ISSN 1510-3943

MINISTERIO DE GANADERÍA
AGRICULTURA Y PESCA
REPÚBLICA ORIENTAL DEL URUGUAY

Análisis sectorial y cadenas productivas
Temas de política
Estudios

ANUARIO 2016 opypa

MINISTERIO DE GANADERÍA,
AGRICULTURA Y PESCA
REPÚBLICA ORIENTAL DEL URUGUAY

OPYPA

Oficina de Programación y Política Agropecuaria

Análisis sectorial y cadenas productivas

Temas de política

Estudios

MINISTERIO DE GANADERÍA, AGRICULTURA Y PESCA

MINISTRO
Ing. Agr. Tabaré Aguerre

SUBSECRETARIO
Ing. Agr. Enzo Benech

DIRECTOR GENERAL
Dr. Alberto Castelar

OFICINA DE PROGRAMACIÓN Y POLÍTICA AGROPECUARIA
Ing. Agr. Adrián Tambler
Subdirector

Coordinación general de la publicación: Mónica Balparda
Coordinación de la Sección Cadenas Productivas: Gonzalo Souto
Coordinación de la Sección Temas de Política: Adrián Tambler
Coordinación de la Sección de Estudios: José Bervejillo

Los artículos en esta publicación están disponibles en: www.mgap.gub.uy/opypa

Avances en la evaluación de políticas agropecuarias en el MGAP[1]

Verónica Durán Fernández[2]
Emilio Aguirre[3]
Patricia Artía[4]
Juan Baraldo[5]
Darío Fuletti[6]
Elisa Hernández[7]

El Ministerio de Ganadería, Agricultura y Pesca (MGAP) está dando un fuerte impulso a la institucionalización de la evaluación de las políticas agropecuarias. Se han fortalecido las capacidades en el Área de Evaluación de Políticas Agropecuarias, un equipo técnico especializado que funciona desde principios de 2014 en la órbita de OPYPA, y ya se pueden observar significativos logros en este proceso. Este documento se organiza en tres partes. En la primera se resumen las principales razones y fundamentos de la institucionalización de la evaluación de políticas dentro del MGAP, el diseño institucional y el enfoque de las evaluaciones que se están desarrollando. En segundo lugar se comentan los avances en la agenda de evaluación que ha sido priorizada por las autoridades. Por último se destacan los principales logros de estos primeros dos años y medio de trabajo, así como las lecciones aprendidas y desafíos por delante. En el anexo metodológico se sintetizan las principales técnicas de evaluación de políticas que forman parte de la "caja de herramientas" del Área de Evaluación.

1. Pasos hacia la institucionalización de la evaluación de políticas agropecuarias en el MGAP

1.1. ¿Por qué evaluar en el MGAP?

El Ministerio de Ganadería, Agricultura y Pesca (MGAP) tiene una larga experiencia en el monitoreo y seguimiento de diversos programas y proyectos, de donde surge valiosa

[1]*Se agradecen los comentarios del equipo técnico de AGEV-OPP.*
[2]*Ec. Coordinadora del Área de Evaluación de Impacto en OPYPA, vduran@mgap.gub.uy.*
[3]*Ec. Especialista en Evaluación de Impacto en OPYPA, emaguirre@mgap.gub.uy.*
[4]*Dra. Lic. Antrop., Especialista en Evaluación de Políticas en OPYPA, partia@mgap.gub.uy.*
[5]*Ec. Especialista en Evaluación de Impacto en OPYPA, jbaraldo@mgap.gub.uy.*
[6]*Ec. Especialista en Evaluación de Impacto en OPYPA, dfuletti@mgap.gub.uy.*
[7]*Ec. Especialista en Evaluación de Impacto en OPYPA, mhernandez@mgap.gub.uy.*

información sobre la marcha de las intervenciones y también sobre los productos y servicios que llegan a los productores agropecuarios beneficiarios.

Las unidades que ejecutan las herramientas de política tienen sus propios equipos técnicos para el monitoreo, que se generan información y análisis sobre la implementación de los proyectos, básicamente relativa al cumplimiento de las metas de ejecución presupuestal y al logro de los productos y servicios de la intervención. Así, el ministerio cuenta con abundante información sobre: el monto de recursos invertidos en las diferentes intervenciones, la cantidad de productores que fueron apoyados para realizar inversiones prediales, los montos y unidades físicas de diferentes inversiones prediales apoyadas (soluciones de agua, alambrado para empotreramiento, etc.), los diferentes cursos y capacitaciones prestados, la cantidad de productores que participaron en los cursos, entre otras.

Cuadro 1. Importancia de la evaluación de políticas

¿Por qué evaluar?	
Se necesita evidencia sobre si las intervenciones son efectivas: • Justificación de los programas. • Escalamiento de intervenciones que logran impactos • Comparar diferentes opciones de política Mejorar las herramientas de política • Diseño (elegibilidad, beneficios a otorgar) • Implementación Rendir cuentas y ser transparentes hacia la ciudadanía	Evaluaciones para la toma de decisiones basada en evidencia.

Sin embargo, no son muchas las políticas ni las herramientas de política que han sido evaluadas desde el punto de vista del logro de sus objetivos propiamente dichos; esto implica que es poco lo que se conoce sobre los efectos que tienen las medidas de política agropecuaria sobre los productores y las organizaciones rurales. Para las autoridades ministeriales es de gran interés disponer de evidencia robusta sobre en qué medida las intervenciones logran los propósitos o impactos buscados, ya que ello está muy vinculado con la utilidad o pertinencia de las diferentes herramientas de política para alcanzar los objetivos estratégicos del MGAP.

Por ejemplo, además de conocer la cantidad de soluciones de agua que se construyeron gracias al apoyo de determinado programa, es muy útil saber si debido a la medida de política los productores beneficiarios aumentaron su productividad, o sus ingresos. O además de identificar el monto invertido en asistencia técnica y capacitación en cierta

intervención, es clave saber si esos conocimientos sirvieron para que los productores beneficiarios modificaran sus prácticas de manejo y adoptaran tecnología.
Por eso OPYPA propuso a principios de 2014 construir y fortalecer capacidades internas a través de un equipo técnico especializado, con la misión de asesorar a las autoridades ministeriales sobre los resultados, consecuencias e impactos de las herramientas de política agropecuaria.

El fin último de la evaluación, como una instancia más del ciclo de políticas públicas, es la mejora de las intervenciones, tanto en lo que hace a su diseño como a la implementación. Asimismo, las evaluaciones permiten rendir cuentas y ser transparentes hacia los productores y la ciudadanía en términos generales.

Cuadro 2. Misión y visión del Área de Evaluación de Políticas Agropecuarias

Visión	Misión
La evaluación como instrumento para el diseño/rediseño de políticas, • permite mejorar los instrumentos e intervenciones y aumentar la efectividad del gasto público • legitima y hace más transparente la política pública.	Asesorar a las autoridades del MGAP sobre los resultados, consecuencias e impactos de las herramientas de política que implementa y otros organismos del Poder Ejecutivo, Sugerir mejoras, nuevos diseños y formulaciones de los instrumentos de política, de modo de alcanzar en forma más eficiente los objetivos estratégicos buscados.

1.2. Diseño institucional: el Área de Evaluación de Políticas Agropecuarias

La estrategia de fortalecer las capacidades internas de evaluación en un área especializada dentro de OPYPA busca aprovechar un conjunto de fortalezas y oportunidades, aunque no está exenta de algunos riesgos y desventajas que son necesario mitigar

Importa destacar que, si bien las evaluaciones son internas al MGAP, se desarrollan fuera de las unidades que ejecutan e implementan las herramientas de política, de modo que puede afirmarse que no constituyen autoevaluaciones.
Son varias las **ventajas** de la construcción de capacidades internas de evaluación. Una de las principales es la posibilidad de desarrollar una “cultura de la evaluación” en el MGAP.
Asimismo, en las evaluaciones internas se aprovechan los conocimientos de técnicos y expertos del MGAP, ya que es fundamental entender profundamente la teoría del cambio del programa y el contexto de las intervenciones para lograr buenas evaluaciones.

Por último, es posible que las recomendaciones de mejora de las políticas y reportes de evaluación sean más aceptados si provienen "desde dentro" del MGAP.

