veremos, este no parece ser el caso.

Por un lado, tal como lo muestra la tabla 3.8, en lo que se refiere a diferencias en la línea de base para un grupo representativo de los indicadores de impacto, no parece haber contrastes significativos entre los hogares entrevistados y no entrevistados. En la línea de base no existen diferencias en los principales indicadores vinculados a la mejora de la vivienda, el consumo de alimentos en el hogar o las capacidades financieras de los hogares. Se identifica una diferencia estadísticamente significativa en el *stock* inicial de algunos animales menores —menor cantidad de cuyes entre los hogares no entrevistados—, y el menor consumo de menestras y legumbres.

¿Hasta qué punto estas diferencias identificadas en las características de los hogares no entrevistados respecto a los hogares entrevistados, o las diferencias en los valores iniciales de las variables resultado entre ambos grupos, pueden influir en la evaluación del impacto de Haku Wiñay en estos hogares? Para explorar si podría existir algún sesgo en nuestros resultados debido a la pérdida muestral ocurrida en la revisita, hemos calculado la prueba estadística sugerida por Beckett y otros (1988), conocida como prueba de BGLW. La intuición detrás de la prueba es evaluar hasta qué punto existen diferencias significativas en la línea de base, en los determinantes de cada variable de impacto, entre quienes han sido entrevistados y quienes no han podido ser entrevistados. La prueba conjunta que evalúa si los parámetros de la ecuación de los determinantes de cada variable de impacto son similares entre el total de la muestra y el subgrupo de los revisitados constituye la base de la prueba estadística.

Tal como muestra la tabla 3.9, la prueba estadística para casi todos los indicadores de impacto evaluados sugiere que los determinantes de los indicadores de impacto no son distintos entre el grupo de hogares revisitados y quienes no fueron entrevistados. Esto sugiere no solo que la pérdida de la muestra es pequeña, sino, además, que no parece generar

Tabla 3.8
Diferencia de medias en variables de impacto medidas en la línea de base, entre hogares entrevistados y no entrevistados en la revisita

Variables de impacto	Medias		Diferencia
	Hogares entrevistados	Hogares no entrevistados	
Cuenta con cocina mejorada (%)	60,5	57,9	-2,6
Los animales menores o chanchos ingresan a los espacios donde las personas cocinan, comen o duermen (%)	20,8	15,8	-5,0
Los miembros del hogar realizan un uso productivo de la basura (compost) (%)	18,5	10,5	-7,9
Los miembros del hogar tratan el agua antes de beberla (%)	80,6	89,5	8,9
Los miembros del hogar se lavan las manos al inicio del día (%)	56,3	47,4	-8,9
Consumen hojas verdes (número de veces al año)	135,1	116,8	-18,3
Consumen menestras/legumbres y semillas (número de veces al año)	209,5	167,6	-42,0*
Consumen proteína animal (número de veces al año)	169,2	198,3	29,1
Consumen huevos (número de veces al año)	250,5	234,7	-15,7
Índice de Herfindahl: diversidad de la dieta anual (9 rubros)	0,97	0,96	-0,01
Gasto total de alimentos comprados en el mercado durante el último mes (S/.) a precios del 2013	219,40	250,70	31,38
Hogares que cuentan con tierra bajo riego (%)	14,02	15,79	1,77
Número de cuyes	6,2	2,9	-3,3 *
Número de gallinas	5,3	4,4	-0,9
Capacidad de duplicar ingresos trabajando más (%)	55,6	68,4	12,8
Algún miembro del hogar mantiene abierta una cuenta de ahorros (%)	7,2	5,3	-2,0
Conoce el fondo de seguros de depósitos (%)	0,9	0,0	-0,9
Si cuenta con un monto importante, prefiere ahorrarlo en el banco (%)	13,1	15,8	2,7

Nota: *** p<0,01, ** p<0,05, * p<0,1

sesgos significativos en la evaluación de impacto que se presentará en el capítulo siguiente. Únicamente en el caso del indicador de diversificación de la dieta se identifica una diferencia significativa.

Tal como sugieren Fitzgerald y otros (1998), estimar la significancia conjunta en el *probit* que incorpora los principales determinantes (observables) de la pérdida muestral, y evaluar si existe un impacto diferenciado de los observables en las variables resultado en la línea de base, permite discernir si existe o no un sesgo vinculado a la pérdida muestral. Aquí, el análisis de ambas pruebas sugiere que no es el caso.

