

Regiones +  
Competitivas




Colombia+  
Competitiva

Cooperación suiza para la competitividad

Programa de fortalecimiento de capacidades  
Módulo 9: “Evaluación de impacto” - Conceptos básicos

  
swisscontact

 Schweizerische Eidgenossenschaft  
Confédération suisse  
Confederazione Svizzera  
Confederaziun svizra  
  
Embajada de Suiza en Colombia  
Cooperación Económica y Desarrollo (SECO)

technopolis  
group

 Competitiveness

 OCYT  
OBSERVATORIO COLOMBIANO DE  
CIENCIA Y TECNOLOGÍA

 COLOMBIA  
POTENCIA DE LA  
VIDA

 Departamento Nacional  
de Planeación - DNP

# Test de entrada



Colombia+  
Competitiva

Cooperación suiza para la competitividad



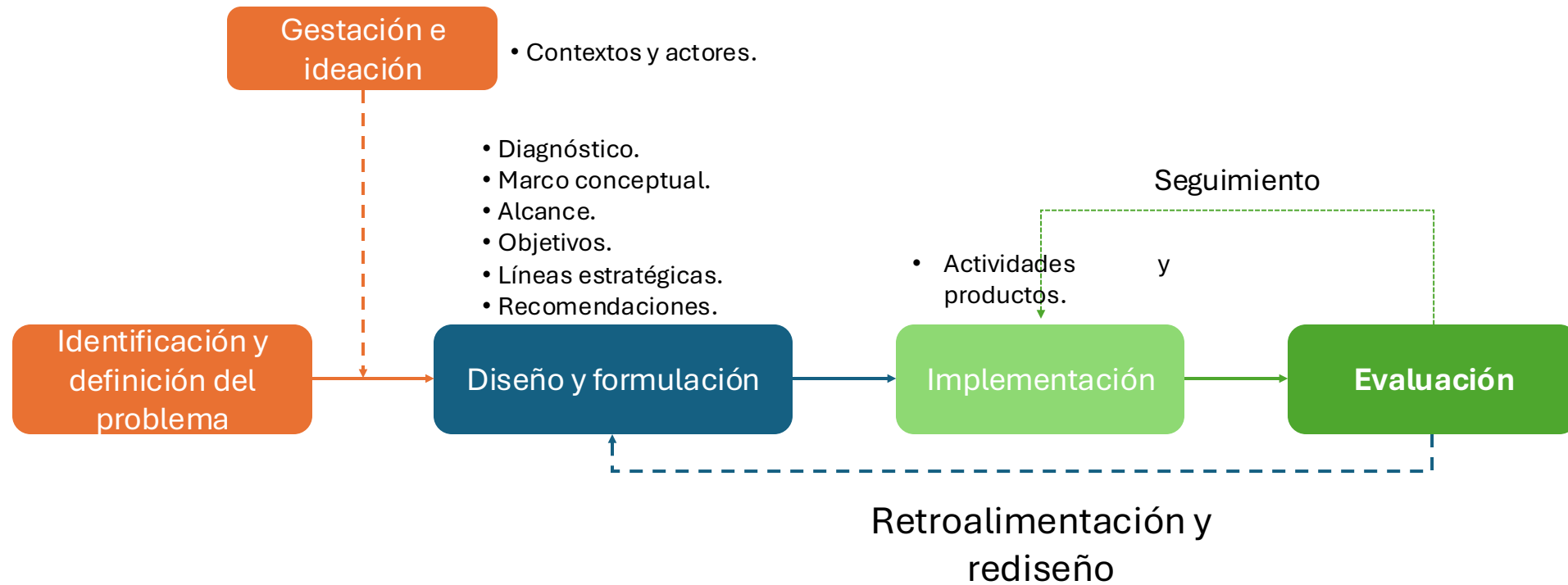


# Agenda Módulo 9: “Evaluaciones de impacto”

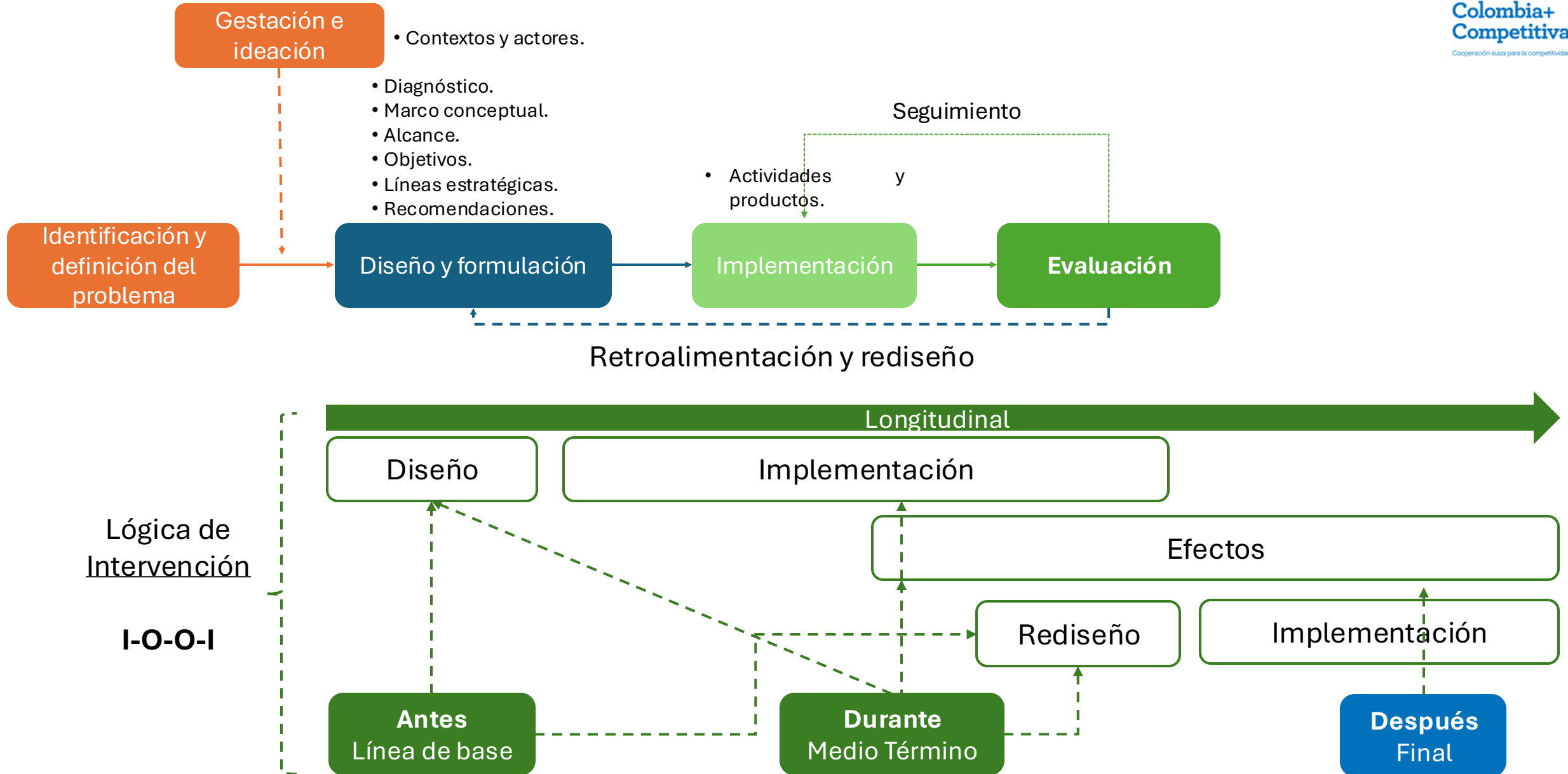
Conceptos básicos para el diseño e implementación de evaluaciones de impacto.

1. **¿Qué es la evaluación de impacto? ¿Qué es la inferencia causal?**
2. Diseños experimentales.
3. Diferencia en Diferencias.
4. Emparejamiento.
5. Regresión discontinua.
6. ¿Qué técnica usar y cuándo?
7. ¿Evaluar o no evaluar?
8. ¿Cómo estructurar una evaluación de impacto desde el diseño?  
Retos y realidades.