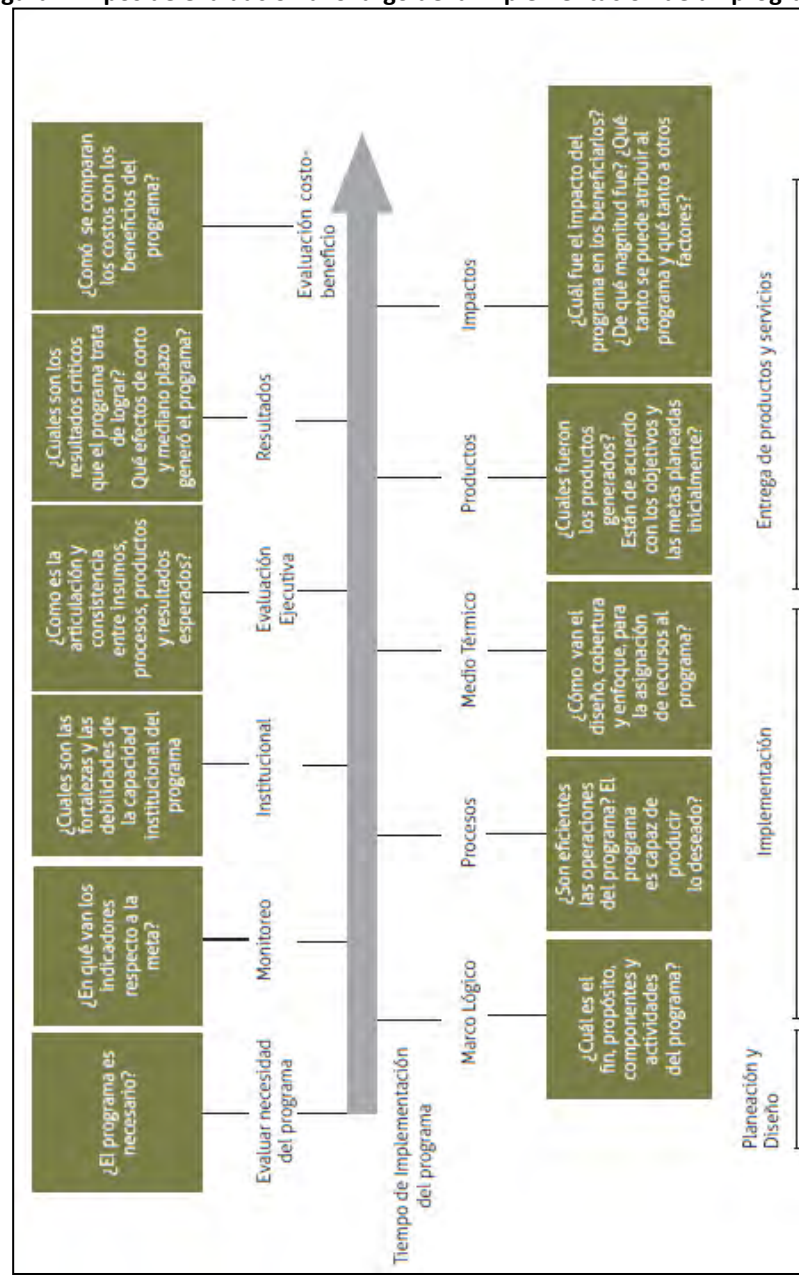
Sin embargo, las evaluaciones hechas dentro del MGAP plantean el **desafío** de mitigar los riesgos de autocomplacencia. Para ello, el Área de Evaluación incorporó un socio estratégico que fortalece la imagen y da garantías técnicas acerca de la independencia de la evaluación (una mirada externa): la Dirección de Gestión y Evaluación del Estado (AGEV)[8], quienes son un referente en evaluación de políticas públicas en Uruguay. La cooperación entre OPYPA y AGEV está formalmente respaldada en un convenio marco firmado por las autoridades del MGAP y de OPP, con el objetivo de aunar capacidades para llevar adelante evaluaciones de las políticas agropecuarias. El convenio establece que, en la medida de sus posibilidades, ambas partes aportarán los recursos humanos y materiales necesarios para llevar adelante evaluaciones de programas implementados por el MGAP y que los informes serán de coautoría de ambas instituciones. Además, el convenio establece cláusulas de confidencialidad, regula el uso y la propiedad conjunta de las bases de datos que se generen y prevé la participación de terceros organismos públicos o privados en las evaluaciones.

La colaboración técnica entablada entre ambos organismos ha permitido también la incorporación de buenas prácticas y lecciones aprendidas en materia evaluación de intervenciones públicas a partir de la experiencia acumulada por AGEV-OPP en otros ámbitos de política pública.

El apoyo de AGEV ha sido clave en el proceso de creación y fortalecimiento de capacidades de evaluación en el MGAP, aunque su participación no debería limitarse a estas primeras etapas de trabajo conjunto, sino que parece deseable que el vínculo institucional continúe profundizándose como un aporte a la institucionalización de la evaluación y mejora continua de los servicios prestados desde el sector público.

[8] *La Dirección de Gestión y Evaluación del Estado (AGEV) está ubicada dentro de la Dirección de Presupuesto Control y Evaluación de la Gestión, perteneciente a la Oficina de Programación y Presupuesto (OPP).*

Figura 1. Tipos de evaluación a lo largo de la implementación de un programa



Fuente: "Guía metodológica para la Evaluación de políticas públicas de Sinergia" (2002).

1.3. Enfoque de trabajo del Área de Evaluación

Las herramientas y medidas de política pública pueden ser evaluadas desde distintos enfoques y utilizando diversas herramientas metodológicas que aportan resultados y visiones complementarias. De acuerdo a la "Guía metodológica para la Evaluación de políticas públicas de Sinergia" (2002) las evaluaciones pueden dividirse de acuerdo a la fase en la que se encuentra el programa, dado que se realizan antes, durante o tras finalizar su implementación.

Los trabajos del Área de Evaluación de Políticas Agropecuarias se concentran en la **evaluación ex post** de las políticas y sus herramientas, con el objetivo de obtener **resultados,** aislar su **impacto** y estimar su relación **costo efectividad**. Quedan así excluidos del ámbito de trabajo del área otros tipos de evaluaciones y seguimientos, también sumamente importantes, en particular los que refieren al control y monitoreo de la ejecución de las medidas de intervención.

Este tipo de evaluaciones tienen como fin principal atribuir impactos a un programa y examinar si está generando los efectos deseados. Para ello, estas evaluaciones responden, entre otras, a las siguientes preguntas: ¿Qué le habría pasado a los beneficiarios de la intervención si el programa no hubiese existido?; ¿Qué tanto de los efectos sobre los beneficiarios se puede atribuir al programa y qué tanto a otros factores?; ¿Qué tan efectivo fue el programa en lograr sus objetivos? ¿Hubo cambios, y de qué magnitud, en los problemas que intenta solucionar el programa? ¿Se justifican los costos? ¿Se podría modificar el diseño para mejorar sus efectos?

2. Agenda de evaluación priorizada

El Área de Evaluación de Políticas Agropecuarias acordó con el Ministro un conjunto de evaluaciones priorizadas, que están siendo desarrolladas en estos momentos. Todas las herramientas de política en evaluación están alineadas con algunos de los objetivos estratégicos del MGAP. Asimismo, esto permitió elaborar una Agenda Plurianual de Evaluación conjunta con la Oficina de Planeamiento y Presupuesto a través de la Dirección de Gestión y Evaluación.

Las evaluaciones priorizadas vinculadas con el objetivo estratégico **adaptación de los sistemas de producción a la variabilidad climática** son:

- Proyecto Ganaderos Familiares y Cambio Climático (GFCC)
- Programa Agua para la Producción Animal (APA)

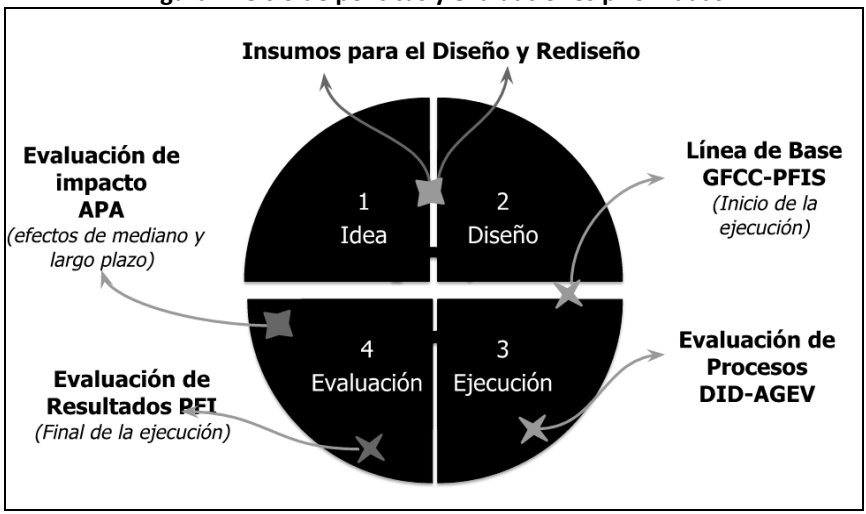
Para el objetivo **desarrollo rural - inserción competitiva de la agricultura familiar en cadenas de valor** las intervenciones priorizadas son:

- Planes de Fortalecimiento Institucional (PFI)
- Producción Familiar Integral y Sustentable (PFIS)

Cada una de estas intervenciones se encuentra en diferentes etapas del ciclo de políticas públicas, de modo que los informes y hallazgos del proceso de evaluación aportan información y recomendaciones para el diseño, rediseño e implementación de las diversas medidas de política, en función de la etapa que se esté transitando (Figura 2). Por ejemplo: los programas GFCC y PFIS, si bien ya iniciaron su ejecución, están en etapa de levantamiento de la línea de base. El PFI, que en los próximos meses estaría finalizando su ejecución, se ubica en etapa de evaluación de resultados; y para la intervención APA, que ya finalizó la implementación, se realizó una evaluación de impacto y se aportaron insumos para el diseño de un nuevo programa con similares objetivos.

A su vez, durante la implementación de algunas herramientas de política del MGAP, AGEV-OPP está realizando evaluaciones de Diseño, Implementación y Desempeño (DID), herramienta destinada a facilitar la mejora de la gestión pública. Estos estudios se orientan a analizar el desempeño de una intervención, para identificar aspectos a ser mejorados y realizar recomendaciones para su puesta en práctica.

Figura 2: Ciclo de políticas y evaluaciones priorizadas



Fuente: Elaboración propia.

A continuación se presentan las herramientas de política en evaluación: luego de una breve descripción de cada programa, se presenta la estrategia de evaluación y en los casos que corresponda, algunos resultados preliminares.

