

# Curso de Econometría de Evaluación de Impacto: Unidad 1

---

Andrés Mendoza – Viviana Carriel

Basado en:

Handbook on Impact Evaluation: Quantitative methods and Practices (2010)

# 1. Introducción a la Evaluación de Impacto

---

# ¿Por qué evaluar?



No se suele abordar si los programas alcanzaron los resultados propuestos

# Evaluación de resultados vs. Evaluación de Impacto

**Evaluación:** Son exámenes periódicos y objetivos de un proyecto o política (programada, en curso o completada).

Se usan para responder preguntas específicas relacionadas con el diseño, la ejecución y los resultados.

Las evaluaciones pueden abordar tres tipos de preguntas:

1. *Descriptivas* → ¿Qué está ocurriendo?
2. *Normativas* → Compara lo que está ocurriendo con lo que debería ocurrir.
3. *Sobre causa y efecto* → Examina los resultados e intenta determinar la diferencia que la intervención ejerce sobre estos resultados.

*La evaluación de resultados se centra en responder a las preguntas 1 y 2.*

## Evaluación de resultados vs. Evaluación de Impacto

La **Evaluación de Impacto** consiste en evaluar los cambios en el bienestar de las personas que pueden atribuirse a un proyecto, programa o política particular.

¿Cuál es el impacto o efecto causal directamente *atribuible* al programa o política?

Cualquier método elegido debe estimar el denominado *contrafactual*. En la práctica se utiliza un *grupo de comparación*.

## Evaluación de Impacto para decidir sobre políticas

- Ayuda a responder a la pregunta: ¿es un programa efectivo comparado con la ausencia del mismo?
- Cuando un programa se puede implementar de varias maneras: ¿Cuál es la más efectiva?

*Discusión: opinión estudiantes.*

# La decisión de evaluar

No todos los programas justifican una evaluación de impacto.

Para justificar la movilización de recursos técnicos y financieros para una evaluación de impacto de gran calidad, el programa a evaluar debe ser:

- **Innovador:** pone a prueba un nuevo enfoque.
- **Replicable:** puede ampliarse y aplicar a una situación diferente.
- **Estratégicamente pertinente:** es una iniciativa emblemática, cubre a un gran número de personas o podría generar un ahorro considerable.
- **No comprobado:** se sabe muy poco sobre la efectividad del programa a nivel mundial.
- **Influyente:** los resultados se usarán para informar decisiones fundamentales de políticas.

# Evaluación prospectiva vs. Evaluación retrospectiva

- Las evaluaciones prospectivas se realizan al mismo tiempo que se diseña el programa y forman parte de la implementación del mismo.
- Las evaluaciones prospectivas de impacto tienen más probabilidades de producir resultados más sólidos y creíbles, debido a que tiene más probabilidades de generar contrafactuales válidos.