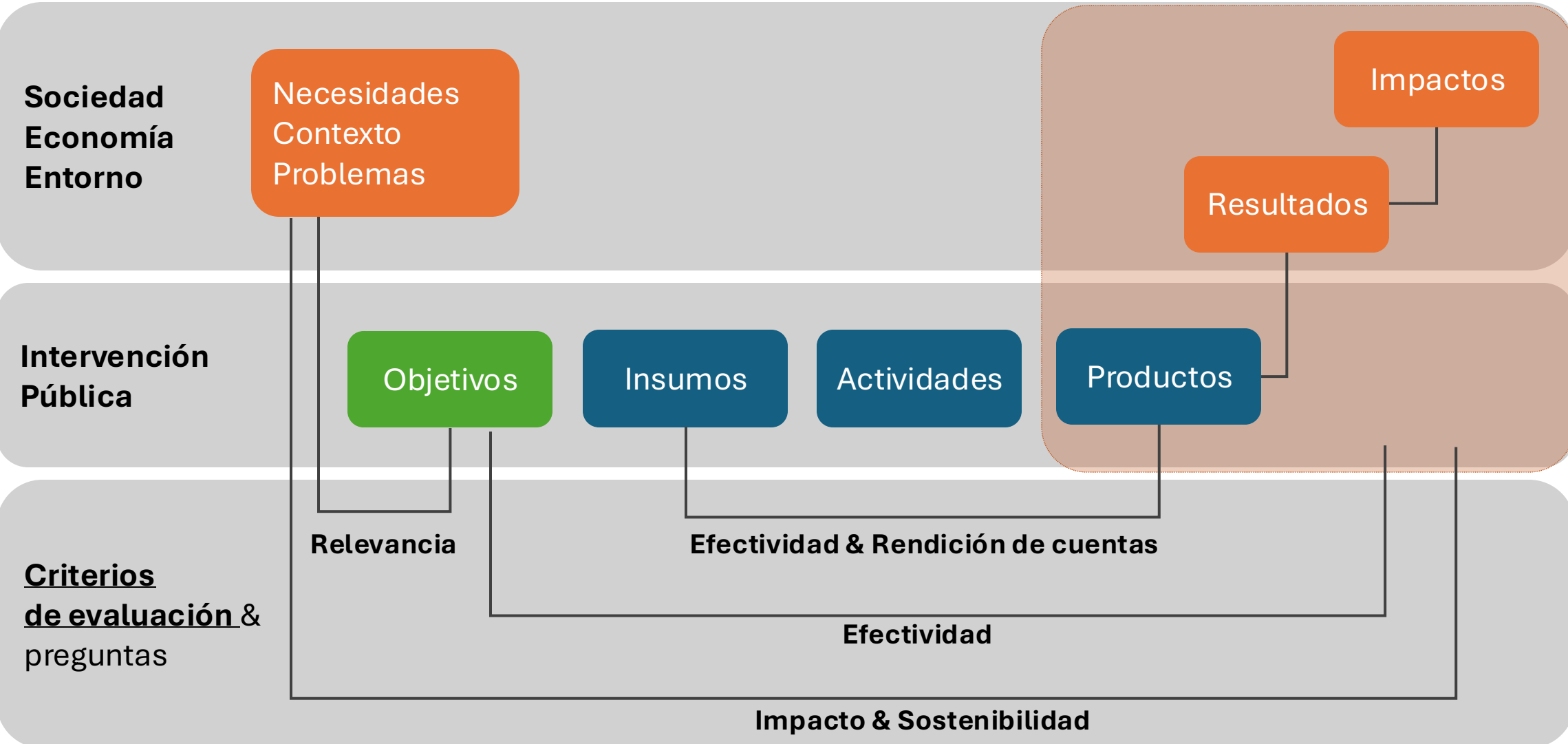
# Ciclo de las políticas públicas



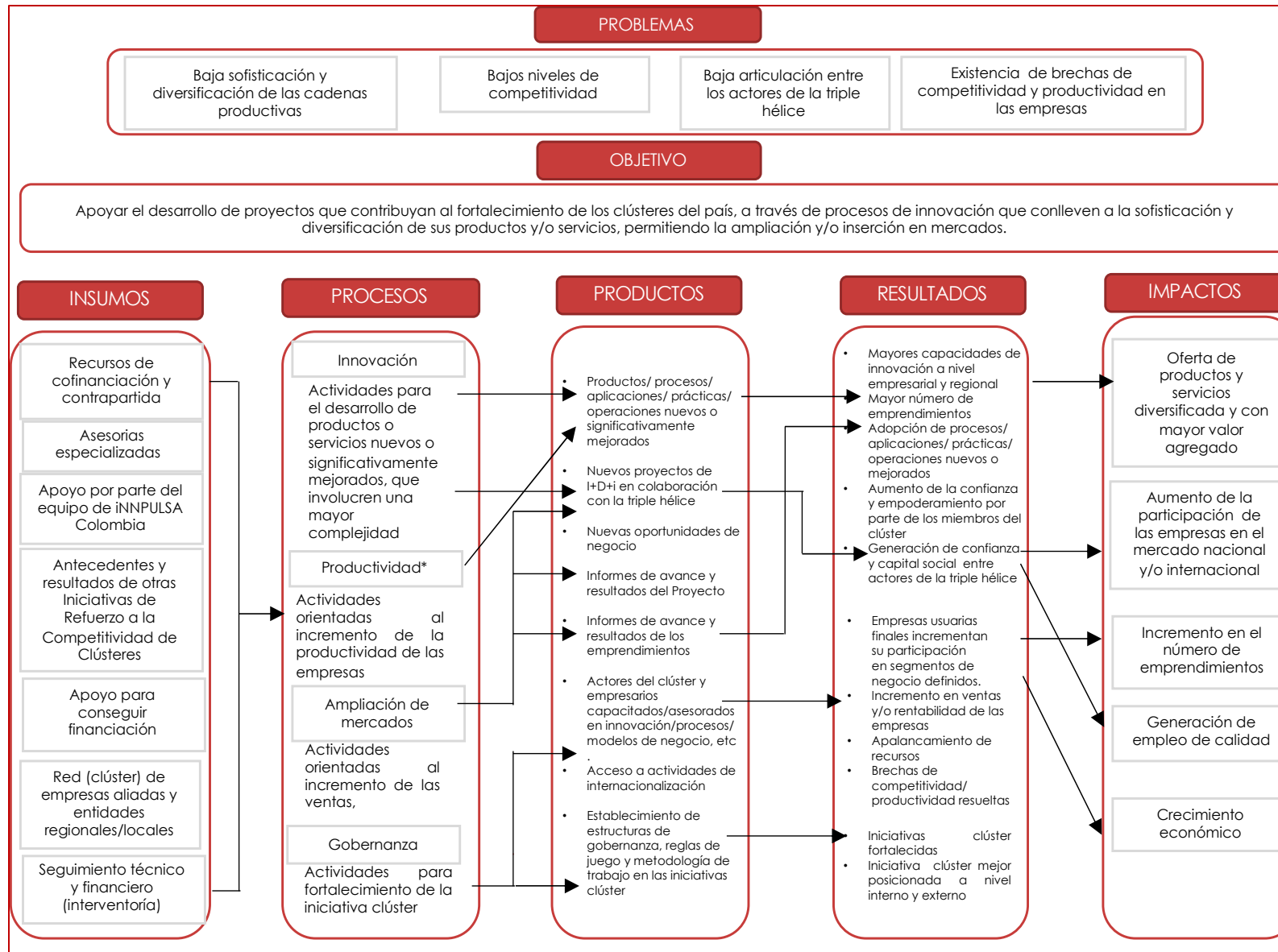
# Ciclo de las políticas públicas



# Lógica de intervención



# Ejemplo - Evaluación Clústeres



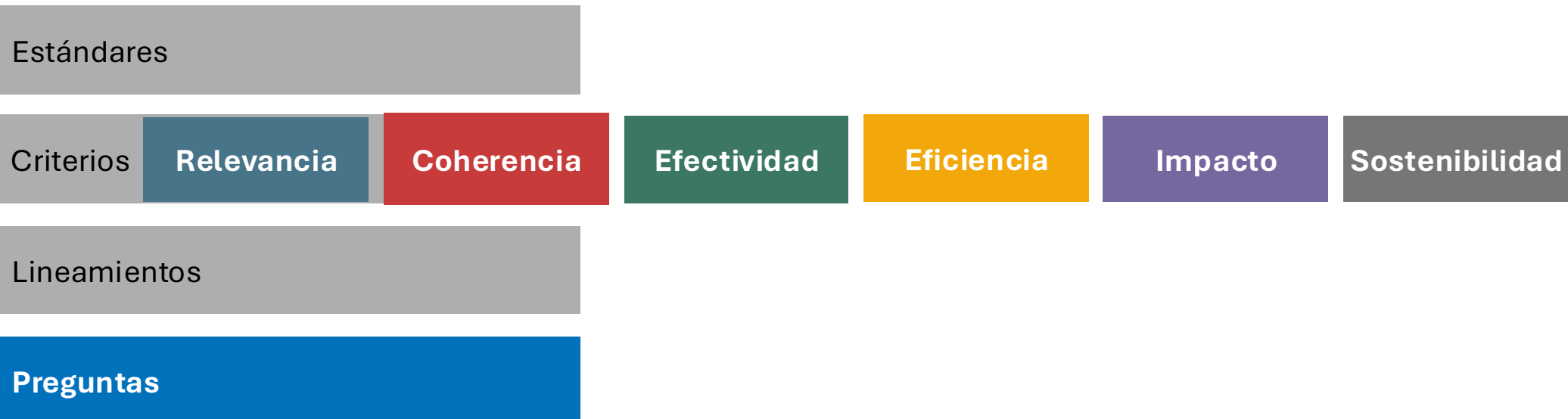


# Evaluación (de impacto)

- **Apreciación** sistemática y objetiva de un proyecto, programa o política, de su diseño, puesta en práctica y sus resultados; para determinar la pertinencia, el logro de objetivos, eficiencia, eficacia, impacto y sostenibilidad (OCDE, 2002).
- Uso de herramientas cuantitativas, con soporte en información cualitativa, para medir los **efectos** de programas e iniciativas, basados en un riguroso **análisis contrafactual** y un análisis de contexto.
- **Inferencia causal**: uso de la teoría, y de un profundo conocimiento de aspectos institucionales, para estimar el impacto de eventos y decisiones sobre un resultado de interés (Cunningham, 2021).

¿Qué habría pasado con los “beneficiarios” en ausencia del programa o la intervención?

**La correlación no implica causalidad**



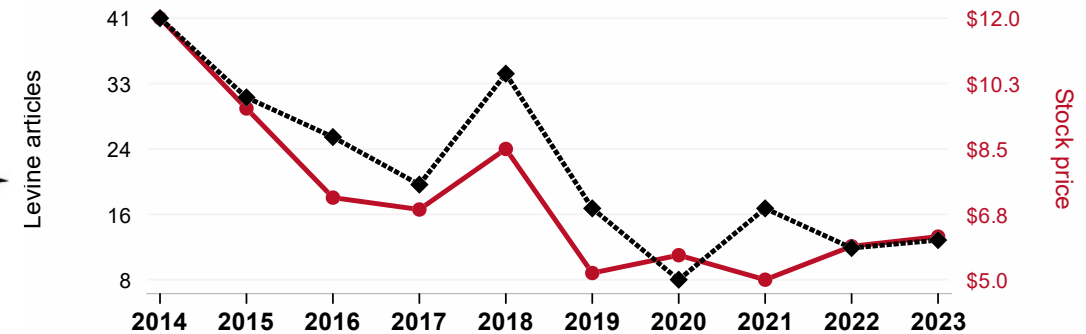
# Correlación vs Causalidad



## Bloomberg Money Stuff articles about insider trading

correlates with

### Banco Bilbao Vizcaya Argentaria's stock price (BBVA)



◆ Articles published by Matt Levine on Bloomberg.com with 'insider' in the headline or summary. · Source: Bloomberg  
● Opening price of Banco Bilbao Vizcaya Argentaria (BBVA) on the first trading day of the year · Source: LSEG Analytics (Refinitiv)

2014-2023,  $r=0.904$ ,  $r^2=0.817$ ,  $p<0.01$  · [tylervigen.com/spurious/correlation/5892](http://tylervigen.com/spurious/correlation/5892)

Figure 1.1: No correlation doesn't mean no causality. Artwork by Seth.