2.1. Estrategia de Evaluación del Programa "Ganaderos Familiares y Cambio Climático (GFCC)"

2.1.1. Descripción del programa

El proyecto Ganaderos Familiares y Cambio Climático (GFCC), implementado por la Dirección General de Desarrollo Rural (DGDR) del MGAP, busca atender la vulnerabilidad de los productores ganaderos familiares ante la variabilidad climática en las dos Unidades de Paisaje[9] (UP) más afectadas por esta problemática: "Sierras del Este" y "Basalto superficial". El programa comenzó a implementarse en 2013 y aún no ha finalizado, por lo que los trabajos de evaluación se concentraron en explicitar la teoría del programa y sus indicadores y recoger la línea de base mediante una encuesta.

Los postulantes deben presentar proyectos individuales o colectivos, elaborados por técnicos privados habilitados por la DGDR. Los planes tendrán un plazo de ejecución de hasta 2 años y permitirán el financiamiento, de manera no reembolsable, de hasta 80% del monto de las actividades presentadas si están registrados como productores familiares, con un máximo de 8.000 dólares por productor. Una característica novedosa del proyecto es que el 15% de esta partida deberá destinarse a "fondos rotatorios" para fortalecer organizaciones del medio rural. El presupuesto total del programa es de 10 millones de dólares.

El propósito del proyecto es incrementar la resiliencia climática de los productores familiares ganaderos de las UP; es decir, aumentar la capacidad de amortiguar el impacto de eventos climáticos adversos sobre el nivel de productividad de los sistemas ganaderos. El proyecto consiste en utilizar el apoyo del Fondo de Adaptación al Cambio Climático y el conocimiento acumulado en la academia y el MGAP para la realización de inversiones prediales y la divulgación de tecnologías de manejo entre los productores de las UP mencionadas.

Se espera como resultado de corto plazo que estas actividades estimulen la adopción de mejores prácticas de manejo, una mayor participación de productores en redes y la incorporación de información agroclimática en las decisiones de producción. En un plazo más largo, estos cambios permitirían aumentar la sostenibilidad del campo natural y reducir la caída de la productividad ante eventos climáticos extremos.

---

[9] De acuerdo al documento del proyecto, "la definición de unidad de paisaje sigue la definición adoptada por la Convención Europea del Paisaje: "Paisaje" se define como una zona o área según la perciben los pobladores locales o visitantes, cuyas características visuales y carácter resultan de la acción de factores naturales y/o culturales (es decir, humanos). (...). También se resalta que un paisaje constituye una unidad en sí mismo con componentes naturales y culturales, con sus servicios de ecosistemas; todos los factores tomados en su conjunto, no en forma separada".

2.1.2. Estrategia de evaluación

El MGAP cuenta con sistemas de información que contienen datos muy relevantes para realizar este tipo de evaluaciones. Por un lado, el Sistema Nacional de Información Ganadera (SNIG) registra todos los movimientos del ganado vacuno desde el establecimiento de origen del animal hasta el frigorífico y contiene además información del uso de la tierra del predio. Por otro lado, el Sistema Nacional de Información Agropecuaria (SNIA) provee información de clima, suelo, agua y vegetación, que serán utilizadas como variables de control.

Debido a que estos sistemas no contienen información sobre prácticas de manejo del ganado y las pasturas ni sobre la participación de los productores en redes (algunos de los objetivos que el programa aspira a modificar en el corto plazo), en 2015 se realizó una encuesta de línea de base a productores ganaderos familiares, tanto beneficiarios como no beneficiarios (candidatos a grupo de control). Con base en la información disponible o con posibilidades de construirse, se presentan los indicadores para evaluar los resultados e impactos del programa (Cuadro 3).

La encuesta de línea de base confirmó que existían diferencias importantes entre los beneficiarios y los no beneficiarios en el momento previo al programa. En estos casos, una comparación directa de resultados entre ambos grupos luego de la intervención no sería una buena estimación del impacto porque tendría un sesgo por autoselección.

En este contexto, se utilizará el método de Diferencias en Diferencias, evaluando su combinación con *Propensity Score Matching* u otro tipo de pareamiento (ver anexo metodológico). Esta combinación se considera adecuada para evaluar la mayor parte de los indicadores de resultado e impacto. Los indicadores de resultados se computarán con base en las encuestas, por lo cual sólo se dispondrá de información en dos momentos del tiempo (la encuesta deberá repetirse luego de finalizado el programa, en 2018). Por el contrario, en el caso de la productividad de carne, la tasa de eficiencia y la sustentabilidad, será posible comparar su evolución a lo largo del tiempo, pues el SNIG y SNIA son registros continuos.

Cuadro 3. Indicadores de resultados e impactos del GFCC

	Indicadores	Fuentes
Resultados	**Adopción de practicas de manejo**	**Encuestas**
	Participación en asociaciones de productores	**Encuestas**
	Participación en grupos de productores	**Encuestas**
	Utilización de información agroclimática	**Encuestas**
	Percepción capacidad de respuesta individual a sequías	**Encuestas**
	Acceso a asistencia técnica	**Encuestas**
Impactos	Productividad (kg carne/há/año) ● Nivel en años sin eventos extremos (posible aumento de la producción por un mejor aprovechamiento de los recursos) ● Varianza de la producción (por lograr una producción más estable) ● Menor incidencia de las variables climáticas	**SNIG**
	Eficiencia Reproductiva: Tasa de destete (%)	**SNIG**
	Sustentabilidad: NDVI	**SNIA**

Fuente: Elaboración propia.

2.1.3. Estado actual de la evaluación

Debido a que el proyecto GFCC aún está en ejecución, la evaluación se encuentra actualmente en fase de diseño. Las actividades más importantes en esta fase consisten en la consolidación y depuración en una base única de los datos necesarios de los registros de SNIG y SNIA. Además, se deberá corroborar que se cumplen los supuestos que implica la estrategia empírica elegida[10] y definir el mejor tipo de emparejamiento para disminuir las diferencias entre el grupo de tratados y controles. La primera evaluación de efectos de mediano plazo podrá realizarse luego de la finalización del programa, en 2018.

[10] *Por ejemplo, las tendencias paralelas en los indicadores de resultados entre el grupo de beneficiarios y el grupo de control.*

2.2. Estrategia de Evaluación del Programa "Producción familiar, integral y sustentable (PFIS)"

2.2.1. Descripción del programa

En 2014 la Dirección General de Desarrollo Rural (DGDR) del MGAP realizó una convocatoria a la presentación de propuestas para mejorar la sustentabilidad económica, social y ambiental de productores agropecuarios familiares, pequeños (no familiares) y medianos. El programa se encuentra actualmente en plena etapa de implementación, por lo que la estrategia de evaluación incluye la reconstrucción de la línea de base, mediante la realización de una encuesta.

La intervención busca mejorar la productividad de los establecimientos beneficiarios y generar mayor capacidad de adaptación y resiliencia a fenómenos climáticos adversos mediante la promoción para la adopción de nuevas tecnologías productivas y organizacionales. Se plantearon dos objetivos intermedios, que llevarían a los resultados finales: la adopción de innovación tecnológica por parte de estos productores y el manejo responsable de los recursos naturales. Por lo tanto, este llamado combina dos componentes, uno referido a promover el desarrollo tecnológico-productivo y organizacional; y un segundo componente que busca promover la conservación de los recursos naturales y la adaptación al cambio climático.

El apoyo económico que recibe cada productor tiene un máximo no reembolsable de hasta 16.000 dólares (hasta 8.000 dólares por cada componente). El plazo para la ejecución de propuestas es de 18 meses, iniciando en los primeros meses de 2015.

A la convocatoria se presentaron 3.609 propuestas de todos los rubros productivos, de las cuales fueron aprobadas 2.215 (61%). La mayoría de las propuestas presentadas (1.869) corresponden al rubro ganadería, de las cuales fueron aprobadas 1.046.

2.2.2. Estrategia de evaluación

Debido a las restricciones en la información disponible y los limitados recursos para evaluación, se decidió realizar una evaluación de impacto únicamente para el rubro ganadería.

Siguiendo el protocolo de evaluación de OPYPA, desde el inicio se convocó un grupo ad-hoc con técnicos de OPYPA, DGDR, UGP y AGEV-OPP. Como primer paso en el proceso de la evaluación, este grupo técnico trabajó para explicitar la teoría de cambio del programa y definir los indicadores de impacto y de caracterización.

La evaluación debe partir de una correcta identificación de: ¿qué cambios se espera lograr en los productores con el PFIS?, ¿cómo se producirían esos cambios (cuál es la cadena causal)? y ¿cómo medir estos cambios esperados?

Para medir los indicadores de la evaluación -fundamentalmente los que refieren a la adopción de tecnologías y medidas de manejo de las pasturas y el ganado- se consideró imprescindible realizar una encuesta que permita reconstruir la información en la línea de base (año 2015). Para ello se cuenta con financiamiento aportado por AGEV-OPP.

Coincidentemente, OPYPA está liderando la realización de una Encuesta Ganadera Nacional financiada por diversas instituciones (OPYPA, BCU, INIA, ANII y SUL), fundamentalmente para relevar costos y tecnología, de modo que se consideró oportuno sumar los recursos y necesidades de la evaluación de impacto del PFIS a este esfuerzo interinstitucional. DIEA es la oficina especializada responsable técnica del relevamiento.

En base a la información disponible o con posibilidades de construirse, se presentan los indicadores para evaluar los resultados e impactos del programa (Cuadro 4).