Información de línea base



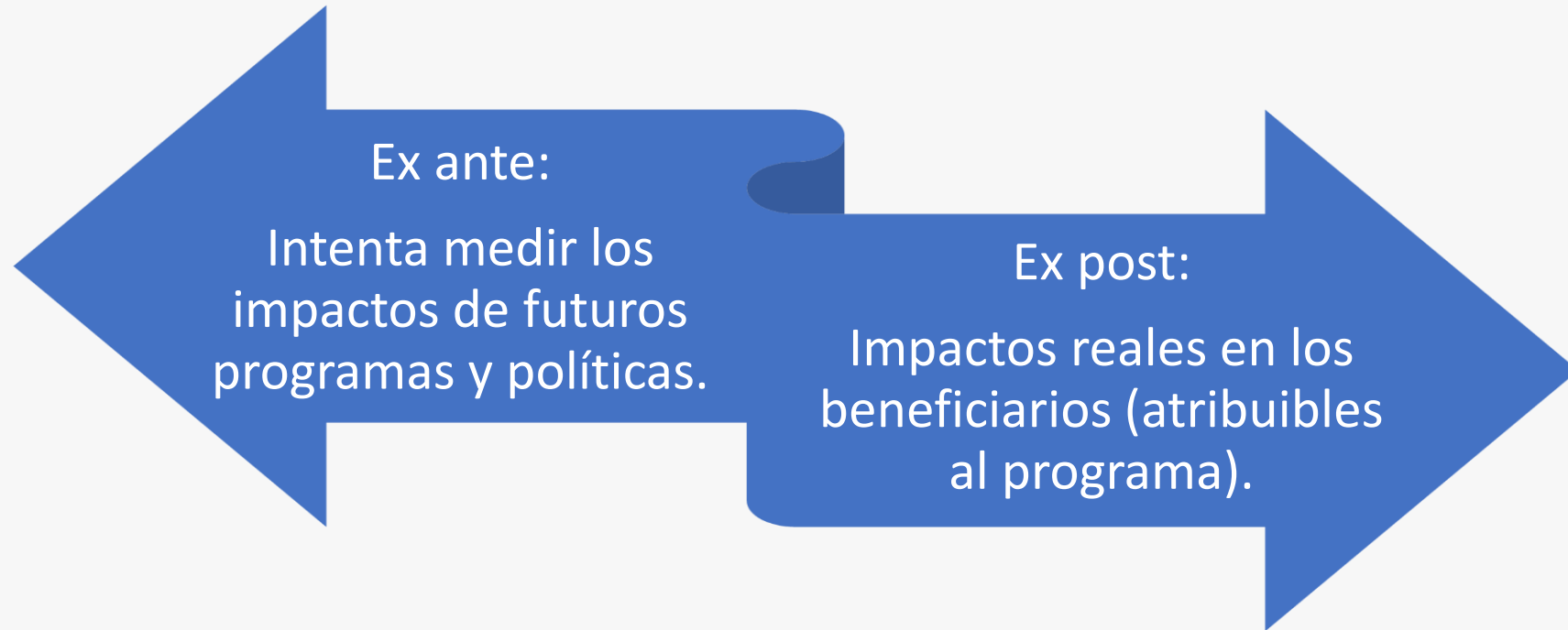
# Evaluación prospectiva vs. Evaluación retrospectiva

- Las evaluaciones retrospectivas examinan el impacto del programa después de su implementación, lo que genera grupos de tratamiento y de comparación ex post.
- Como consecuencia, la viabilidad de una evaluación retrospectiva depende del contexto y nunca está garantizada. Incluso cuando son factibles, las evaluaciones retrospectivas suelen usar métodos cuasi experimentales y depender de mayores suposiciones, lo que hace que puedan generar evidencias más discutibles.
- Los diseños de investigación cuasiexperimentales contrastan hipótesis causales (como los experimentales). El programa o política se considera como una «intervención» en la que se comprueba en qué medida un tratamiento logra sus objetivos.

# Evaluación prospectiva vs. Evaluación retrospectiva

- Un diseño cuasiexperimental carece, por definición, de distribución aleatoria. La asignación a las condiciones (tratamiento versus ningún tratamiento o comparación) se lleva a cabo por autoselección (los participantes eligen el tratamiento), por la selección efectuada por los administradores (por ejemplo, funcionarios, profesores, autoridades, etc.) o por ambas vías.
- Los diseños cuasiexperimentales identifican un grupo de comparación lo más parecido posible al grupo de tratamiento en cuanto a las características del estudio de base (previas a la intervención).

# Evaluación de impacto ex ante vs. Evaluación de impacto ex post



# Evaluación de impacto contrafactual

- Establecer si una determinada intervención produce los efectos deseados en alguna dimensión de interés preestablecida.
- El objetivo primordial es responder a una pregunta "¿hizo la diferencia?", identificando y estimando los efectos causales a través de métodos contrafácticos.
- La cuestión central de la Evaluación de impacto contrafactual es: ¿cuánta diferencia produce un tratamiento?

# Evaluación de impacto basada en teoría (TBIE)

- En TBIE, la causalidad es a menudo vista como un problema de contribución, no de atribución. A menudo se cita el análisis de la contribución causal (Mayne, 2001; Leeuw, 2003), que pretende demostrar si la intervención evaluada es o no una de las causas del cambio observado.
  - El análisis de la contribución se basa en cadena de argumentos lógicos que se verifican a través de un cuidadoso trabajo de campo. Implica la identificación sistemática y la investigación de explicaciones alternativas para los impactos observados.
- 
- ¿Por qué el programa funciona? Qué factores inciden en su éxito?
  - Abarca más allá de una evaluación contrafactual, es integral.
  - Si esto funciona ¿por qué funciona?