- **Causalidad:** se da cuando un fenómeno afecta directamente a otro. Es el camino entre la causa y el efecto.
- **Correlación:** significa que existe una relación mutua entre dos fenómenos. Sin embargo, esta puede presentarse en ausencia de causalidad.

Problema Fundamental de la Inferencia Causal.



Contrafactual (Supuestos)

(Holland, 1986)

# Ejemplo - Evaluación Clústeres

Criterios de evaluación	Preguntas orientadoras
Relevancia	<ul style="list-style-type: none"> <li><input type="checkbox"/> ¿En qué medida los objetivos y componentes de la convocatoria (innovación, ampliación de mercados, fortalecimiento de las iniciativas clúster) están alineados con las necesidades, los desafíos, problemas y oportunidades de la región?</li> <li><input type="checkbox"/> ¿En qué medida los objetivos de la convocatoria están alineados con las necesidades, los desafíos, problemas y oportunidades de las iniciativas clúster?</li> </ul>
Coherencia (externa)	<ul style="list-style-type: none"> <li><input type="checkbox"/> ¿En qué medida el apoyo de iNNovaClúster se articula con las estrategias a largo plazo de la iniciativa clúster?</li> <li><input type="checkbox"/> ¿En qué medida otros instrumentos de política contribuyen al desarrollo de las iniciativas clúster?</li> </ul>
Efectividad	<ul style="list-style-type: none"> <li><input type="checkbox"/> ¿En qué medida se alcanzaron los productos esperados, según tipos de productos?</li> <li><input type="checkbox"/> ¿En qué medida se lograron los resultados intermedios para las empresas y para la Iniciativas clúster?</li> <li><input type="checkbox"/> ¿En qué medida se lograron los resultados finales para las empresas y para la Iniciativas clúster?</li> </ul>
Diseño del instrumento	<ul style="list-style-type: none"> <li><input type="checkbox"/> ¿En qué medida el diseño de las convocatorias ha respondido adecuadamente a las necesidades de las empresas y de las iniciativas clúster? (tiempos, recursos de cofinanciación, etc.)</li> <li><input type="checkbox"/> ¿El diseño de las convocatorias ha permitido responder adecuadamente a las necesidades de las regiones en los que se están ejecutando los proyectos? (tiempos, recursos de cofinanciación, etc.)</li> </ul>
Eficiencia de los procesos administrativos	<ul style="list-style-type: none"> <li><input type="checkbox"/> ¿Se están usando los recursos (financieros, humanos, etc.) de manera eficiente?</li> </ul>
Sostenibilidad	<ul style="list-style-type: none"> <li><input type="checkbox"/> ¿En qué medida se prevé que, una vez terminado el proyecto, la iniciativa tendrá continuidad?</li> <li><input type="checkbox"/> ¿En qué medida los resultados de las intervenciones realizadas han generado impactos/beneficios que, se prevé, se mantendrán en el mediano o largo plazo?</li> </ul>

# Evaluación (de impacto)

¡Con las evaluaciones generamos recomendaciones basadas en los resultados y conclusiones que apuntan al mejoramiento de los programas evaluados!

## Seleccionamos la política a evaluar

1



Identificamos las intervenciones públicas prioritarias para ser evaluadas, teniendo en cuenta:

- Importancia estratégica
- Oportunidad
- Factibilidad
- Disponibilidad de información.

## Definimos qué y como vamos a evaluar

2



Teniendo en cuenta la intervención pública seleccionada. Para ello:

- Estudiamos la intervención
- Identificamos los actores
- Diseñamos la cadena de valor.

## Contratamos o desarrollamos

3



- En el caso de las evaluaciones externas, se realiza a través de un concurso público de méritos.
- En el caso de las evaluaciones internas, se desarrolla por el equipo de la Subdirección de Evaluación.

## El consultor recopila la información

4



- Se recopila la información necesaria para responder a las preguntas de la evaluación.
- La información se recopila a través de diferentes fuentes, como encuestas, entrevistas, observación o análisis de documentos.

## Socializamos el resultado de la evaluación

5



- Se generan recomendaciones de política a partir de los resultados de las evaluaciones.
- Se socializan de los productos de evaluación a los actores interesados.
- Se plantean acciones puntuales para impulsar el uso de los resultados de las evaluaciones.

### Transversal

#### Cómite técnico de seguimiento

Conformado por la DSEPP, la dirección técnica del DNP a cargo del sector y las entidades ejecutoras de la política o programa a evaluar.

# Evaluación (de impacto) - ¿Para qué?

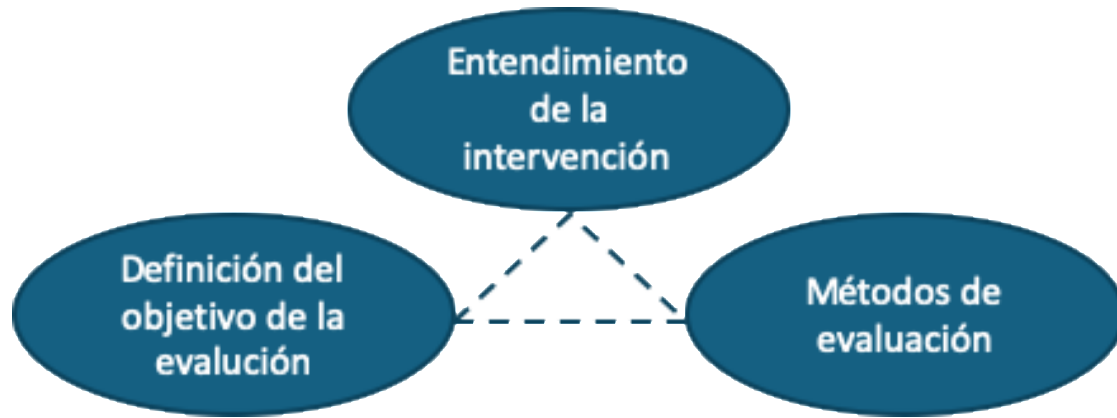
- La teoría por sí sola no puede guiar la política pública → Por definición, la teoría surge de la necesidad de limitar la complejidad del mundo para poder entenderlo.
- Se requiere evidencia de aquello que funciona.
- Evaluación de costo-beneficio.
- Mejorar la implementación de los programas en curso.
- Aprendizaje y gestión de conocimiento.

## Rendición de cuentas:

- Los soportes de la eficacia de los programas o intervenciones son necesarios para las **revisiones del gasto** y en respuesta al escrutinio y el desafío de los organismos públicos de rendición de cuentas.

# Evaluación (de impacto) - Etapas del proceso

## 1. Alcance de la evaluación.



- Métodos mixtos (costos y beneficios).
- Recursos limitados (priorización de respuestas de evaluación).
- Costos, tiempos, necesidades de información.

### “Teoría de cambio”

- Problemática que soluciona la intervención.
- Objetivos de la intervención.
- Cadena de causalidad de la intervención.
- Grupos poblacionales beneficiados.
- Condiciones necesarias para el éxito de la intervención.
- Supuestos y evidencia.

### “Cadena de valor”

## 2. Enfoque/preguntas de la evaluación.

## 4. Implementación de la evaluación.

## 3. Método de la evaluación – Estrategia de recolección de información.

## 5. Uso y difusión de hallazgos.

# Evaluación (de impacto) – Tipos de evaluación (etapa)

Antes



- **Apoyar el diseño de la intervención.**  
Revisión de evaluaciones existentes.

- ¿Cómo se espera que funcione la intervención?
- ¿Qué pruebas lo respaldan?
- ¿Por qué podría no funcionar?
- ¿Cuáles son los riesgos?

Durante



## **Influencia en decisiones de ajuste.**

- ¿La intervención se está ejecutando según lo previsto?
- ¿La intervención está funcionando de la misma forma en diferentes zonas o grupos poblacionales?
- ¿Se observan consecuencias imprevistas?
- ¿Qué indicios se tienen respecto a los resultados esperados?

Después



## **Conclusiones sobre el diseño, implementación y resultados.**

- ¿Funcionó la intervención?
- ¿Cuáles son las lecciones aprendidas?
- ¿Hubo algún impacto negativo o imprevisto?