La evaluación de impacto requiere construir un grupo de tratamiento y uno de control o comparación. Para aplicar los métodos de Diferencias en Diferencias con *Propensity Score Matching* (ver detalles en el anexo metodológico) el grupo de control se seleccionará del universo de todos los establecimientos que no recibieron el subsidio. Lo novedoso de esta evaluación de impacto es que, además de esas técnicas, se cumplen los requisitos que hacen posible aplicar el Diseño de Regresión Discontinua: 1) el subsidio fue asignado en forma discontinua en base a un puntaje, con un punto de corte exógeno[11]; 2) el número de productores que quedaron a ambos lados del punto de corte es similar; 3) las características de los productores a ambos lados del corte son similares. De modo que, bajo esta metodología, el grupo de control estará formado por los postulantes rechazados, en un entorno del umbral para la asignación del subsidio.

[11] Para que un proyecto se aprobara debería obtener una calificación mínima de 60 puntos en la evaluación técnica (sobre un máximo de 100); no obstante los fondos fueron insuficientes en el segundo llamado para atender a todas las propuestas que superaron ese mínimo, por lo que se decidió correr el punto de corte y financiar aquellas que superaron los 66 puntos. La clave para aplicar el método de Regresión Discontinua es que este nuevo punto de corte no fue manipulado por los candidatos ni por los técnicos que hicieron las calificaciones, porque era impredecible (fue exógeno).

Cuadro 4. Indicadores de resultados e impactos del PFIS

Indicadores		Fuentes
Impactos	**Productividad ganadera:** kg carne equivalente/ha pastoreo/año	SNIG
	Eficiencia reproductiva: 1) Tasa de destete; 2) Tasa de procreo; 3) Eficiencia reproductiva global[12]	Encuesta Ganadera
	Capacidad adaptativa individual: Percepción de la capacidad de respuesta individual a sequías por parte de los productores.	
Resultados	**Incorporación de prácticas de manejo:** % de productores que incorpora buenas prácticas de manejo del rodeo y campo natural	
	Capacidad adaptativa individual: Medidas de adaptación a la variabilidad climática.	
	Fortalecimiento de la asociatividad: Participación de los productores en grupos y organizaciones.	

Fuente: Elaboración Propia.

La Encuesta Ganadera Nacional que se está llevando adelante tiene como unidad de análisis la explotación agropecuaria, según la definición del Censo Agropecuario, por lo que el marco muestral se construyó a partir del Censo Agropecuario 2011.[13] La muestra de 1.506 casos fue realizada por el INE, con una sub muestra específica para la evaluación de impacto del PFIS (Cuadro 5). Las variables de estratificación fueron: región y tamaño del rodeo (ovino y bovino, en unidades ganaderas).

Paralelamente, se trabajó en la definición conceptual del formulario, con el fin de medir stock y prácticas de manejo de ganado vacuno y ovino, uso y aprovechamiento de la tierra, participación en capacitaciones y asistencia técnica, y conocimiento y prácticas de adaptación a la variabilidad del cambio climático.

A la fecha del presente informe se está comenzando el trabajo de campo de la Encuesta.

[12] *Número de terneros en stock por cada 100 hembras mayores de 1 año en el stock del año anterior. Este indicador es conocido como "parámetro de eficiencia reproductiva" o PER.*

[13] *Los criterios de la Encuesta Ganadera para definir el marco muestral fueron: 1) actividad principal ganadera vacuna u ovina; 2) sin actividad de lechería comercial; 3) con al menos 7 unidades ganaderas (UG).*

Cuadro 5. Muestra Encuesta Ganadera

No beneficiaries (candidatos control)		Beneficiarios para Dif en Dif y PSM		Beneficiarios y Controles para RD		Total	
Marco	Muestra	Marco	Muestra	Marco	Muestra	Marco	Muestra
24.302	957	987	272	277	277	25.566	1.506

Fuente INE (2016).

2.3. Estrategia de Evaluación de Resultados del Programa "Propuestas de Fortalecimiento Institucional para el Desarrollo Rural Sustentable (PFI)"

2.3.1. Descripción del programa

La intervención tiene como propósito u objetivo general "apoyar a las organizaciones del medio rural mediante el financiamiento parcial de equipos multidisciplinarios, de asistencia técnica e inversiones, de forma de mejorar sus capacidades para analizar, proponer y ejecutar, desde la lectura de sus territorios de influencia, acciones que contribuyan al desarrollo, apoyando la integración a la organización de productores, trabajadores y otros pobladores rurales".[14]

Como criterio general de elegibilidad los beneficiarios potenciales fueron todas aquellas organizaciones de la sociedad civil vinculadas al desarrollo rural, estrictamente privadas, con personería jurídica, de base territorial, con diversos grados de desarrollo institucional.

Estas organizaciones deberán ser representativas de: productores agropecuarios organizados, con énfasis en las organizaciones de productores familiares, pescadores artesanales, asalariados rurales, pobladores rurales en general, incluyendo organizaciones de diferente naturaleza: comunitarias, centros sociales y culturales, cooperativas, ONGs y otras del medio rural y de habitantes de centros poblados menores. También son beneficiarias potenciales aquellas organizaciones de segundo grado que demanden fortalecimiento para aplicar específicamente en apoyo de organizaciones de base local, menos desarrolladas, mediante la canalización de respaldos técnicos y acompañamiento institucional que permitan aumentar su grado de incidencia en el territorio en términos de promoción, organización social y desarrollo de servicios.

[14] *Según la Convocatoria pública a presentar propuestas de "Fortalecimiento institucional para el desarrollo rural sustentable", DGDR - MGAP.*

La población objetivo potencial se estima entre 300 y 500 organizaciones, de las cuales 207 presentaron propuestas y 174 fueron beneficiarias, de modo que la tasa de cobertura del programa es elevada.

2.3.2 Estrategia de evaluación de resultados

Desde el inicio de la evaluación se convocó un grupo ad-hoc con técnicos de OPYPA, DGDR, UGP y AGEV-OPP para discutir conjuntamente los objetivos e indicadores de la evaluación, facilitar el acceso a la información existente y la articulación de la recogida de datos y recoger su profundo conocimiento sectorial y acerca del diseño y la implementación de la intervención.

Definición de Objetivos e Indicadores para la evaluación

Para identificar y explicitar los objetivos del programa que serían evaluados y los indicadores de resultados para la evaluación, el grupo ad hoc mencionado trabajó sobre la matriz de objetivos, con base en una metodología desarrollada por AGEV.

A los efectos de la evaluación, se consideró que el propósito que persigue la intervención tiene las siguientes tres dimensiones:

1. Mejora de los procesos socio organizativos y de gestión de las organizaciones rurales.
2. Mejora de la calidad y cantidad de servicios que ofrecen a sus socios.
3. Manejo sostenible de los recursos naturales y la adaptación a la variabilidad y cambio climático de sus socios.

Fuentes de información y técnicas de recogida de datos

Los insumos básicos de información para la evaluación son de tres tipos:

1) Datos vinculados a la implementación del Proyecto: registros administrativos, documentos y entrevistas con autoridades y referentes de la unidad ejecutora que diseñó e implementa el programa.

2) Datos recabados a través de técnicas cualitativas para una muestra de organizaciones beneficiarias. Para conocer las percepciones de los actores[15] respecto a los resultados y beneficios que obtuvieron por la intervención se están aplicando los siguientes instrumentos:

[15] *La metodología procura que participe la pluralidad de actores relacionados con el PFI: socios, técnicos privados, comisión directiva y gerentes.*

- Entrevistas semi estructuradas y en profundidad, dirigidas a integrantes de las Comisiones Directivas, gerente y el equipo técnico de las organizaciones beneficiarias incluidas en la muestra.
- Grupos focales con socios: se convoca a aproximadamente 10 a 14 socios de cada organización, seleccionados mediante sorteo por el equipo evaluador.

3) Encuesta al conjunto de organizaciones beneficiarias. Al finalizar la ejecución del programa, (en junio de 2017) se aplicará una encuesta de resultados auto administrada a todas las organizaciones beneficiarias.

2.4. Evaluación de Impacto del Programa "Agua para la Producción Animal (APA)"[16]

2.4.1. Descripción del Programa

El Programa "Agua para la Producción Animal (APA)" fue lanzado en febrero de 2011 con el propósito de mejorar las capacidades de los productores para administrar agua y recursos forrajeros, de modo que pudieran enfrentar las sequías en mejores condiciones.

Las sequías constituyen el principal factor de vulnerabilidad para la producción de carne y leche, debido a la importancia del agua como bebida de los animales y, fundamentalmente, por su incidencia en la alimentación animal, a través de su impacto sobre el crecimiento de las pasturas.

Una adecuada gestión del agua para abrevadero, además de asegurar su disponibilidad para los animales, genera condiciones favorables para que los productores adopten cambios en las prácticas de manejo así como para que logren una mayor estabilidad productiva.

La intervención consistió en el otorgamiento de apoyo económico de hasta 7.300 dólares por productor para la realización de las actividades asociadas a las soluciones de agua propuestas por los productores. Conjuntamente al apoyo para inversiones, se agregaron honorarios profesionales para asistencia técnica.

De un total de 2.089 propuestas presentadas fueron aprobadas 1.894, de las cuales resultaron ejecutadas efectivamente 1.755. Del total de proyectos ejecutados, 64% correspondió a productores ganaderos y 36% a productores lecheros. Las principales

[16]La Evaluación de impacto del programa APA fue elaborada por el Área de Evaluación de Políticas Agropecuarias de OPYPA conjuntamente con AGEV-OPP y la Ec. Paola Cazulo, consultora contratada por FAO a tales efectos.

actividades emprendidas fueron inversiones para almacenamiento y distribución (77% del monto total financiado), el alumbramiento de fuentes subterráneas mediante perforación (44%) y la adquisición de equipos de bombeo (41%).