# Evaluación de impacto basada en teoría (TBIE)

- La importancia de TBIE se deriva del hecho de que una gran cantidad de información adicional, además del efecto causal cuantificable, es útil para los responsables de formular políticas y tomar decisiones ante los ciudadanos.
- La cuestión de por qué un conjunto de intervenciones produce efectos, tanto intencionados como no intencionados, para quién y en qué contexto, es tan relevante, importante e igualmente desafiante, si no más, que la cuestión "hizo la diferencia".
- La evaluación basada en la teoría es un enfoque en el que se presta atención a las teorías de los encargados de formular políticas; es decir, recopilaciones de suposiciones e hipótesis (empíricamente probadas) que están enlazadas lógicamente entre sí.

# Evaluación de impacto basada en teoría (TBIE)

- Un elemento esencial de la eficacia de las políticas son los mecanismos que hacen que la intervención funcione. La evaluación basada en la teoría explora los mecanismos que los creadores de políticas creen que hacen efectiva la política y la compara con la evidencia basada en la investigación.
- La evaluación basada en la teoría tiene en su núcleo dos componentes vitales. El primero es conceptual, el segundo empírico:

*Conceptualmente*, las evaluaciones basadas en la teoría articulan una teoría de políticas o programas.

*Empíricamente*, las evaluaciones basadas en la teoría buscan probar esta teoría, investigar sí, por qué o cómo las políticas o programas causan resultados previstos u observados.

## Teorías del Cambio:

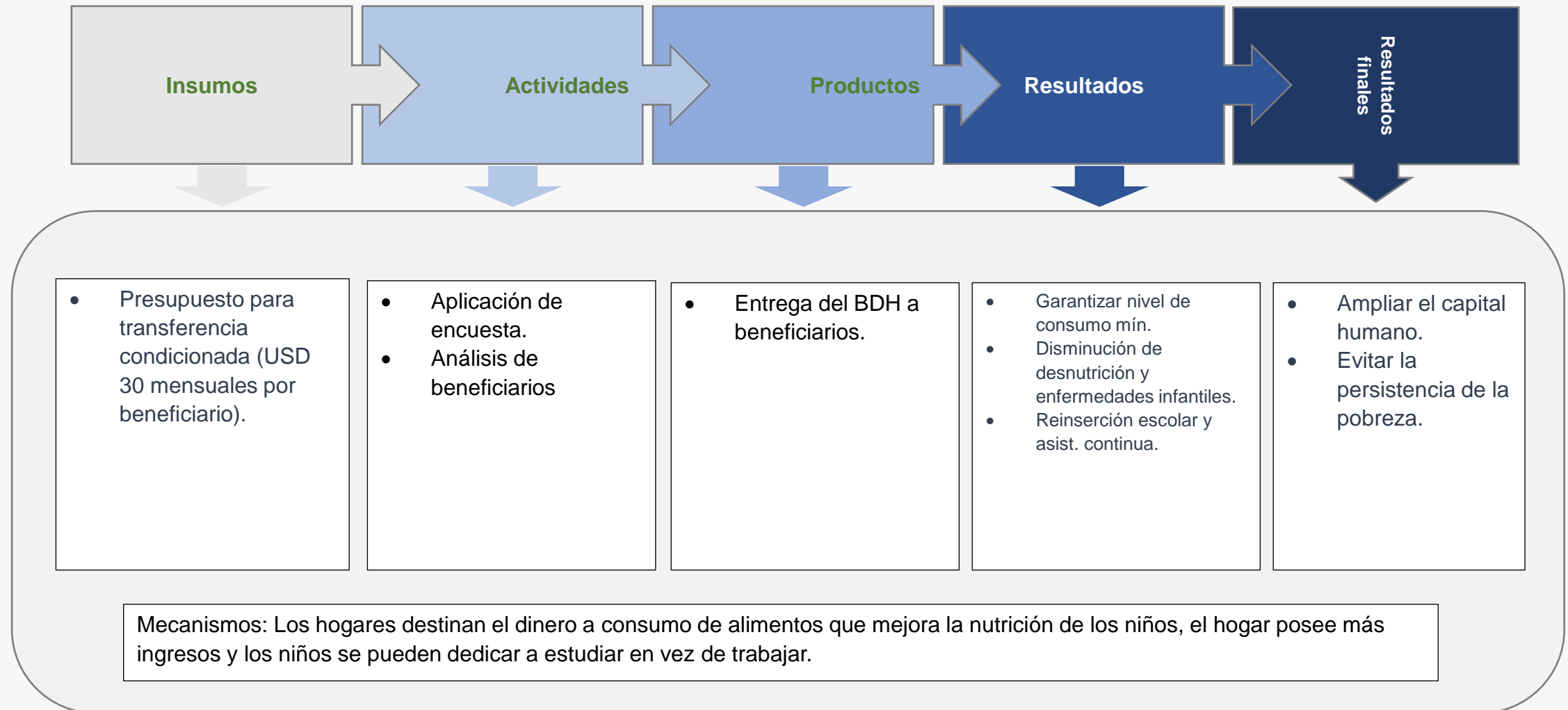
- Una teoría del cambio es una descripción de cómo se supone que una intervención conseguirá los resultados deseados.
- El mejor momento para desarrollar una teoría del cambio para un programa es al principio del diseño, cuando se puede reunir a las partes interesadas para que definan una visión común del programa, de sus objetivos y de la vía para lograrlos.