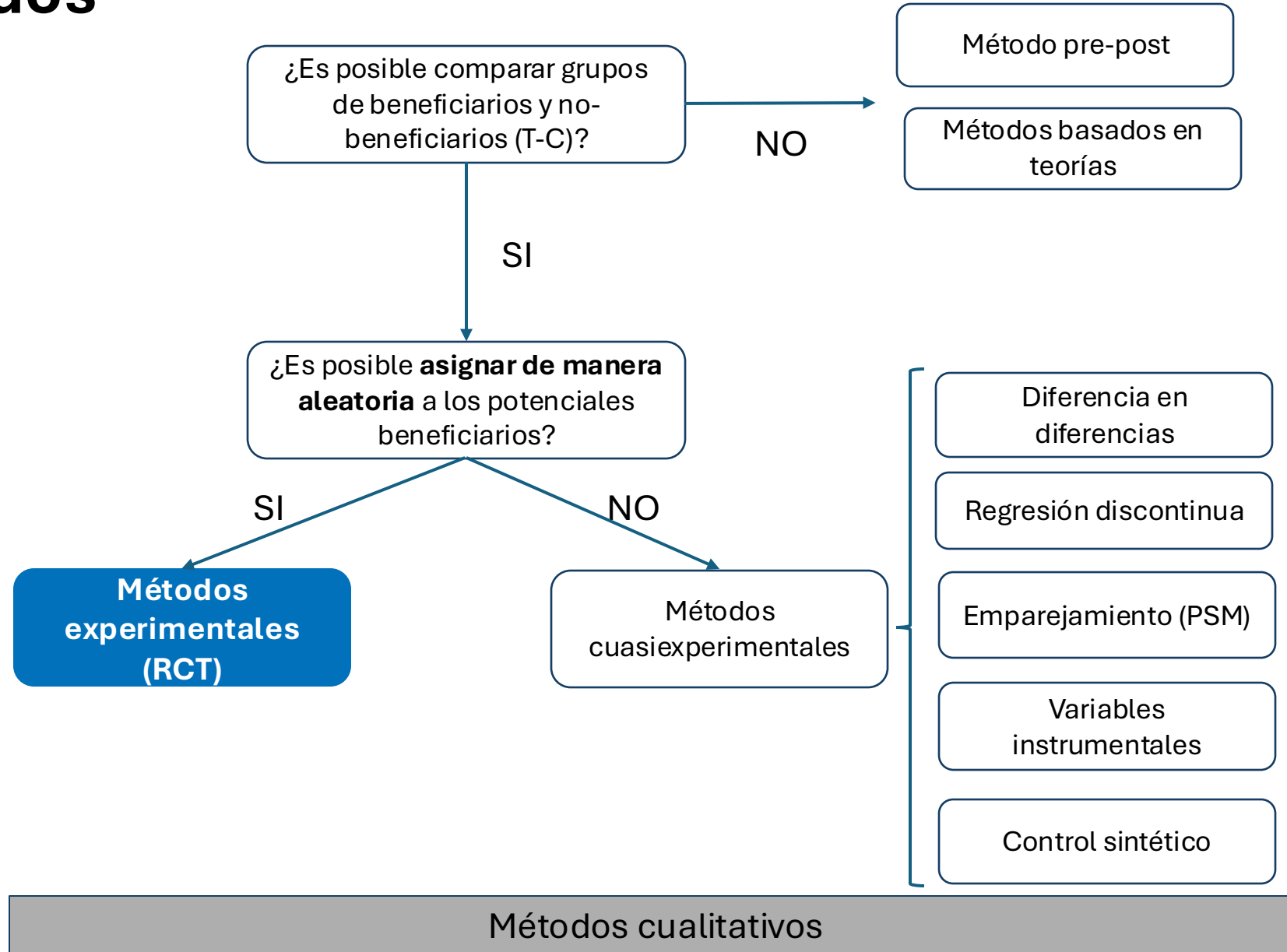


# Agenda Módulo 9: “Evaluaciones de impacto”

Conceptos básicos para el diseño e implementación de evaluaciones de impacto.

1. ¿Qué es la evaluación de impacto? ¿Qué es la inferencia causal?
2. **Diseños experimentales.**
3. Diferencia en Diferencias.
4. Emparejamiento.
5. Regresión discontinua.
6. ¿Qué técnica usar y cuándo?
7. ¿Evaluar o no evaluar?
8. ¿Cómo estructurar una evaluación de impacto desde el diseño?  
Retos y realidades.

# Métodos



# Experimentos sociales para luchar contra la pobreza



Colombia+  
Competitiva

Cooperación suiza para la competitividad



# El “Mundo Ideal” del evaluador de política

- ¿Cómo saber si un programa sirve?
  1. Clone a una persona.
  2. A uno de ellos asígnelo aleatoriamente a un tratamiento
  3. Al otro déjelo en el grupo de control.
- Obtenga dos muestras aleatorias de la misma población, y asigne a una muestra al tratamiento y a la otra al grupo de control.
  - Usted debe verificar si la muestra está balanceada en observables
  - En caso de que sí, usted puede argumentar por la **ley de los grandes números** que también están balanceados en no observables.

# Efectos causales

- Imagine que usted quiere saber cuál es el efecto que tienen los seguros de salud sobre las personas.
- ¿Por qué no puede comparar a las personas que toman un seguro *versus* los que no toman un seguro?
- Pensemos que el gobierno asigna este seguro de salud aleatoriamente.

# Efectos Causales

*(Mostly Harmless Econometrics Chapter 2)*

- El efecto del seguro sobre un individuo  $i$  está dado por:
  
- Problema en el mundo real:
  
- Los efectos causales pueden ser diferentes entre individuos por lo que nos interesan los efectos promedio:

# Efectos Causales

(*Mostly Harmless Econometrics Chapter 2*)

- El efecto del seguro sobre un individuo  $i$  está dado por:

$$\kappa_i = Y_{1i} - Y_{i0}$$

- $Y_{1i}$ : nivel de salud para  $i$  con seguro ( $D_i=1$ )
  - $Y_{0i}$ : nivel de salud para  $i$  **sin seguro** ( $D_i=0$ )
- 
- Problema en el mundo real:
    - Sólo puedo observar  $Y_{1i}$  o  $Y_{0i}$
  
  - Los efectos causales pueden ser diferentes entre individuos por lo que nos interesan los efectos promedio:
    - **Average treatment effect (ATE):**  $E[Y_{1i} - Y_{0i}]$
    - **Average treatment effect on the treated (ATT):**  $E[Y_{1i} - Y_{0i} | D_i = 1]$

# Sesgo de selección

- Una diferencia de medias entre tratados y controles nos puede llevar a conclusiones equivocadas:

$$E[Y_i | D_i = 1] - E[Y_i | D_i = 0]$$

**Tabla 1.2. Resultados y tratamientos para Khuzdar y María**

	Khuzdar Khalar	María Moreno
Resultado potencial sin seguro $Y_{0i}$	3	5
Resultado potencial con seguro $Y_{1i}$	4	5
Tratamiento (opción elegida en cuanto a seguro) $D_i$	1	0
Resultado real en cuanto a estado de salud $Y_i$	4	5
Efecto del tratamiento $Y_{1i} - Y_{0i}$	1	0

ATE

ATT

# Sesgo de selección

*(Mostly Harmless Econometrics Chapter 2)*

- Podemos escribir el resultado observado como

$$Y_i = Y_{0i} + (Y_{1i} - Y_{0i}) \cdot D_i$$

- ¿Qué podemos observar aquí?



Colombia+  
Competitiva

Cooperación suiza para la competitividad



# ¿De qué nos sirve la asignación aleatoria?

- Aún nos interesa la diferencia de medias observable:

$$E[Y_i | D_i = 1] - E[Y_i | D_i = 0]$$

- **PERO** la aleatorización **desaparece el sesgo de selección**

$$\{E[Y_{0i} | D_i = 1] - E[Y_{0i} | D_i = 0]\}$$

- Porque  $Y$  y  $D$  son independientes.
- Y así concluimos que:

# ¿De qué nos sirve la asignación aleatoria?

- Aún nos interesa la diferencia de medias observable:

$$E[Y_i | D_i = 1] - E[Y_i | D_i = 0]$$

- **PERO** la aleatorización **desaparece el sesgo de selección**

$$\{E[Y_{0i} | D_i = 1] - E[Y_{0i} | D_i = 0]\}$$

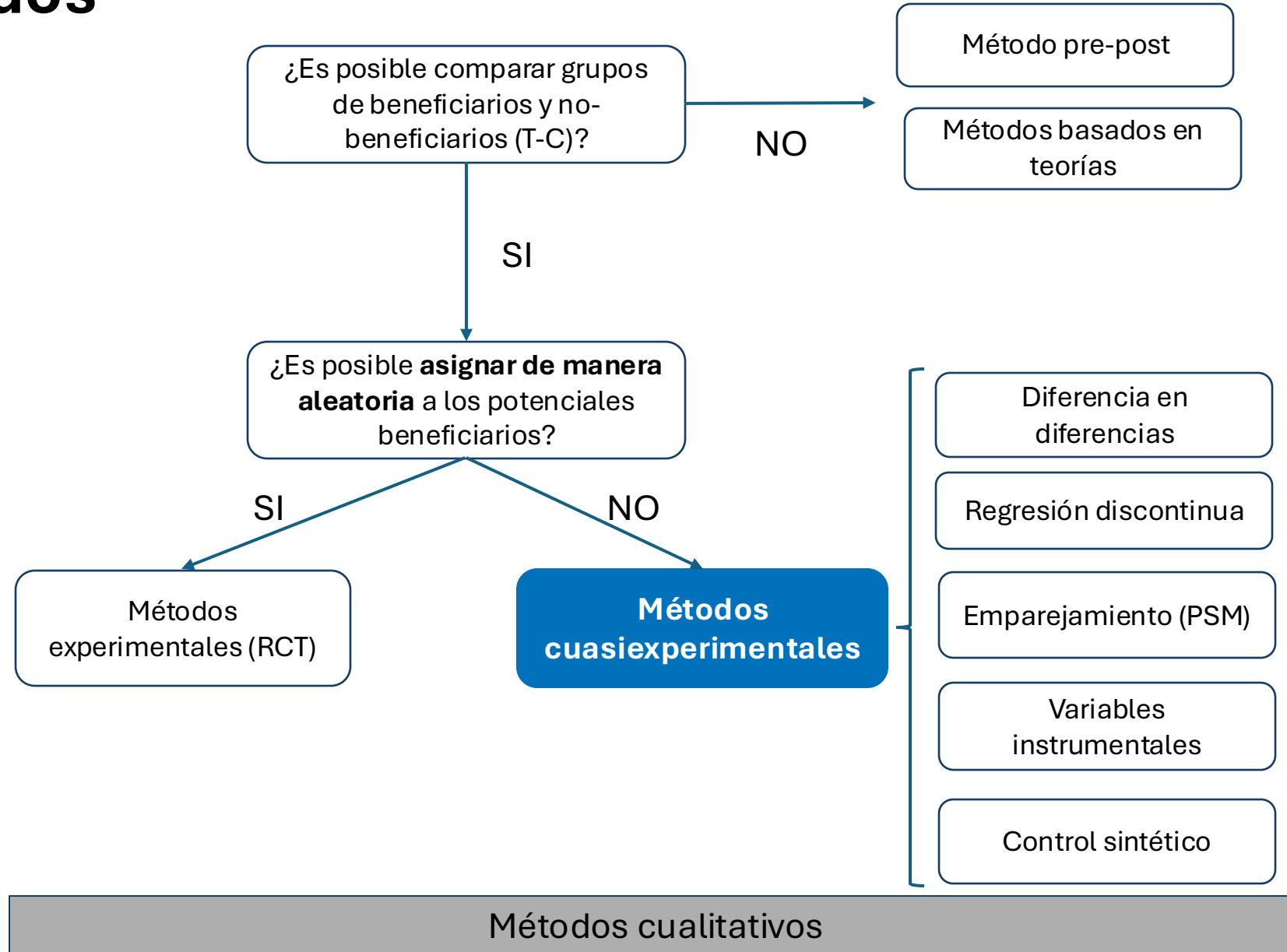
- Porque  $Y$  y  $D$  son independientes.
- Y así concluimos que:

$$\begin{aligned} E[Y_i | D_i = 1] - E[Y_i | D_i = 0] &= E[Y_{1i} - Y_{0i} | D_i = 1] \\ &= E[Y_{1i} - Y_{0i}] \end{aligned}$$