2.4.2. Estrategia de evaluación

El programa terminó de implementarse en 2013, alcanzando un elevado nivel de ejecución y de cumplimiento de metas. Más allá de los buenos resultados en materia de ejecución, el MGAP se propuso analizar los impactos de la intervención sobre el desempeño productivo de los beneficiarios como reflejo de su resiliencia climática, lo que constituye el objetivo último del programa.

Debido a la disponibilidad de información, se consideró oportuno restringir el ámbito de la evaluación a un sub-universo de productores lecheros, dejando para otra etapa el análisis de los impactos sobre los productores ganaderos de carne.

La evaluación se propuso determinar si los apoyos económicos otorgados por el MGAP a los productores lecheros tuvieron efectos sobre la productividad de los beneficiarios, una vez que estos pudieron solucionar parcial o totalmente los problemas relativos al abastecimiento hídrico en la parcela. Las variables de resultado sobre las que se evaluó el impacto del programa fueron: la producción total de leche del establecimiento, la productividad lechera (medida como litros por hectárea lechera, litros por vaca masa y litros por vaca en ordeñe) y el stock de vacas.[17]

El método utilizado fue el de Diferencias en Diferencias con *Propensity Score Matching*, explicado en el anexo metodológico.

Fuentes de información

La evaluación se llevó adelante en base exclusivamente a datos administrativos: las declaraciones juradas de DICOSE de los años 2009 a 2015[18] y la base de datos de la intervención, con información referente a los beneficiarios, sus proyectos y los detalles de su ejecución.

---

[17] Como superficie lechera se consideró la superficie total excepto las hectáreas dedicadas a monte y huerta. Las vacas masa refieren a la suma de vacas en ordeñe y vacas secas

[18] Como es conocido, las declaraciones juradas de DICOSE incluyen: el giro de actividad de la empresa, la superficie explotada, la forma de tenencia de la tierra, los padrones del establecimiento, el uso del suelo, las existencias y la propiedad del ganado. Además, las empresas lecheras declaran la producción anual y el destino de la leche, entre otros datos.

Unidad de análisis

La unidad de relevamiento de la información de DICOSE no siempre coincide exacta ni biunívocamente con un productor, un establecimiento o una explotación agropecuaria. Por el contrario, muchos productores agropecuarios son titulares de más de un registro de DICOSE, mientras que en diversos establecimientos agropecuarios coexisten más de un productor y más de un número de DICOSE. Esta situación complejiza la definición de la unidad de análisis a tener en cuenta como referencia para las evaluaciones de impacto que quieran medir efectos sobre los productores o las explotaciones agropecuarias.

La principal decisión fue la de analizar únicamente el impacto del programa sobre el grupo de establecimientos cuyos titulares se caracterizan por tener un único número de DICOSE, de modo que para este grupo de productores, su número de DICOSE identifica en forma biunívoca al productor, establecimiento o explotación. Sin embargo, ello limita la interpretación de los resultados a dicha sub población.

Otra de las restricciones que se aplicó al panel de datos fue ajustarlo al universo objetivo del programa APA. Esto consiste en truncar la base de datos según las condiciones de elegibilidad establecidas en las bases del llamado.

2.4.3. Algunos resultados preliminares

Antes de la intervención, los productores beneficiarios y los no beneficiarios tenían algunas similitudes en características relevantes: superficie explotada, régimen de tenencia de la tierra y grado de concentración geográfica en la cuenca lechera primaria. Sin embargo, los beneficiarios tenían algunas características más acentuadas que los no beneficiarios en el momento previo al inicio del programa: un rodeo de mayor tamaño, una mayor proporción de remitentes a industria y una mayor proporción de inscriptos en el Registro de Productores Familiares de la DGDR, entre otras.

Entre 2009 y 2015 tanto los beneficiarios como los no beneficiarios registraron un notable dinamismo de la producción, que en promedio creció a una tasa de 4,3% acumulativo anual. Sin embargo, la expansión productiva de los beneficiarios en el período posterior al programa en comparación con el período previo fue sustancialmente mayor que la de los no beneficiarios.

El resultado de la evaluación de impacto muestra indicios de que el programa APA tuvo un impacto positivo sobre la productividad por hectárea lechera: los beneficiarios aumentaron su producción de leche por hectárea en relación a una situación (contrafáctico) en la que el programa no hubiera existido. También se encontraron impactos positivos atribuibles al programa sobre el stock de vacas masa y sobre la producción total de leche de los beneficiarios. No se encontraron, sin embargo, efectos

del programa estadísticamente significativos sobre la producción por vaca. No obstante, las pruebas de robustez ensayadas y las limitaciones que impone la unidad de análisis utilizada, sugieren ser cuidadosos en la interpretación de los resultados.

Es preciso no perder de vista que el ejercicio realizado es una evaluación de alcance local. Esto quiere decir que los efectos encontrados se restringen al conjunto de productores "mono-DICOSES", aquellos que son titulares de un solo número de DICOSE. Esta restricción deja fuera de la evaluación al menos a 30% del total de receptores del programa.

3. Principales logros, lecciones aprendidas y desafíos

En estos dos años y medio desde que comenzó a trabajar el Área de Evaluación de Políticas Agropecuarias se alcanzaron significativos logros, entre los que se destacan:

- **Fortalecimiento de un equipo técnico multidisciplinario dentro de OPYPA** (economistas y antropóloga) especializado en evaluación de políticas y programas, que además se nutre de los conocimientos y experiencia del resto del equipo técnico de OPYPA. Asimismo, se lleva adelante un plan de formación y capacitación de alta especialización, que incluye intercambio con especialistas internacionales y de la región y la academia a nivel nacional. Interesa destacar que para estos logros fue muy valioso el apoyo de FAO -mediante una consultoría para fortalecer la evaluación de impacto del programa Agua para la Producción Animal- así como del Plan Nacional de Adaptación (NAP-Ag, PNUD y FAO), mediante el financiamiento de cursos de especialización en metodologías cuantitativas para la evaluación de impacto y otras técnicas.

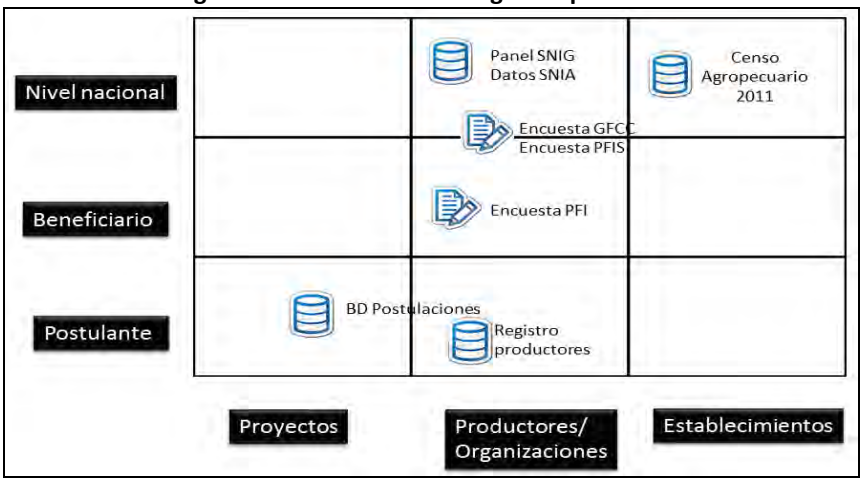
- **Formalización y profundización del Convenio OPYPA-AGEV (OPP)**. En julio de 2015 las autoridades firmaron un convenio marco, que permitió que los equipos técnicos de las dos oficinas trabajaran en forma mancomunada en la evaluación de impacto del programa Agua para la Producción Animal, además del diseño de la evaluación del programa para la Producción Familiar, Integral y Sustentable. Asimismo, AGEV-OPP financiará 400 encuestas ganaderas que permitirán construir la línea de base del PFIS para su evaluación de impacto. El convenio también motivó la realización de varias evaluaciones de Diseño-Implementación-Desempeño (DID) por parte de AGEV a algunas delas evaluaciones que forman parte de la agenda priorizada, lo que aporta una rica complementariedad de enfoques y técnicas.

- **Convenio MGAP-FCS para pasantías, con estudiantes de la licenciatura en desarrollo y ciencia política**. En agosto de 2015 las autoridades del MGAP y de la

Facultad de Ciencias Sociales (UDELAR) firmaron un acuerdo que permite que estudiantes de las licenciaturas en desarrollo y en ciencia política realicen prácticas educativas en el MGAP como parte de su trabajo final (en sustitución de la tradicional monografía o tesis). En el transcurso de 2015 y 2016 tres estudiantes participaron de la evaluación de resultados del programa Propuestas de Fortalecimiento Institucional.

- **Acceso y construcción de fuentes de información integradas.** Desde el Área de Evaluación se ha avanzado en el acceso y armado de una base de datos SQL que integra distintas fuentes de información para las evaluaciones. Es conocido que Uruguay dispone de buenos y extensos registros tanto administrativos como de fuentes estadísticas. Además de los Censos Agropecuarios Nacionales (el último de 2011) y las encuestas agropecuarias que lleva adelante la DIEA periódicamente, el país cuenta con un sistema de declaraciones juradas anuales de ganado y tierra, que computa diariamente las entradas y salidas de ganado (DICOSE-SNIG). Adicionalmente, las unidades ejecutoras de las intervenciones llevan diversos registros administrativos de beneficiarios y postulantes (Registro de Productores Familiares, bases de datos de postulantes y beneficiarios de programas focalizados, etc.).

Figura 3. Bases de datos integradas para evaluación



Fuente: Elaboración propia.

El Área de Evaluación de Políticas Agropecuarias ha integrado bases de datos de los postulantes, de los beneficiarios e información a nivel nacional (Figura 3). De este modo se elaboran indicadores clave para las evaluaciones a nivel micro (por productor).