## La cadena de resultados:

- Una teoría del cambio puede diseñarse de varias maneras, por ejemplo, mediante modelos teóricos, modelos lógicos, marcos lógicos, modelos de resultados y cadenas de resultados.
- Una cadena de resultados da una definición lógica y plausible de cómo una secuencia de insumos, actividades y productos relacionados directamente con el proyecto interactúan y establecen las vías por las que se logran los impactos.

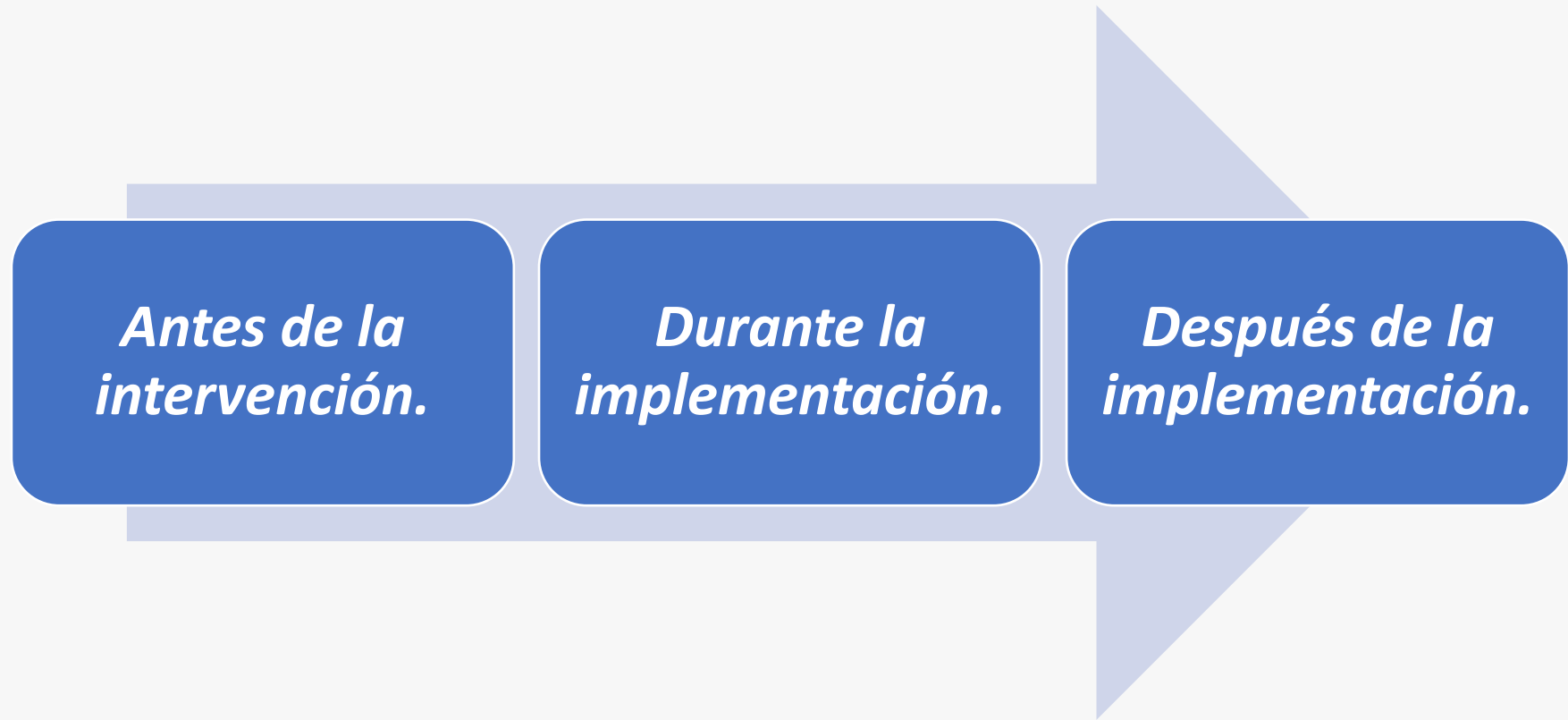
## La cadena de resultados, Bono de Desarrollo Humano:



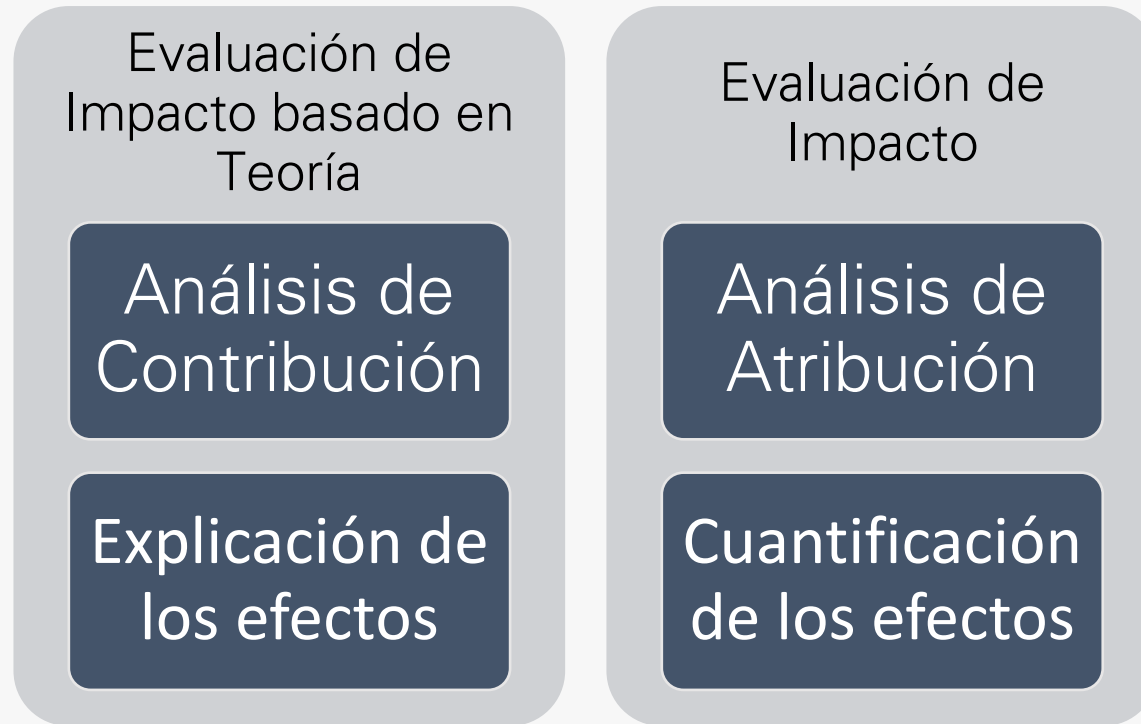
## Hipótesis para la evaluación:

- Una vez que se haya descrito la cadena de resultados, se puede formular la hipótesis que se quiere comprobar mediante la evaluación de impacto.
- El BDH contribuye a una mayor inserción escolar de los niños en los hogares beneficiarios.

¿Qué y cuándo puede contribuir la evaluación basada en teoría?



# Evaluación de impacto contrafactual vs. El basada en teoría





## 2. El Modelo Contrafactual

---

# Inferencia Causal

- La pregunta básica de la evaluación de impacto es esencialmente un problema de inferencia causal.
- Evaluar el impacto de un programa sobre una serie de resultados es equivalente a evaluar el efecto causal del programa sobre ellos.
- Aunque las preguntas sobre causa y efecto son habituales, determinar que una relación es causal no es tan sencillo.
- Las evaluaciones de impacto nos ayudan a atribuir causalidad al establecer empíricamente en qué medida cierto programa, y *solo ese programa*, ha contribuido a cambiar un resultado.