# Desventaja de los métodos experimentales

- Hacer una buena evaluación con base en un experimento social controlado es costoso monetaria y políticamente.
  - Los métodos experimentales pueden ser más costosos que los no experimentales.
- ¿Injusticia fundamental?
  - Por pertenecer al grupo de control se está excluyendo a un segmento de la población, igualmente vulnerable, de los beneficios de la intervención.
- Efectos de largo plazo versus efectos de corto plazo.
- Externalidades o efectos de equilibrio general.
- Solamente se puede evaluar el programa **tal** como está diseñado.
- Posibilidad de generalizar los resultados.
- Imposibilidad para diseñar un experimento.

# Métodos





# Agenda Módulo 9: “Evaluaciones de impacto”

Conceptos básicos para el diseño e implementación de evaluaciones de impacto.

1. ¿Qué es la evaluación de impacto? ¿Qué es la inferencia causal?
2. Diseños experimentales.
3. **Diferencia en Diferencias.**
4. Emparejamiento.
5. Regresión discontinua.
6. ¿Qué técnica usar y cuándo?
7. ¿Evaluar o no evaluar?
8. ¿Cómo estructurar una evaluación de impacto desde el diseño?  
Retos y realidades.

# Contrafactuales

- El objetivo de cualquier evaluación de impacto de un programa es construir un **contrafactual** válido.
- Por esto los experimentos son tan útiles: *si el tratamiento se asigna aleatoriamente, las observaciones en promedio **solo difieren precisamente en ser tratado o no.***
- El efecto de un tratamiento en la variable  $Y_i$  para el individuo  $i$  es:
$$\kappa_i = Y_{1i} - Y_{i0}$$
  - $Y_{1i}$ : resultado si la persona es tratada ( $D_i=1$ ).
  - $Y_{0i}$ : resultado si la persona no es tratada ( $D_i=0$ ).
- Para un individuo observamos  $Y_{1i}$  o  $Y_{0i}$ , pero nunca ambas.
- Estamos interesados en:
  - **Average treatment effect (ATE):**  $E[Y_{1i} - Y_{0i}]$
  - **Average treatment effect on the treated (ATT):**  $E[Y_{1i} - Y_{0i} | D_i = 1]$



# Diferencia en Diferencias

- Recupera una **variación** de la variable independiente que es “**casi aleatoria**” al comparar grupos de **tratamiento** y **control** en el tiempo.
- En 1855, Jhon Snow estudió las epidemias de cólera en Londres.
  - Snow quería establecer que la cólera se transmitía por el consumo de agua potable contaminada (en lugar de por inhalar “mal aire”).



# Diferencia en Diferencias

- Recupera una **variación** de la variable independiente que es “**casi aleatoria**” al comparar grupos de **tratamiento** y **control** en el tiempo.
- En 1855, Jhon Snow estudió las epidemias de cólera en Londres.
  - Snow quería establecer que la cólera se transmitía por el consumo de agua potable contaminada (en lugar de por inhalar “mal aire”).

Snow comparó los cambios en las tasas de mortalidad por cólera en distritos atendidos por dos compañías de agua. En 1849, Lambeth water company movió sus tuberías de toma río arriba, más arriba en el Támesis, por encima del principal punto de descarga de aguas residuales.

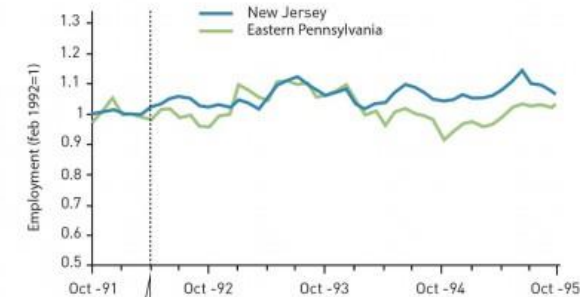
# Card and Krueger, 1994 - ¿El salario mínimo afecta el empleo?

- Experimento “natural”.
- ¿Por qué no comparar pasado y presente para una sola unidad?
  - Si no tenemos un experimento puede haber muchos eventos que afecten a la unidad entre los dos momentos que se observa.
    - E.g.: la ciudad se destruye, pero al mismo tiempo el país sufre un choque presupuestal por el precio del petróleo
- ¿Por qué no comparar tratamiento respecto al control?
  - Es difícil encontrar buenos controles si no son inducidos experimentalmente.
  - Podría ocurrir que la unidad de tratamiento y control fuesen diferentes antes del tratamiento.
    - E.g.: las ciudades comparadas pueden diferir en distancia a mercados, calidad fiscal, entre otros.

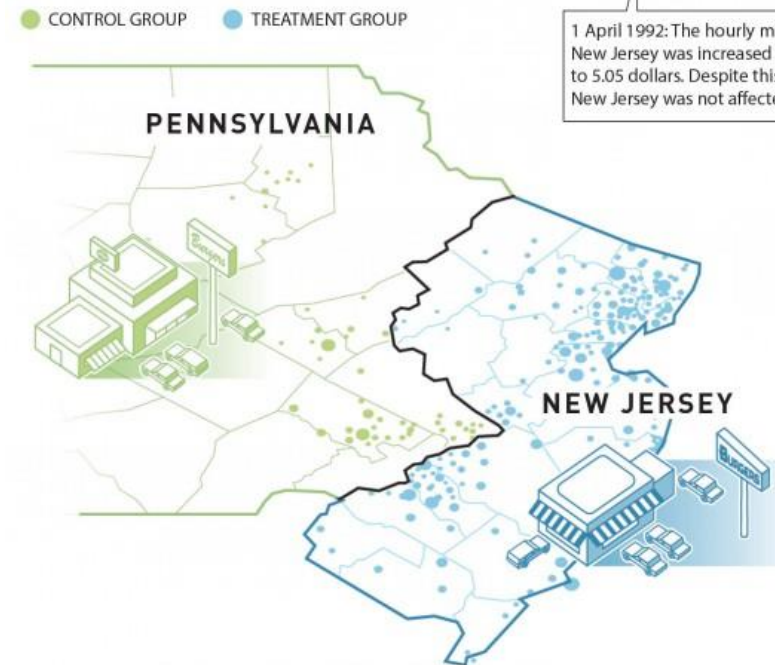
## The effect of increasing the minimum wage

Card and Krueger used a natural experiment to study how increasing the minimum wage affects employment.

The researchers identified a treatment group (restaurants in New Jersey) and a control group (restaurants in eastern Pennsylvania) to measure the effect of increasing the minimum wage.



1 April 1992: The hourly minimum wage in New Jersey was increased from 4.25 dollars to 5.05 dollars. Despite this, employment in New Jersey was not affected.



# Diferencia en Diferencias (DiD)

- La idea básica tras DiD es mezclar dos dimensiones para estimar un efecto causal (al menos el **ATE**)
  - Primera Dimensión: Tiempo.
  - Segunda Dimensión: (usualmente) Ubicación.
- En esencia:
  - Hay características que no cambian en el tiempo (e.g., la distancia a otras ciudades), y hay cosas que no cambian entre ubicaciones geográficas (e.g., choques macroeconómicos).
  - Si lo único que cambia en el tiempo y entre ubicaciones es el **tratamiento**, entonces la doble diferencia elimina las demás potenciales variables explicativas o *confounders*.