Tabla 3.9
Prueba de BGLW: diferencias entre los determinantes de quienes fueron entrevistados y quienes no fueron entrevistados en la revisita

	Valor de la prueba estadística	
Cuenta con cocina mejorada	0,00	
Los animales menores o chanchos ingresan a los espacios donde las personas cocinan, comen o duermen	0,00	
Los miembros del hogar realizan un uso productivo de la basura (compost) (%)	0,00	
Los miembros del hogar tratan el agua antes de beberla (%)	0,00	
Los miembros del hogar se lavan las manos al inicio del día (%)	3,31	
Consumo de hojas verdes (número de veces al año)	0,97	
Consumo de menestras/legumbres y semillas (número de veces al año)	1,07	
Consumo de proteína animal (número de veces al año)	1,08	
Consumo de huevos (número de veces al año)	0,95	
Índice de Herfindahl: diversidad de la dieta anual (9 rubros)	2,61	**
Gasto total de alimentos comprados en el mercado en el último mes (S/.) a precios del 2013	0,84	
Hogares que cuentan con tierra bajo riego	0,00	
Número de cuyes	0,50	
Número de gallinas	0,61	
Activos e ingresos suficientes (autoevaluación: 1-10)	1,37	
Capacidad de duplicar los ingresos trabajando más	0,00	
Algún miembro del hogar mantiene abierta una cuenta de ahorros	0,89	
Si cuenta con un monto importante, prefiere ahorrarlo en el banco	2,32	

Nota: *** p<0,01, ** p<0,05, * p<0,1.

Referencias bibliográficas

- Beckett, Sean; William Gould, Lee Lillard y Finis Welch (1988). The panel study of income dynamics after fourteen years: an evaluation. *Journal of Labour Economics*, 6(4), 472-492.
- Duflo, Esther; Rachel Glennerster y Michael Kremer (2008). Using randomization in development economics research: a toolkit. En T. Paul Schultz and John Strauss (Eds.). *Handbook of Development Economics* (pp. 3895-3962). Vol. 4. Amsterdam; London: North Holland.
- Fitzgerald, John; Peter Gottshalk y Robert Moffitt (1998). An analysis of sample attrition in panel data: the Michigan panel study on income dynamics. *Journal of Human Resources*, 33(2), 251-299.
- Hainmueller, Jens (2012). Entropy Balancing for Causal Effects: A Multivariate Reweighting Method to Produce Balanced Samples in observational studies. *Political Analysis*, 20(1), 25-46. Recuperado de <http://pan.oxfordjournals.org/content/20/1/25.abstract>
- Hainmueller, Jens y Yiqing Xu (2013). ebalance: a Stata package for entropy balancing. *Journal of Statistical Software*, 54(7), 1-18.
- Imbens, Guido y Jeffrey Wooldridge (2009). Recent developments in the econometrics of program evaluation. *Journal of Economic Literature*, 47(1), 5-86.
- Lechner, Michael (2010). The estimation of causal effects by difference-in-difference methods. *Foundations and Trends in Econometrics*, 4(3), 165-224.
- Miguel, Edward y Michael Kremer (2004). Worms: identifying impacts on education and health in the presence of treatment externalities. *Econometrica*, 72(1), 159-217.
- Rubin, Donald B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, 66(5), 688-701. doi:10.1037/h0037350.
- Lechner (2010).

Anexo

Tabla 1
Diferencias entre los hogares tratados que ganaron concursos de negocios rurales inclusivos y los hogares del grupo control