- **Profundización de los vínculos con la academia, sector privado y sector público, a nivel nacional e internacional.** Desde 2015 el Área de Evaluación de Políticas del MGAP ha venido impulsando en diversos ámbitos la conformación de redes de evaluación que permitan fortalecer capacidades técnicas, adoptar buenas prácticas y desarrollar debates metodológicos. En ese marco, OPYPA propició que el tema evaluación ex post de políticas se incluyera en la agenda del Grupo de Trabajo de Análisis de Políticas (GT1) de REDPA, con el apoyo técnico del IICA. Además, las autoridades del MGAP y del Ministerio de Agricultura de Chile firmaron en noviembre de 2016 un Convenio Marco entre sus oficinas de políticas, OPYPA y ODEPA, para trabajar conjuntamente algunos temas de especial interés, entre los que la evaluación de políticas tiene un lugar preponderante. Por otra parte, el Área de Evaluación de Políticas Agropecuarias se ha beneficiado del apoyo técnico de instituciones internacionales como FAO, PNUD y BID, con quienes se ha iniciado una fructífera relación en torno a la evaluación de políticas agropecuarias. Este entramado de relaciones con oficinas técnicas de otros organismos no puede más que fortalecerse en los próximos años para continuar profundizando la institucionalización de la evaluación de políticas en el MGAP.

Entre las lecciones aprendidas y desafíos identificados en estos primeros dos años y medio de trabajo en el Área de Evaluación de Políticas Agropecuarias se encuentran algunas sugerencias que podrían enriquecer el ciclo de políticas públicas en el MGAP:

- **Mejorar el diseño de las herramientas de política en el MGAP para que sean evaluables**, mediante una mayor integración de las etapas de diseño y evaluación. Muchas veces durante las evaluaciones surgen preguntas y necesidades que, si no fueron definidas en forma completa y precisa durante el diseño de la intervención, hacen muy difícil la evaluación ex post.

  Es de notable utilidad para desarrollar evaluaciones completas y sólidas tener claridad sobre la teoría del programa, que permita relacionar consistentemente las actividades con los cambios esperados en la población objetivo, así como explicitar la focalización y cuantificar las poblaciones objetivo y potencial; a su vez, esto podría mejorar el desempeño de la herramienta de política en sí misma (Agosin et al, 2014).

  Para ello, las diversas áreas de OPYPA pueden contribuir con la elaboración de diagnósticos, información y análisis en la fase de diseño de las herramientas de política. En particular, sería sumamente favorable que el Área de Evaluación de Políticas participe en la formulación de las medidas de política, contribuyendo con evidencia empírica y asegurándose la información requerida para las evaluaciones.

- **Aprovechar la información existente en el MGAP para nutrir las evaluaciones**. Si bien se ha avanzado en el acceso y construcción de un sistema de información específico para las evaluaciones de políticas agropecuarias, hace falta mucho trabajo para que estos grandes volúmenes de información puedan ser utilizados eficientemente para evaluaciones de impacto. Entre los desafíos se encuentra:
    - Coordinar la aplicación de los mismos identificadores en las diversas fuentes de información, a nivel de productor, establecimiento y empresa.
    - Mejorar la disponibilidad de registros administrativos y de inscripción a programas anteriores (por ejemplo para construir trayectorias).
    - Elaborar indicadores de impacto y de resultados que puedan ser relevados a partir de las encuestas de rutina de DIEA (medidas de manejo, características sociales de los productores, etc.).
    - Incorporar en los formularios de postulación de los programas preguntas que permitan medir los avances en las variables objetivo de los beneficiarios.

  La interrelación de las bases de datos disponibles y la mejora de los formularios de encuestas y registros podrían significar un importante ahorro de costos para evaluación de impacto y una notable mejora de la calidad de la información para la toma de decisiones de política agropecuaria.

- **Desplegar técnicas mixtas cuali-cuantitativas** para comprender los efectos y las razones que los explican y poder realizar recomendaciones de política pertinentes. Algunas de las evaluaciones priorizadas se están desarrollando con técnicas cuantitativas y una evaluación se está implementando exclusivamente con técnicas cualitativas. El desafío del Área de Evaluación es desarrollar una caja de herramientas metodológicas adecuada, que permita integrar ambos métodos, y dar pasos hacia el trabajo bajo un enfoque de métodos mixtos.

- **Conveniencia de amalgamar los procesos de formulación y evaluación de las políticas.** El desafío del Área de Evaluación es centrar todos los esfuerzos de evaluación en lograr informes y recomendaciones útiles y orientadas a la mejora de la formulación de las intervenciones. Por eso resulta fundamental que los resultados del trabajo analítico estén disponibles cuando se deban tomar decisiones acerca del futuro de las herramientas de política. A su vez, para que las autoridades puedan tomar en cuenta los resultados de las evaluaciones, los informes deberán incluir recomendaciones claras y prácticas y acompasar, en la medida de las posibilidades, los momentos del ciclo de políticas.

4. Anexo metodológico

Entre las metodologías a utilizar en las evaluaciones ex post de las políticas se destacan los métodos econométricos de evaluación de impacto y las metodologías cualitativas. A continuación se presentan algunas de las herramientas que se están utilizando en el MGAP para la evaluación de los programas.

4.1. Métodos econométricos de evaluación de impacto de políticas

Las metodologías de evaluación de impacto buscan realizar inferencia causal[19] de un cierto tratamiento T, sobre un conjunto de variables de resultados Y, sobre el cual se quiere medir el impacto. Si se supone que el tratamiento de una intervención de política es binario y T es una variable que vale 1 si el individuo fue beneficiario y 0 si no lo fue, entonces el efecto del programa para una unidad i sobre la variable Y, es la diferencia en la variable de respuesta entre los resultados potenciales con y sin el programa:

$$(1) \quad TE_i = Y_i(1) - Y_i(0)$$

Pero en la práctica sólo es posible observar a cada individuo en uno solo de los dos estados de resultados potenciales, por lo que no es posible estimar directamente el efecto del tratamiento para cada individuo i.

El efecto promedio del programa sobre los tratados (ATT por sus siglas en inglés) mide la diferencia entre la media de resultados obtenidos en el grupo de tratamiento E[Y=1|T=1], con respecto a cómo hubieran sido en promedio los resultados de ese grupo en el escenario hipotético contrafáctico de que el programa no hubiera existido E[Y=0|T=1].

$$(2) \quad ATT = E[Y(1)|T = 1] - E[Y(0)|T = 1]$$

Observar que mientras el primer término de la ecuación 2 es fácilmente estimable, el segundo término (cómo les hubiera ido a los individuos que recibieron la intervención si ella no hubiera existido) es inobservable, y se requiere realizar algunos supuestos para poder estimarlo.

El desafío de la evaluación de impacto es cómo construir un grupo con individuos no participantes del programa, llamado grupo de control, que sirva para una correcta comparación con los beneficiarios y cumpla la función del contrafáctico. Una forma de

[19] *Este apartado se basa en el modelo causal de Rubin. Ver Imbens y Rubin (2015) por mayores detalles.*

obtener un buen grupo de control es aplicando métodos experimentales. Es decir, asignando la intervención o programa aleatoriamente entre los individuos de la población objetivo, de modo que, por construcción del método, el hecho de recibir el tratamiento es independiente de la variable de resultado. Si la aleatorización fue exitosa, los grupos de control y de tratamiento deberían ser en promedio idénticos antes de la intervención y la única razón por la que tendrían resultados distintos luego de la medida de política es que ésta haya tenido efectos sobre los beneficiarios.

Cuando la forma en que se asignó el tratamiento de política no fue aleatoria, sino que la adhesión a la medida fue en parte voluntaria para el productor, es usual que el grupo de individuos que no estuvieron afectados por la medida tenga diferencias significativas en sus características previas a la intervención, que los hacen incomparables con los individuos que sí participaron del programa o política. Esto constituye el llamado sesgo de selección.

Así, se recurre a diversos métodos econométricos para sortear este sesgo de selección, seleccionando adecuadamente un conjunto de productores que sirva como grupo de comparación o control del grupo de individuos afectados por la medida de política.

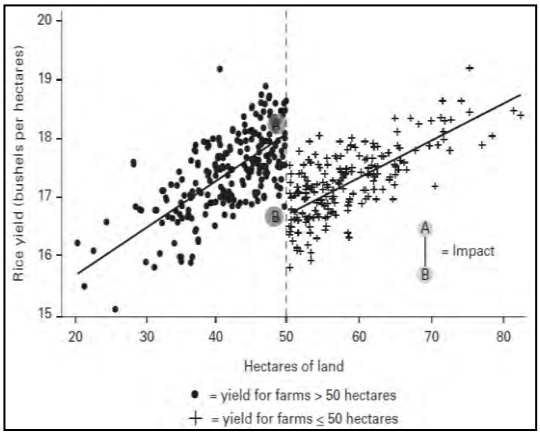
La utilización de uno u otro método depende, por un lado, de las características de la política a evaluar y, por otro, de la disponibilidad de datos con que se cuente a la hora de la evaluación.

Los métodos no experimentales de evaluación de impacto[20] más utilizados en la literatura son: el método de diferencia en diferencias (Dif-en-Dif), el de emparejamiento por la probabilidad de recibir el tratamiento (PSM por sus siglas en inglés), el de variables instrumentales (VI), el diseño de regresión discontinua (RD) y el método de control sintético, entre otros.

[20] Para una introducción a las métodos de evaluación de impacto ver el manual de Gertler et al (2016), o el portal (web) BID de evaluación de impacto. Winters et al. (2012) presentan una introducción centrada en el sector agropecuario.