# Diferencia en Diferencias (DiD)

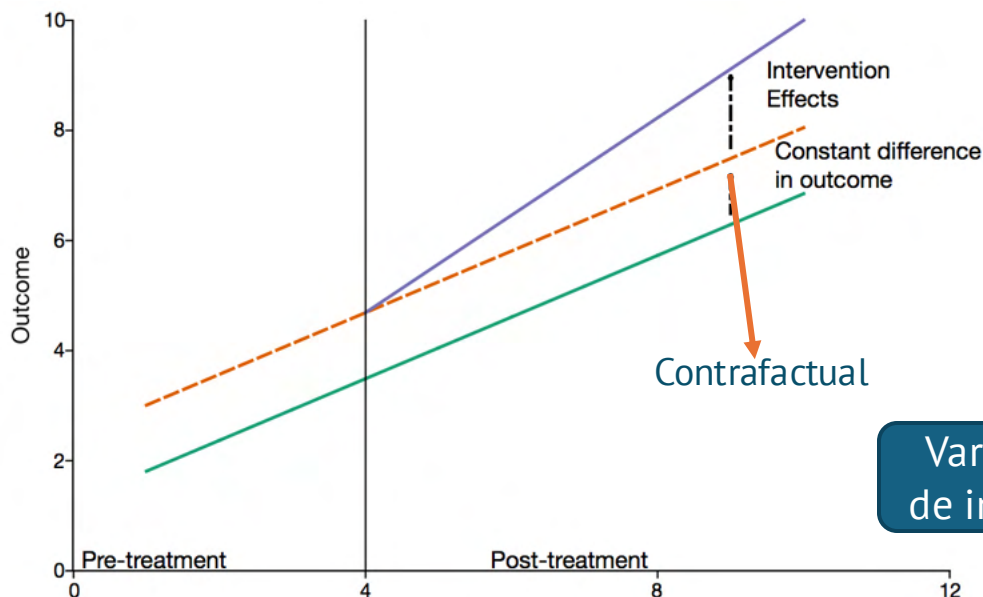
TABLE 5.2.1

Average employment in fast food restaurants before and after the New Jersey minimum wage increase

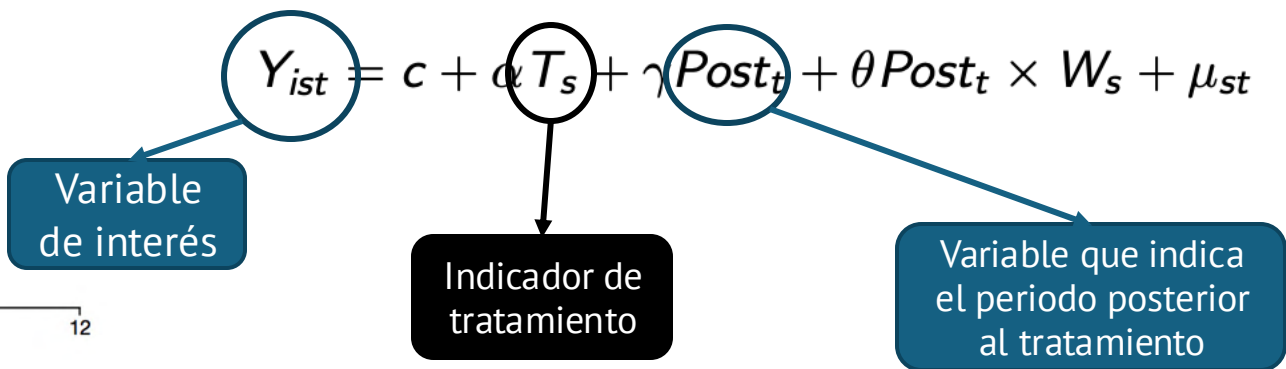
Variable	PA (i)	NJ (ii)	Difference, NJ – PA (iii)
1. FTE employment before, all available observations	23.33 (1.35)	20.44 (.51)	-2.89 (1.44)
2. FTE employment after, all available observations	21.17 (.94)	21.03 (.52)	-.14 (1.07)
3. Change in mean FTE employment	-2.16 (1.25)	.59 (.54)	2.76 (1.36)

*Notes:* Adapted from Card and Krueger (1994), table 3. The table reports average full-time-equivalent (FTE) employment at restaurants in Pennsylvania and New Jersey before and after a minimum wage increase in New Jersey. The sample consists of all restaurants with data on employment. Employment at six closed restaurants is set to zero. Employment at four temporarily closed restaurants is treated as missing. Standard errors are reported in parentheses.

# Diferencia en Diferencias (DiD)



- **Supuesto de identificación clave:** las tendencias habrían sido las mismas en ambos grupos en ausencia del tratamiento.



## Ejemplo

$$Ventas_{it} = c + \alpha T_i + \gamma Post_t + \theta T_i Post_t + \beta_1 Tama\tilde{n}o_{it} + \beta_2 Edad_{it} + \beta_3 Edad_{it}^2 + \beta_4 Protec_{it} + \beta_5 CH_{it} + \dots + \epsilon_{it}$$

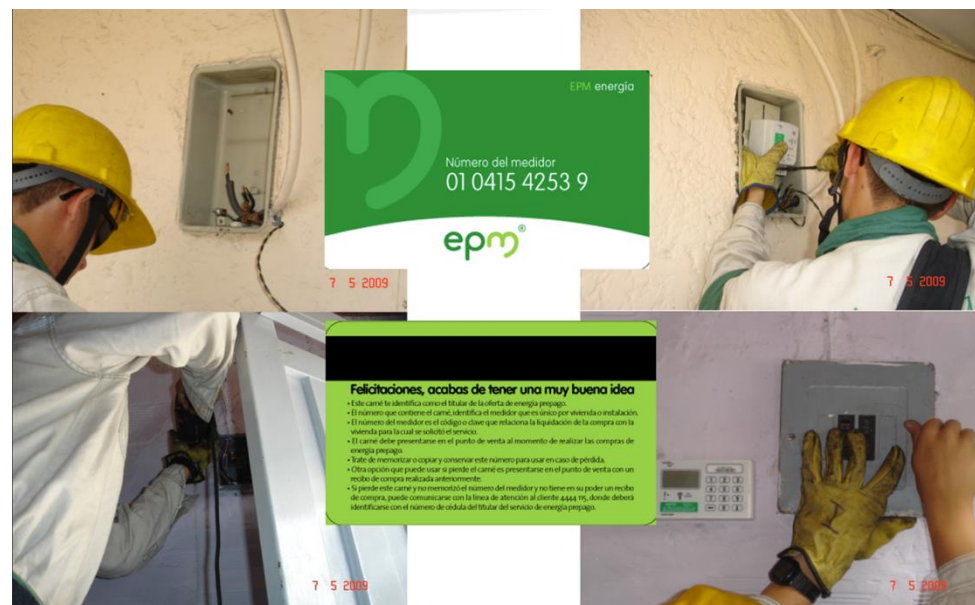
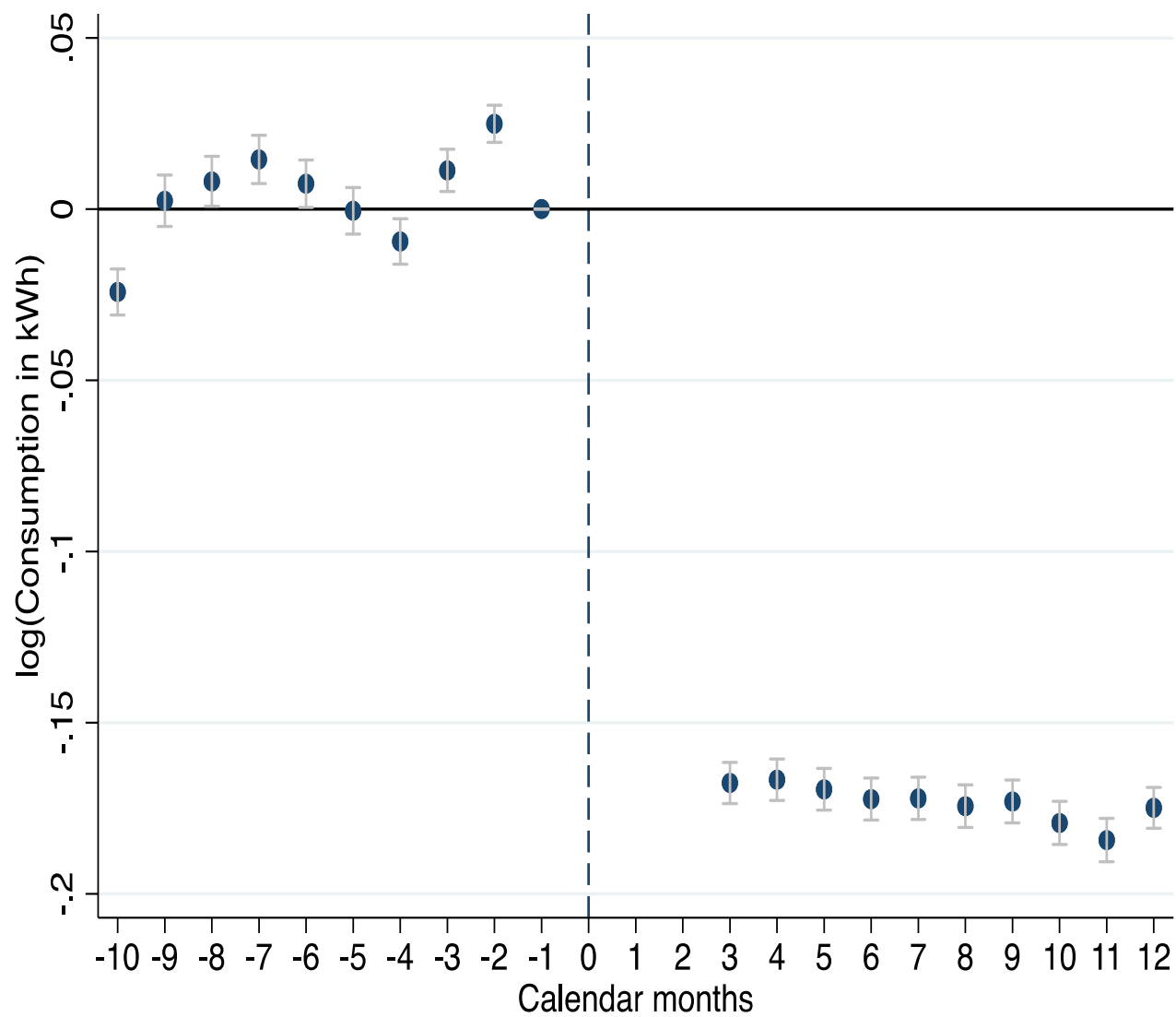
# Diferencia en Diferencias (DiD) - ¿Qué se hace en la práctica?

- Definir y obtener un **buen** grupo de control para el grupo de tratamiento.
- Revisar el supuesto de tendencias paralelas:
  - Con los datos (periodos antes no había diferencias entre los grupos).
  - Según la teoría económica.
- Estimar la ecuación:

$$Y_{it} = \gamma_0 + \gamma_1 \cdot T_i + \gamma_2 \cdot A_t + \delta \cdot T_i \cdot A_t + u_{it}$$

- Interpretar  $\delta$ .
- Realizar pruebas de robustez: otras formas de estimar la ecuación (e.g., más variables de control, excluir algunos de los controles que no sean muy buenos, hacer placebos).