	Sin pesos			Con pesos		
	Hogares en centros poblados controles	Hogares en centros poblados tratados	Sig. ¹	Hogares en centros poblados controles	Hogares en centros poblados tratados	Sig. ¹
Características de los hogares de la muestra						
Hogares en Cajamarca (%)	33,9	49,1	**	23,2	49,1	***
Hogares en Huánuco (%)	34,4	32,1		32,1	32,1	
Años de educación del jefe de hogar	4,6	4,7		4,7	4,7	
Número de miembros del hogar	4,8	5,0		4,6	5,0	**
Hogares en cuyo centro poblado opera alguna ONG (%)	27,2	57,7	***	41,8	57,7	
Hogares cuyo jefe de hogar se dedica a la actividad agrícola (%)	56,9	49,5		49,5	49,5	
Hogares cuya jefa de hogar es mujer (%)	12,2	11,3		11,3	11,3	
Edad del jefe de hogar	41,5	40,7		39,4	40,7	
Indicadores de contexto del CENAGRO 2012 (agregados a nivel de centro poblado)						
Hogares que cuentan con teléfono (%)	18,2	18,9		14,8	18,9	
Hogares alguno de cuyos miembros migra para obtener otros ingresos (%)	32,9	22,4	***	33,5	22,4	
UA en las que se contrata a trabajadores remunerados (%)	59,4	51,8	*	51,8	51,8	
UA que cuentan con algún tipo de riego (%)	18,6	17,0		17,0	17,0	
Promedio del valor del ganado (a precios de 1994)	1029	954		953	953	
Hogares que reportan que la actividad agropecuaria les produce ingresos suficientes (%)	20,3	18,6		18,6	18,6	
UA en las que se utiliza tractor (%)	4,3	7,6		4,6	7,6	
Número de hogares	221	53				

¹ Significancia estadística de la diferencia. ***p<0,01, **p<0,05, *p<0,1.

Tabla 2
Diferencias entre submuestras de hogares tratados
(con menor y mayor intensidad) y sus respectivos grupos control

Características previas a la intervención	(a) Tercio de menor intensidad				(b) Tercio de mayor intensidad				
	Sin pesos		Con pesos		Sin pesos		Con pesos		
	Hogares en centros poblados controles	Hogares Sig. ^{1/} en centros poblados controles	Hogares en centros poblados controles	Hogares Sig. ^{1/} en centros poblados controles	Hogares en centros poblados controles	Hogares Sig. ^{1/} en centros poblados controles	Hogares en centros poblados controles	Hogares Sig. ^{1/} en centros poblados controles	
Hogares en Cajamarca	33,9	31,8	55,8	31,7	*	33,9	44,1	29,3	44,1
Hogares en Huánuco	34,4	30,2	28,8	30,2		34,4	32,2	32,2	32,2
Años de educación del jefe de hogar	4,6	3,9	3,9	3,9		4,6	4,7	4,7	4,7
Número de miembros del hogar	4,8	5,1	4,6	5,1	**	4,8	5,0	4,6	5,0
Hogares en cuyo centro poblado opera alguna ONG	27,2	64,9	30,7	64,9	*	27,2	60,3	45,9	60,3
Hogares cuyo jefe de hogar se dedica a la actividad agrícola	56,9	68,3	68,3	68,3	***	56,9	53,5	53,5	53,5
Hogares cuya jefa de hogar es mujer	12,2	23,8	23,8	23,8	**	12,2	13,6	13,6	13,6
Edad del jefe de hogar	41,5	43,2	42,3	43,2		41,5	40,7	39,8	40,7

Características previas a la intervención	(a) Tercio de menor intensidad				(b) Tercio de mayor intensidad			
	Sin pesos		Con pesos		Sin pesos		Con pesos	
	Hogares en centros poblados controlados	Hogares Sig. ^{1/} en centros poblados controlados	Hogares en centros poblados controlados	Hogares Sig. ^{1/} en centros poblados controlados	Hogares en centros poblados controlados	Hogares Sig. ^{1/} en centros poblados controlados	Hogares en centros poblados controlados	Hogares Sig. ^{1/} en centros poblados controlados
Indicadores de contexto del CENAGRO 2012 (agregados a nivel de centro poblado)								
Hogares que cuentan con teléfono	18,2	20,2	17,2	20,2	18,2	19,8	13,9	19,8
Hogares alguno de cuyos miembros migra para obtener otros ingresos	32,9	22,7	***	33,3	22,7	32,9	20,2	***
UA en las que se contrata a trabajadores remunerados	59,4	45,5	***	45,5	45,5	59,4	48,3	**
UA que cuentan con algún tipo de riego	18,6	15,8		15,8	15,8	18,6	17,5	
Promedio del valor del ganado (a precios de 1994)	1029	1003		1003	1003	1029	932	*
Hogares que reportan que la actividad agropecuaria les produce ingresos suficientes	20,3	27,1	**	23,5	27,1	20,3	18,7	
UA en las que se utiliza tractor	4,3	16,2	***	6,5	16,2	4,3	10,9	***
Número de hogares	221	63			59	221		

1 Significancia estadística de la diferencia. ***p<0,01, **p<0,05, *p<0,1.