- La respuesta a la pregunta básica de la evaluación de impacto, ¿Cuál es el impacto o efecto causal de un programa P sobre un resultado de interés Y?, se obtiene mediante la fórmula básica de la evaluación de impacto:

$$\alpha = (Y|P = 1) - (Y|P = 0)$$



Impacto  
causal

- Pensemos en un ejemplo:

P representa un programa de formación profesional.

Y representa el ingreso.

# Inferencia Causal

- Al comparar a la misma persona consigo misma, en el mismo momento, se eliminaría cualquier factor externo que pudiera explicar también la diferencia en los resultados. En este caso se podría confiar en que la relación entre el programa de formación profesional y el ingreso es causal.
- La fórmula básica de la evaluación de impacto es válida para cualquier cosa que se analice: una persona, un hogar, una comunidad, un negocio, una escuela, un hospital o cualquier otra unidad de observación que pueda beneficiarse o verse afectada por un programa. La fórmula también es válida para cualquier resultado (Y) que esté relacionado con el programa en cuestión.

## *¿Cuál es el problema del contrafactual?*

- Es imposible medir a la misma persona en dos situaciones diferentes en el mismo momento. Una persona habrá participado en un programa, o no lo habrá hecho.
- ¿Cómo se mide lo que habría ocurrido si hubiera prevalecido la otra circunstancia?
- Aunque se puede observar y medir el resultado ( $Y$ ) para los participantes en el programa ( $Y \mid P = 1$ ), no se cuenta con datos para establecer cuál habría sido el resultado en ausencia del programa ( $Y \mid P = 0$ ) (el contrafactual).

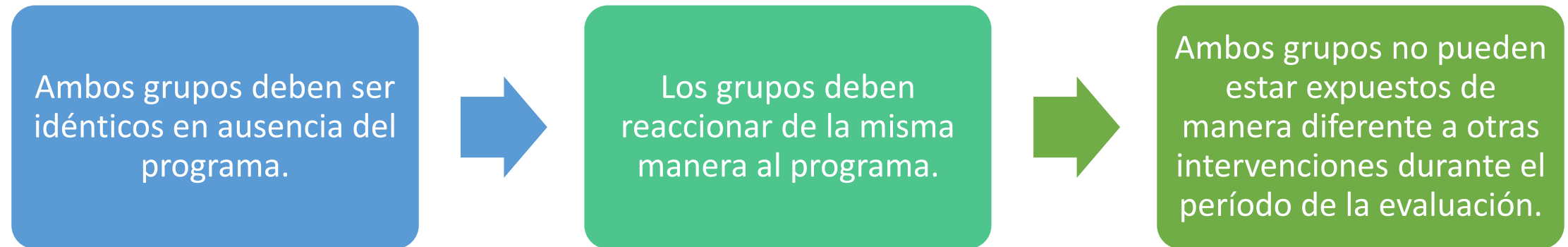
$$\alpha = (Y|P = 1) - (Y|P = 0)$$



- Para ello se recurre habitualmente a grupos de comparación (denominados también “grupos de control”).
- La identificación de esos grupos de comparación es lo medular de cualquier evaluación de impacto, independientemente del programa que se evalúe.