## Effect of switching to the prepaid scheme on user's electricity consumption



**Table 2**  
The impact of prepaid schemes in electricity consumption by stratum.

Variable	log(Consumption in kWh)					
	Stratum 1		Stratum 2		Stratum 3	
	OLS 1	IV 2	OLS 3	IV 4	OLS 5	IV 6
PostPaid	-0.104*** (0.0028699)	-0.0876*** (0.0050623)	-0.199*** (0.0029715)	-0.132*** (0.0145609)	-0.292*** (0.00739)	-0.317 (0.244)
log(p <sup>-</sup> )	0.0213 (0.0010264)	-0.124*** (0.0290956)	0.0023 (0.0017865)	-0.3356*** (0.0801713)	-0.0575*** (0.0202)	0.0685 (1.073)
Observations	4,722,473	4,330,692	4,265,994	3,750,577	939,261	828,371
Clusters	72,754	66,227	56,993	50,267	12,348	10,928
R-squared	0.4545	-	0.4447	-	0.480	-
Adj. R-squared	0.4460	-	0.4372	-	0.473	-
Dwelling FE	✓	✓	✓	✓	✓	✓
Month FE	✓	✓	✓	✓	✓	✓

Notes: Clustered standard errors by dwelling in parentheses. \*\*\* p < 0.01; \*\* p < 0.05; \* p < 0.1. Estimation method: Difference-in-Differences.

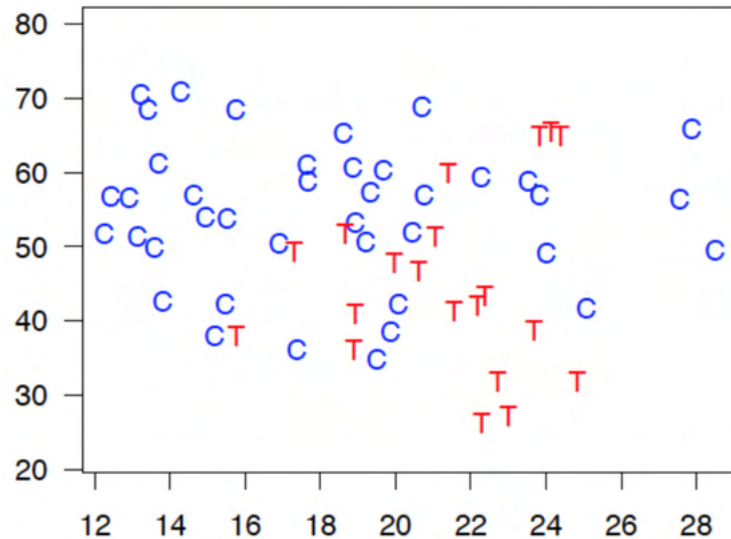


# Agenda Módulo 9: “Evaluaciones de impacto”

Conceptos básicos para el diseño e implementación de evaluaciones de impacto.

1. ¿Qué es la evaluación de impacto? ¿Qué es la inferencia causal?
2. Diseños experimentales.
3. Diferencia en Diferencias.
4. **Emparejamiento.**
5. Regresión discontinua.
6. ¿Qué técnica usar y cuándo?
7. ¿Evaluar o no evaluar?
8. ¿Cómo estructurar una evaluación de impacto desde el diseño?  
Retos y realidades.

# Emparejamiento



- Coarsened exact matching
- Propensity score matching
- Synthetic control

- Eliminar las observaciones que requerirían una extrapolación sustancial y producirían el desbalance, con el fin de:
  - Reducir la dependencia al modelo especificado.
  - Evitar la discreción del investigador.
  - Evitar posibles sesgos.

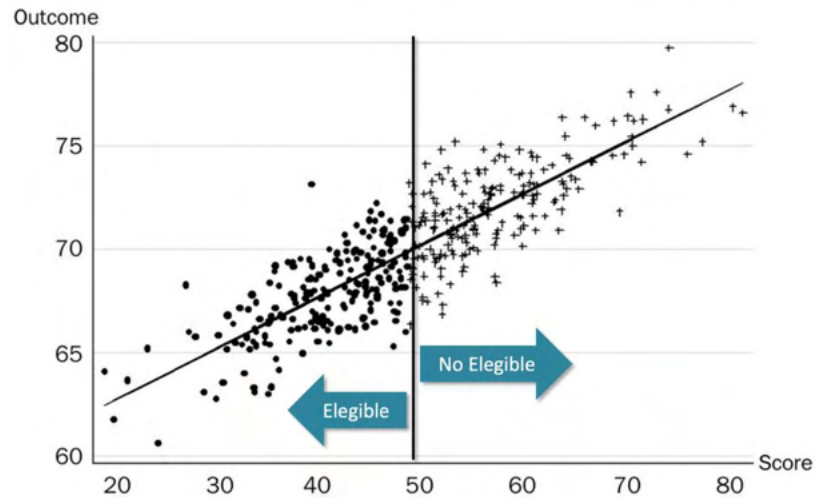


# Agenda Módulo 9: “Evaluaciones de impacto”

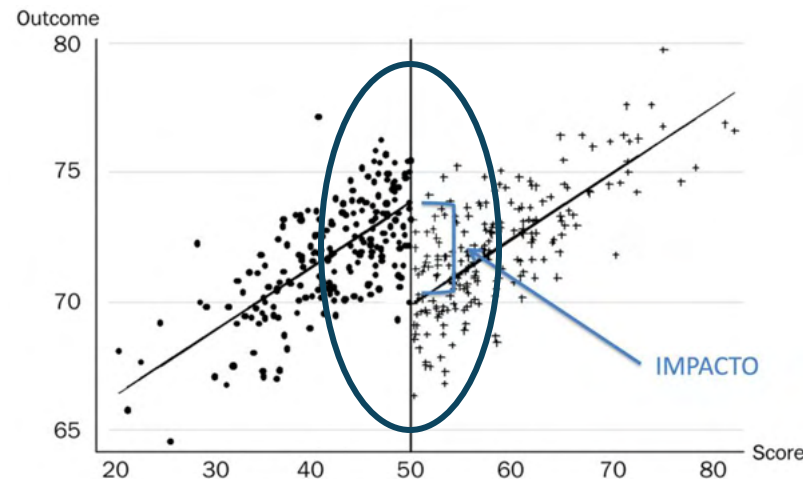
Conceptos básicos para el diseño e implementación de evaluaciones de impacto.

1. ¿Qué es la evaluación de impacto? ¿Qué es la inferencia causal?
2. Diseños experimentales.
3. Diferencia en Diferencias.
4. Emparejamiento.
5. **Regresión discontinua.**
6. ¿Qué técnica usar y cuándo?
7. ¿Evaluar o no evaluar?
8. ¿Cómo estructurar una evaluación de impacto desde el diseño?  
Retos y realidades.

# Regresión Discontinua



Línea base



Post intervención

- Aprovecha los **saltos** en la probabilidad de obtener el tratamiento sobre cierta variable de asignación.
- Es equivalente a tener un experimento justo alrededor del salto o discontinuidad.
- La idea es hacer una diferencia de medias justo alrededor de dicha discontinuidad.
- Cuando las unidades tratadas se pueden ordenar en una dimensión cuantificable (un índice es computable).
- Este índice tiene un **corte de elegibilidad**:
  - Salto en la probabilidad de ser tratado
- Los individuos justo debajo del corte son marginalmente más probables de ser tratados.

# Regresión Discontinua

- Programas sociales basados en puntajes universales:
  - Resultados de pruebas SABER 11 asignan cupos en universidades (o becas).
  - Puntajes en el SISBEN para acceder a beneficios sociales.
  - Edades mínimas o máximas para pensionarse (o ser beneficiario).
- Y otras discontinuidades “menos obvias”:
  - Reglas de diagnóstico aplicadas por médicos (ex. Reglas de hipertensión u obesidad).
  - Cuando un candidato gana una elección muy apretada.
  - “Maimonides rule:” Número máximo de estudiantes que pueden registrarse en un único salón.

# Tipos de Regresión Discontinua

- Discontinuidad nítida.
  - La probabilidad de ser tratado **salta** de 0 a 1.
  - Los resultados vienen de una diferencia de medias por encima y por debajo del punto de corte.
  - Ejemplo: el porcentaje de impuestos de la declaración de renta obedece a una regla contable muy clara **Y** la dirección de impuestos la fiscaliza de forma perfecta.
- Discontinuidad difusa.
  - La probabilidad de ser tratado **no salta** de 0 a 1
  - El salto, que no va de 0 a 1, se utiliza como una variable instrumental.
  - Ejemplo: el porcentaje de impuestos de la declaración de renta obedece a una regla contable muy clara **PERO** el gobierno no tiene la capacidad de fiscalizarla perfectamente, por lo que alguna gente paga menos de lo que debería.

# Regresión Discontinua – Supuestos adicionales



- Por supuesto, validez del instrumento (que la variable de puntaje salte).
- Los individuos no pueden manipular la variable de puntaje.
- No hay saltos en el punto de corte para otras variables (que no sean de resultado).

# Ventajas -Regresión Discontinua

- Es fácil de implementar
  - RD requiere datos en forma de corte transversal.
  - IV necesita encontrar un instrumento, y encontrar los datos.
  - En DiD se necesita un panel para poder verificar qué ocurría antes de la intervención
- La causalidad es sencilla de defender
  - Fuerte evidencia de exogeneidad por el salto.
  - Usualmente es suficiente tener un “gráfico” que muestre la discontinuidad, y su efecto en la variable de resultado.
  - La restricción de exclusión en IV no se puede probar.
  - Las tendencias paralelas en DiD requieren muchos más datos, y sus efectos usualmente no pueden ser probados.