Regresión Discontinua[21] (RD): explota la presencia de un salto en la probabilidad de ser tratado en una variable continua de elegibilidad (puntaje o *score*), denominado punto de corte o umbral. Bajo ciertos supuestos, los productores por debajo del umbral constituyen un buen contrafáctico que permite estimar el impacto del programa en dicho entorno. El método posee la fortaleza de ser muy creíble cuando se cumplen los test de falsificación[22] y de no imponer supuestos paramétricos[23], pero posee dos grandes limitantes. La primera es que puede requerir un tamaño de muestra grande en el entorno del umbral para detectar impactos estadísticamente significativos. La segunda es que la estimación del impacto obtenida tiene validez local, solamente es representativa del impacto del programa en un entorno del punto de corte.

Figura 4. Regresión Discontinua



Fuente: Gertler et al (2016).

Para ilustrar el método, se considera un programa sobre los productores de arroz que buscan mejorar la productividad de aquellos que poseen menos de 50 hectáreas. Para simplificar se supone que la adhesión al programa es perfecta: todos los productores con menos de 50 hectáreas reciben el beneficio y ninguno con más de 50 hectáreas lo hace. La Figura 4 muestra la relación entre el tamaño del establecimiento (abscisas) y el rendimiento del arroz por hectárea (ordenadas), luego de la intervención. Bajo determinados supuestos las unidades no tratadas a la derecha del punto corte (50 hectáreas) constituyen un buen contrafáctico para las unidades a la izquierda (beneficiarios), por lo que el impacto del programa se mide como la diferencia en el rendimiento medio entre el grupo de control y el grupo de beneficiarios en el punto de corte (la distancia entre los puntos A y B), luego de la medida de política.

[21] Ver Skovron y Titiunik (2015) para una presentación del método.

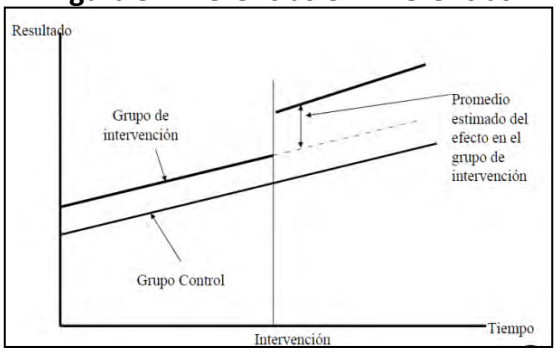
[22] Estas pruebas estadísticas analizan si el puntaje o score fue manipulado estratégicamente y si existen diferencias significativas (previo a la intervención) en alguna variable a ambos lados del punto de corte.

[23] Es no paramétrico porque no se impone una forma funcional concreta.

Diferencias en Diferencias (Dif-en-Dif): este método requiere de datos de panel donde se observe a dos grupos de individuos (tratados y controles) en dos periodos de tiempo.[24] En el período base ningún grupo recibe el tratamiento, mientras en el segundo momento un grupo recibe la intervención. Se compara los cambios en el tiempo (antes y después del tratamiento) en la variable de resultado entre el grupo de participantes del programa (tratamiento) y el grupo de no participantes (control).

El supuesto de identificación del método es que, sin la medida de política, las variables de resultado del grupo de beneficiarios y del grupo de controles tendrían la misma tendencia o que las tendencias serían paralelas.

Figura 5. Diferencias en Diferencias

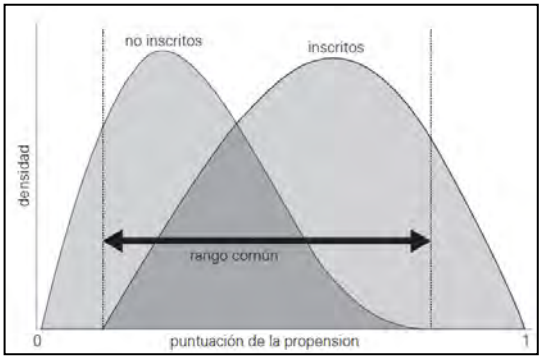


Fuente: Tomado del Seminario Regional de Evaluación de Impacto (Perú 2009).

Para fijar ideas, se supone que la variable de resultado del grupo de tratados no tiene el mismo nivel pero evoluciona en forma paralela a la del grupo de control (ambas presentan la misma pendiente pero parten de un punto inicial distinto). En la Figura 5, se muestra que el impacto de la intervención es la distancia vertical entre la curva de respuesta del grupo de beneficiarios y la línea punteada.

Propensity Score Matching[25] (PSM): Este método empareja participantes y no participantes de un programa con base en la probabilidad condicional de participar o propensión a participar, denominado *propensity score* (PS), dada una serie de características observables[26] de los individuos.

Figura 6. Propensity Score Matching



Fuente: Gertler et al (2016)

[24] Este método es fácilmente generalizable a más de dos periodos de tiempo.
[25] Ver Stuart (2010) para una revisión de la metodología.
[26] Estas características observables deben incluir aquellas variables que se encuentren relacionadas con la probabilidad de ser tratado y con la variable de resultado.

Como se muestra en la Figura 6 el emparejamiento entre observaciones del grupo de tratamiento y de control frecuentemente se puede realizar sólo en una parte de las observaciones, que se denomina el "soporte común". En esa región del *propensity score* es que se compara la variable de resultado sobre la cual se quiere evaluar el efecto del programa, siendo la diferencia en esa variable entre participantes y no participantes emparejados, el impacto estimado del programa dentro del soporte común.

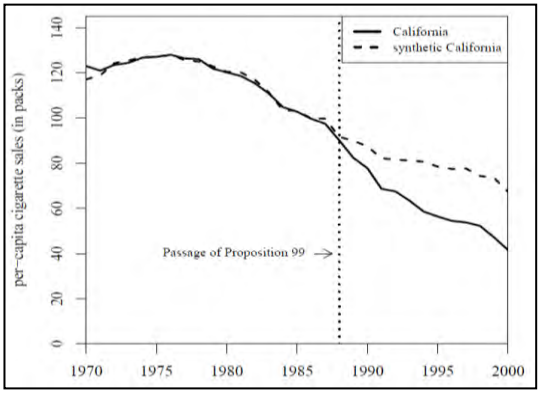
Los métodos de apareamiento en general, PSM en particular, no poseen buenas propiedades cuando existe alguna variable no observable que afecta simultáneamente la decisión de participar en el programa y la variable de resultado de interés. A su vez el método de Dif-en-Dif, no hace explícito el balance entre el grupo de tratamiento y de control. Sin embargo, si existe información de línea de base, es posible combinar un **PSM con Dif-en-Dif.** En este último caso es posible controlar por características inobservables que se mantengan constantes en el tiempo (al analizar la variable de resultado en diferencia), comparando a unidades similares en características observables (dentro del soporte común).

El método de **Variables Instrumentales** asume que existe (por lo menos) una variable conocida como instrumento (z), que se encuentra relacionada con la probabilidad de ser tratado pero no con la variable de resultado (y). Este método tiene la fortaleza de permitir que controles y tratados sean diferentes en factores inobservables (u), pero la debilidad que para su funcionamiento requiere de un buen instrumento y de supuestos que no pueden ser testeados en la realidad.

Como ejemplo supongamos un programa donde los beneficiarios son distintos a los productores que no participan en alguna característica inobservable "u" (por ejemplo: motivación), correlacionada simultáneamente con la probabilidad de participar en el programa y con la variable de resultado ("y", por ejemplo: productividad). Para estimar el impacto del programa debemos poseer de una fuente exógena de variabilidad en el tratamiento, es decir de un instrumento ("z"). Un posible instrumento es brindar de forma aleatoria información sobre el programa, de modo que recibir la información adicional esté correlacionado con la participación en el programa, pero por construcción no estará correlacionado con la variable de resultado. Con VI estimamos el impacto de aquellos productores que cambian su decisión de participar de acuerdo al instrumento.

El método de **Control Sintético** (CS) fue concebido para analizar el impacto a niveles agregados como un país o una región. La idea del método es construir un contrafáctico para la región tratada, usando una combinación de zonas de control (por ejemplo, otros países o regiones que no recibieron la intervención), siempre que este grupo de control así construido satisfaga el supuesto de tendencias paralelas entre tratados y controles antes de la intervención.

Figura 7. Control Sintético



Fuente: Abadie et al. (2012).

Para ilustrar el método, se presenta el ejemplo analizado por Abadie et al. (2012), en el que se quiere estimar el efecto de la suba de impuestos al tabaco en California sobre el consumo per cápita. En la Figura 7, la curva sólida muestra el consumo de tabaco en California y la línea punteada el CS del consumo de tabaco en California sin la suba de impuestos. El CS se construye utilizando una ponderación de otros estados de EEUU. Estos ponderadores surgen de minimizar la distancia en la trayectoria previa del consumo de tabaco (per cápita) antes de la intervención entre California y el CS, en características previas a la intervención.

4.2. Métodos cualitativos de evaluación

Los métodos cualitativos tratan de comprender los procesos, comportamientos y condiciones como las perciben los individuos, grupos u organizaciones evaluadas. Proporcionan información sobre las percepciones de los actores respecto a los resultados y beneficios que obtuvieron en el marco de las intervenciones de política. Una de las ventajas de las evaluaciones cualitativas es su flexibilidad, ya que pueden ser adaptadas específicamente a las necesidades de la evaluación. Ello permite mejorar los resultados de las herramientas de política al proporcionar una mayor comprensión de las percepciones y prioridades de las partes interesadas y las condiciones y procesos que pueden haber afectado los resultados. Según Strauss y Corbin (1990), los resultados de la investigación cualitativa pueden usarse para: aclarar e ilustrar hallazgos cuantitativos, desarrollar política pública y evaluar programas, entre otros.