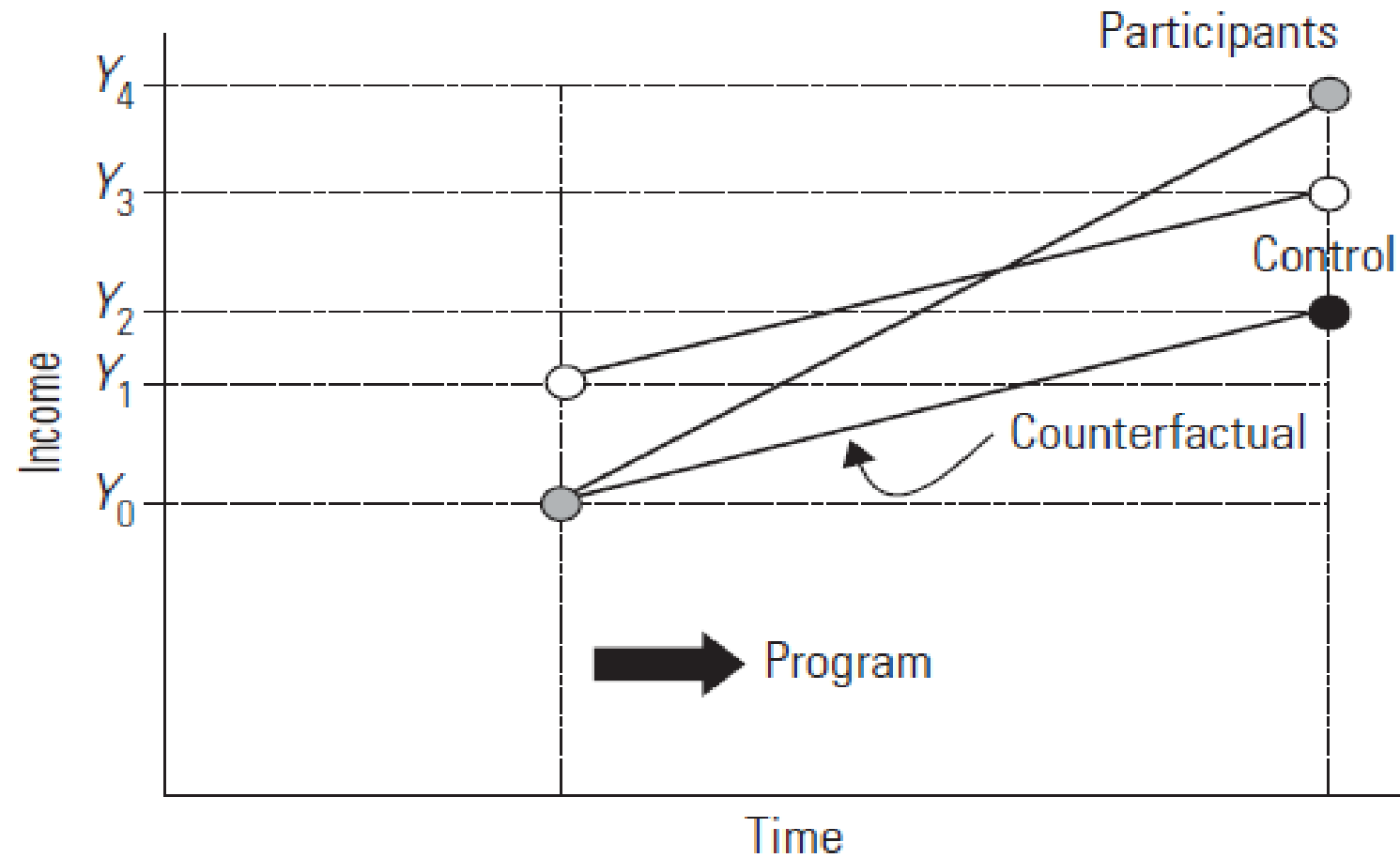
# Estimación del Contrafactual

El reto principal es identificar a un grupo de comparación con las mismas características que el grupo de tratamiento. Al menos en tres aspectos deben ser iguales:



# Dos estimaciones falsas del contrafactual

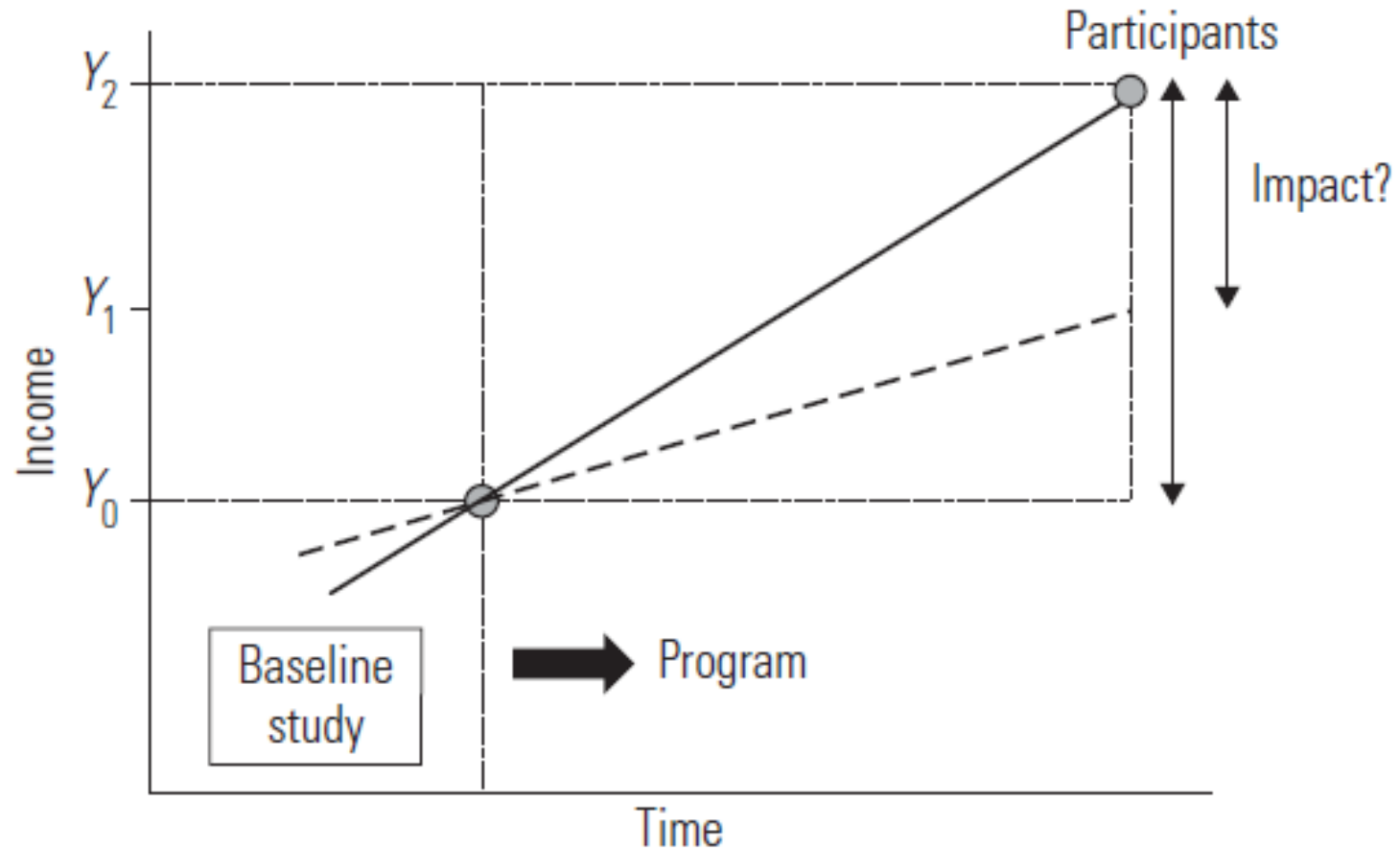
## *Participantes y no participantes*



Fuente: Khandker et al. (2009)

# Dos estimaciones falsas del contrafactual

## *Comparaciones antes y después de los tratados*



Fuente: Khandker et al. (2009)

# El problema del sesgo de selección

- La evaluación de impacto es en esencia un problema de falta de datos, puesto que no se observa el resultado de las personas de no haber participado en el programa.
- Las evaluaciones de impacto exitosas dependen de encontrar un buen grupo de comparación!
- Hay dos enfoques generales a los que recurren los investigadores para imitar el contrafactual de un grupo tratado:
  - (a) crear un grupo de comparación a través de un diseño estadístico, o
  - (b) modificar la estrategia de focalización del programa en sí, para eliminar las diferencias que han existido entre los grupos tratados y no tratados antes de comparar los resultados entre los dos grupos.



# El problema del sesgo de selección

La siguiente ecuación presenta el problema básico de evaluación de los resultados a través de individuos tratados y no tratados.

$$Y_i = \alpha X_i + \beta T_i + \varepsilon_i$$

El problema en esta ecuación es que la asignación del tratamiento no es con frecuencia aleatorio. Esto puede deberse a:

1. Asignación intencional en el programa (éstos se brindan en función con la necesidad de las comunidades y los individuos).
2. Autoselección en el programa (podría basarse en características observadas, inobservadas, o ambas).

En el caso de las características inobservadas, éstas estarán contenidas en el término de error y podrían estar correlacionadas con la variable dummy de tratamiento (variable omitida).

$$Y_i = \alpha X_i + \beta T_i + \varepsilon_i$$

$$Cov(T, \varepsilon) \neq 0$$

El hecho de que el término de error esté correlacionado con la variable de tratamiento, introducirá un sesgo en la estimación del efecto del programa ( $\beta$ ).