Los métodos cualitativos se caracterizan por integran una variedad de técnicas para recoger información. En las evaluaciones de las intervenciones que se están desarrollando en el Área de Evaluación de Políticas Agropecuarias se destacan: grupos focales y entrevistas semiestructuradas.

Grupos focales. El propósito principal del grupo focal es hacer que surjan, en los participantes, actitudes, sentimientos, creencias, experiencias y reacciones relativas a la intervención; esto no sería fácil de lograr con otros métodos. Además, comparados con la entrevista individual, los grupos focales permiten obtener una multiplicidad de miradas y posiciones dentro del contexto del grupo (Gibb, 1997). Por otra parte, el análisis a partir

del grupo focal se centra en los datos que surgen de la interacción dentro del grupo, la cual gira alrededor del tema propuesto por el moderador. Las discusiones en los grupos focales le ofrecen al moderador la flexibilidad necesaria para explorar asuntos que no hayan sido anticipados y acceder a una mayor cantidad de información en un periodo corto de tiempo.

Respecto al número ideal de personas que deben conformar un grupo focal, diferentes autores han propuesto una gran variedad de rangos de participantes. Myers (1998) propone que el número adecuado de participantes se determine según las características y la complejidad del tema a tratar, así como la experiencia y la habilidad del moderador, sin exceder las 12 personas. El grupo de participantes puede ser homogéneo o heterogéneo, dependiendo del estudio; cuando se trata de explorar experiencias compartidas, se aconseja un grupo homogéneo. Además, si el tema lo amerita, puede ser importante escoger personas que no se conozcan entre sí para evitar que se hagan suposiciones o que haya prejuicios.

El primer paso es el diseño de preguntas para el grupo. Estas deben ser abiertas, de forma que permitan a los participantes responder ampliamente. Deben evitarse preguntas que puedan ser resueltas con un sí o un no; por el contrario, es mejor usar preguntas que permitan a los participantes recordar, comparar y clasificar. Además, se recomienda usar preguntas de cierre que permitan extraer conclusiones y resumir sus comentarios.

El rol del moderador es fundamental, ya que debe conducir a los participantes a una conversación. Las responsabilidades básicas de éste son: mantener la discusión en un solo tema y ritmo, involucrar a los individuos en la conversación y establecer un ambiente de confianza que elimine las barreras de la comunicación, utilizar estrategias de control de grupo sutiles para manejar a los “expertos”, a los que tienden a dominar la conversación, a los tímidos, a los confusos, entre otros. Al inicio del grupo focal, el moderador deberá describir el protocolo de la discusión e indicar que la conversación será grabada. Además, debe asegurar a los participantes que habrá confidencialidad y que los nombres no serán revelados en las citas que se utilicen para los reportes posteriores (Beck, et al., 2004; Mayan, 2001)

Entre las características principales de los grupos focales se destaca el hecho de que es una técnica versátil. Se pueden usar en combinación con otros métodos de investigación, como complemento a otros medios para recopilar datos, para triangular o validar la información, o para generar las preguntas apropiadas para una encuesta o un cuestionario.

Entrevistas semiestructuradas. En este caso el entrevistador dispone de un «guion», que recoge los temas que debe tratar a lo largo de la entrevista. Sin embargo, el orden en el que se abordan los diversos temas y el modo de formular las preguntas se dejan a la libre

decisión y valoración del entrevistador, quien, en el ámbito de un determinado tema, puede plantear la conversación como desee, efectuar las preguntas que crea oportunas y hacerlo en los términos que estime convenientes, explicar su significado, pedir al entrevistado aclaraciones cuando no entienda algún punto y que profundice en algún extremo cuando le parezca necesario, estableciendo un estilo propio y personal de conversación.

Análisis de los datos: Después de recogidos y transcriptos, los datos deben ser organizados para encontrarles "sentido".

El análisis cualitativo de datos no es estándar, sino que cada evaluación requiere de un análisis propio o "coreografía de análisis". Es un proceso de reinterpretaciones sucesivas, con el propósito de descubrir conceptos y relaciones entre los datos brutos y organizarlos en un esquema interpretativo. Para ello es necesario un proceso de codificación que se realiza desarrollando conceptos tentativos, que el investigador considera de relevancia potencial para el problema que está estudiando, y etiquetando el texto. Esta fase descriptiva sienta las bases para la etapa interpretativa en la cual se extraen los significados a partir de los datos, se hacen comparaciones y se construyen marcos interpretativos.

En el curso del trabajo un analista puede hallar docenas de conceptos y observar que ciertos conceptos se pueden agrupar bajo un orden abstracto más elevado, basado en su capacidad de explicar lo que está sucediendo. El concepto que agrupa no sólo permite clasificar los objetos, sino que también explica lo que hacen; esto constituye la categoría**.**

La fiabilidad y validez de la investigación cualitativa tiene uno de sus puntos de apoyo en la triangulación y la saturación de la información. Se utilizan múltiples herramientas y diversos informantes para poder cruzar la información obtenida en los diversos niveles, hasta llegar a un punto donde se comienza a producir redundancia de la información. De esta manera se puede afirmar que la información refleja la realidad investigada. La **saturación teórica** sucede cuando los nuevos datos no introducen cambios en las dimensiones de las categorías o conceptos. No hay una norma de corte, pero es importante saturar la información obtenida.

Por otra parte, una herramienta de evaluación mixta que combina información cuantitativa y cualitativa son las **Evaluaciones de Diseño, Implementación y Desempeño (DID)** desarrolladas por la Dirección de Gestión y Evaluación (AGEV-OPP). Las Evaluaciones DID permiten identificar fortalezas y oportunidades de mejora que condicionan la posibilidad de alcanzar los resultados esperados, optimizando tiempo y recursos. Se basan en información existente en los propios organismos públicos (cuantitativa y cualitativa) que es complementada con visitas de campo y entrevistas a informantes clave a efectos

de ayudar a alcanzar una visión integral de las intervenciones públicas y los factores asociados a sus niveles de desempeño.

Al igual que en la evaluación cuantitativa, en el marco de la evaluación cualitativa, la confiabilidad es la probabilidad de obtener los mismos resultados dentro de un mismo periodo de tiempo. La triangulación de la información obtenida de las diversas fuentes de datos ayuda a mejorar la objetividad y verificabilidad de los indicadores cualitativos, minimizando sus sesgos.

La dificultad de establecer y consensuar criterios homogéneos en la investigación cualitativa se debe fundamentalmente a la "pluralidad" de corrientes (Positivismo, Postpositivismo, Teoría crítica, Constructivismo, etc.), enfoques metodológicos o métodos (fenomenología, etnografía, teoría fundamentada, análisis del discurso, investigación acción, etc.).

Los métodos pueden complementarse entre sí y en algunos casos se recomienda el uso de métodos mixtos (cualitativos y cuantitativos) para dar mayor robustez a los resultados. Es ideal combinar métodos cuantitativos y cualitativos, porque esto proporcionará el efecto cuantificable de un proyecto y una explicación de los procesos e intervenciones que originaron estos resultados.

5. Lista de referencias

Abadie, A., Diamond, A., & Hainmueller, J. (2012).Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program. *Journal of the American statistical Association*.

Agosin, Manuel R.; Fernández-Arias, Eduardo; Crespi, Gustavo; Maffioli, Alessandro; Rasteletti, Alejandro; Wagner, Rodrigo; Stein, Ernesto H.; Bassi, Marina; Rucci, Graciana; Urzúa, Sergio; Panizza, Ugo; de Olloqui, Fernando; Casaburi, Gabriel; Pietrobelli, Carlo; Blyde, Juan S.; Volpe Martincus, Christian; Cornick, Jorge; Trejos, Alberto (2014) ¿Cómo repensar el desarrollo productivo?: Políticas e instituciones sólidas para la transformación económica .

Beck, M., Bryman, A. y Futing, L. (2004). The Sage Encyclopedia of Social Science Research Methods. New Delhi: SAGE Publications.

Gertler, Paul J.; Martinez, Sebastian; Premand, Patrick; Rawlings, Laura B.; Vermeersch, Christel M. J.(2016). Impact Evaluation in Practice, Second Edition. Washington, DC: Inter-American Development Bank and World Bank.

Gibb, A. (1997) Focus group. Social Research Update, 5 (2).

Imbens, G. W., & Rubin, D. B. (2015). Causal inference in statistics, social, and biomedical sciences. Cambridge University Press.

Kitzinger J. (1995) 'Introducing focus groups', British Medical Journal 311: 299-302.

Mayan, M. (2001). Una introducción a los métodos cualitativos: módulo de entrenamiento para estudiantes profesionales. Tesis de Grado. Universidad Nacional Autónoma de México.

Meyers, G. (1998) Displaying opinions: topics and disagreement in focus groups. Language in Society, 27 (1), 85-111.

Winters, Paul; Salazar, Lina; Maffioli, Alessandro (2012). Designing Impact Evaluations for Agricultural Projects. Office of Strategic Planning and Development Effectiveness, Inter-American Development Bank (BID).

SINERGIA (2012). Guía para la Evaluación de Políticas Públicas Departamento Nacional de Planeación (Colombia).

Skovron, C., & Titiunik, R. (2015). A Practical Guide to Regression Discontinuity Designs in Political Science. working paper, University of Michigan.

Strauss, A., Corbin, J. (1990). Basics of qualitative research. Thousand Oaks.

Stuart, E. A. (2010). Matching methods for causal inference: A review and a look forward. Statistical science: a review journal of the Institute of Mathematical Statistics