# Desventajas - Regresión Discontinua

- Se necesita una base de datos GRANDE.
  - Se necesitan muchas observaciones cerca al punto de corte para tener suficiente poder estadístico y detectar una diferencia.
- Es un LATE (Local Average Treatment Effect)
  - Los resultados de una RD aplican para observaciones cerca al punto de corte.
  - Si la regresión discontinua es difusa, la interpretación tiene el mismo problema que la estimación de variables instrumentales.
- El LATE puede ser problemático al interpretar los resultados:
  - Asignación basada en un índice de pobreza: los resultados solo son válidos para aquellos muy cerca a ser considerados pobres.
  - Pruebas de desempeño: resultados válidos para estudiantes muy cercanos a ganarse la beca, pero no a los estudiantes de menor rendimiento.



# Agenda Módulo 9: “Evaluaciones de impacto”

Conceptos básicos para el diseño e implementación de evaluaciones de impacto.

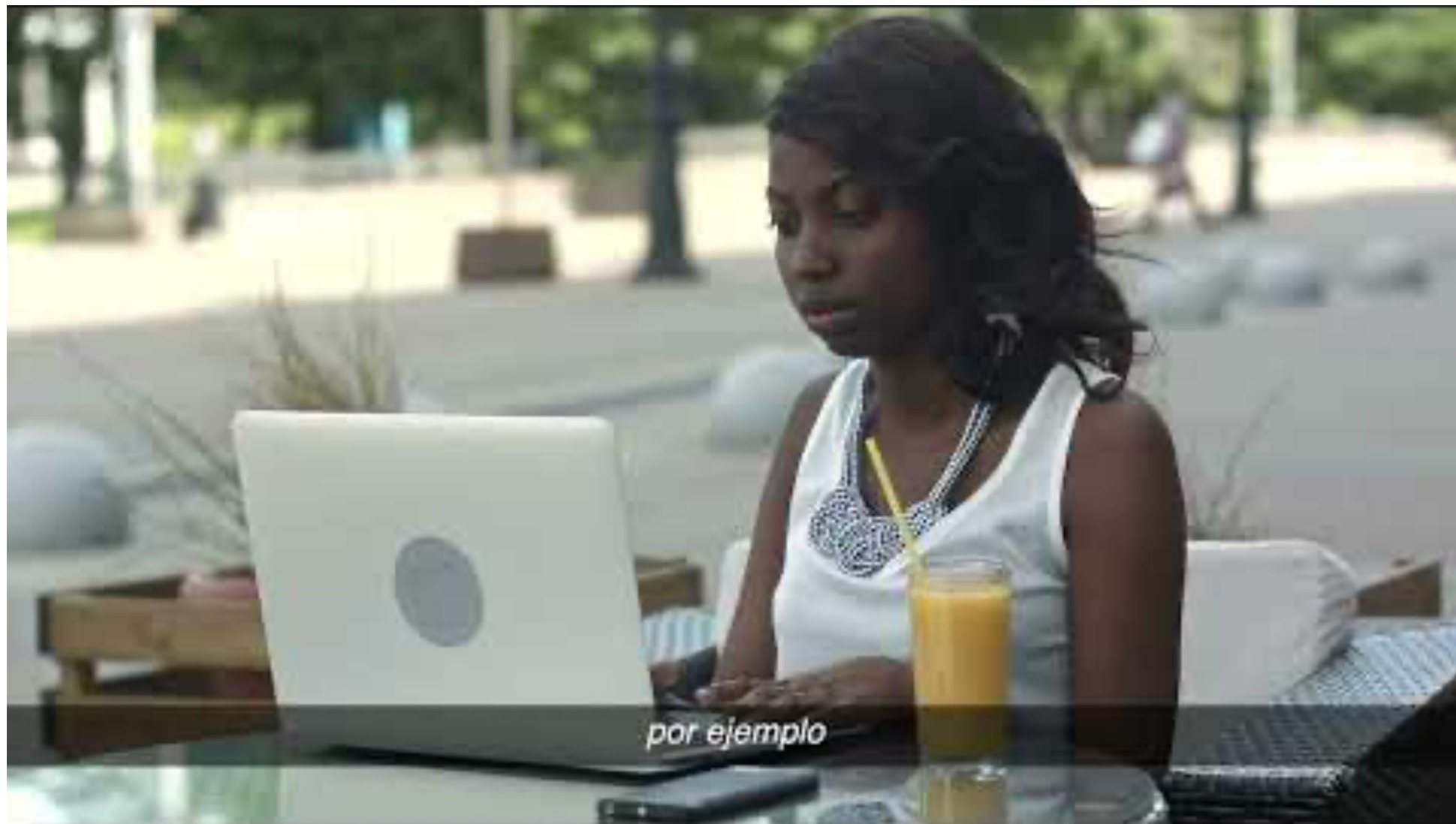
1. ¿Qué es la evaluación de impacto? ¿Qué es la inferencia causal?
2. Diseños experimentales.
3. Diferencia en Diferencias.
4. Emparejamiento.
5. Regresión discontinua.
6. **¿Qué técnica usar y cuándo?**
7. **¿Evaluar o no evaluar?**
8. ¿Cómo estructurar una evaluación de impacto desde el diseño?  
Retos y realidades.

# Programa Fábricas de Productividad



Colombia+  
Competitiva

Cooperación suiza para la competitividad



## PROGRAMA

# FÁBRICAS DE PRODUCTIVIDAD

## OBJETIVO

Mejorar la **productividad**, competitividad, rentabilidad y capacidades de las empresas a través de **asistencia técnica especializada**.



“DANE: Departamento Administrativo Nacional de Estadística; EAC: Encuesta Anual de Comercio; EAM: Encuesta Anual Manufacturera; EAS: Encuesta Anual de Servicios.”



DEPARTAMENTO NACIONAL DE PLANEACIÓN

## EVALUACIÓN DE IMPACTO PROGRAMA FÁBRICAS DE PRODUCTIVIDAD

**Objetivo:** Evaluar el impacto del Programa en las empresas beneficiarias y formular recomendaciones para mejorar su impacto.

### MÉTODOS MIXTOS DE EVALUACIÓN

#### Cuantitativo

- Análisis estadísticos descriptivos del Programa y metodología de diferencia en diferencias.
- Procesamiento de bases de datos del DANE, (EAM, EAC y EAS).
- Impacto en las empresas según región, tamaño y sector.

#### Cualitativo

##### Fuentes primarias

- 77 entrevistas con cámaras de comercio, Colombia Productiva, empresas, gestores, extensionistas y entidades territoriales.

##### Fuentes secundarias

- Análisis de documentos y normativas.
- Enfoque etnometodológico e interaccionismo simbólico



# RESULTADOS DE LA EVALUACIÓN



DEPARTAMENTO  
NACIONAL DE PLANEACIÓN

Gran parte de las **empresas entrevistadas** consideran que **cumplieron los objetivos** y se obtuvieron mejoras e impactos en su **productividad y rentabilidad**.

Esto debido al **conocimiento y prácticas adquiridas con el PFDP**, como la estandarización y optimización de procesos, menores costos y tiempos de producción, entre otros.

Solo se observan **diferencias significativas** en los impactos diferenciados por **línea de servicio** o región en el **sector turismo**

En las **empresas pequeñas** se observan **efectos significativos** en los **indicadores de impacto, productividades** medias, marginales, laboral, de capital y PTF en el corto plazo.

Las **empresas pequeñas** consideran que cuentan con **menor disponibilidad de tiempo y personal** para adquirir e implementar todos los conocimientos y prácticas que pueden aportar extensionistas.



# RESULTADOS DE LA EVALUACIÓN



DEPARTAMENTO NACIONAL DE PLANEACIÓN

Las **líneas de servicio con mayor percepción positiva** sobre el impacto en el componente cualitativo fueron: Productividad Operacional, Transformación Digital, Productividad Laboral y Gestión de Calidad.

Se evidencia una **adecuada articulación entre los actores internos** del PFDP. Se identificaron **casos exitosos de articulación con actores externos del sector privado**, algunas **entidades territoriales**, entre otros.

En el corto plazo **no se observa un efecto significativo** en el balance de los **indicadores de impacto**, ni en las **productividades medias, marginales o la PTF en empresas medianas y grandes**. Es posible que se puedan observar impactos a mediano y largo plazo

Los actores entrevistados perciben que el **PFDP impacta positivamente el entorno de negocios intra regional**. Además, se identificaron iniciativas de **articulación entre empresas beneficiarias a nivel intraregional**

Esto puede deberse a factores como el **poco tiempo entre la intervención y las bases de datos**, efectos asociados al **COVID 19** o la **dificultad metodológica** para calcular la **multiplicidad de factores** que componen la productividad de las empresas grandes.

# RECOMENDACIONES DE LA EVALUACIÓN



DEPARTAMENTO NACIONAL DE PLANEACIÓN

**01** Fomentar la participación de empresas repitentes.

**02** Concentrar las intervenciones de PFDP en las empresas pequeñas.

**03**

Usar el mapeo de instrumentos de la metodología ArCo para identificar y coordinar los programas de extensionismo.

**04**

Continuar documentando y divulgando buenas prácticas del PFDP.

**05**

Precisar/Definir mejor el rol que se espera de las entidades territoriales en PFDP.

**06**

Incrementar la participación de las empresas en los espacios de articulación existentes (ben.-ben. y no ben.-ben.).

**07**

Priorizar a nivel regional las intervenciones por sector, de acuerdo con los sectores definidos por las Comisiones Regionales de Competitividad e Innovación.

**08**

Ofrecer un extensionismo tecnológico más sofisticado para empresas grandes y medianas.

**09**

Evaluar la posibilidad de implementar intervenciones grupales.

**10**

Ampliar el tiempo entre la medición de línea de base y el seguimiento para realizar una evaluación de impacto de medio término.

**11**

Fomentar la información con que se cuenta para la medición de las productividades en articulación con otras entidades interesadas.

**12**

Diseñar rutas de atención diferenciadas según el tamaño de las empresas.

# ¿Qué técnica usar y cuándo?

- Eclecticismo metodológico.
- Conocimiento profundo de los mecanismos y variables que determinan la elegibilidad y participación en un programa.
- Circunstancias prácticas y políticas.
- Criterio de evaluador.
- Complementariedad entre evaluadores y ejecutores de programas.

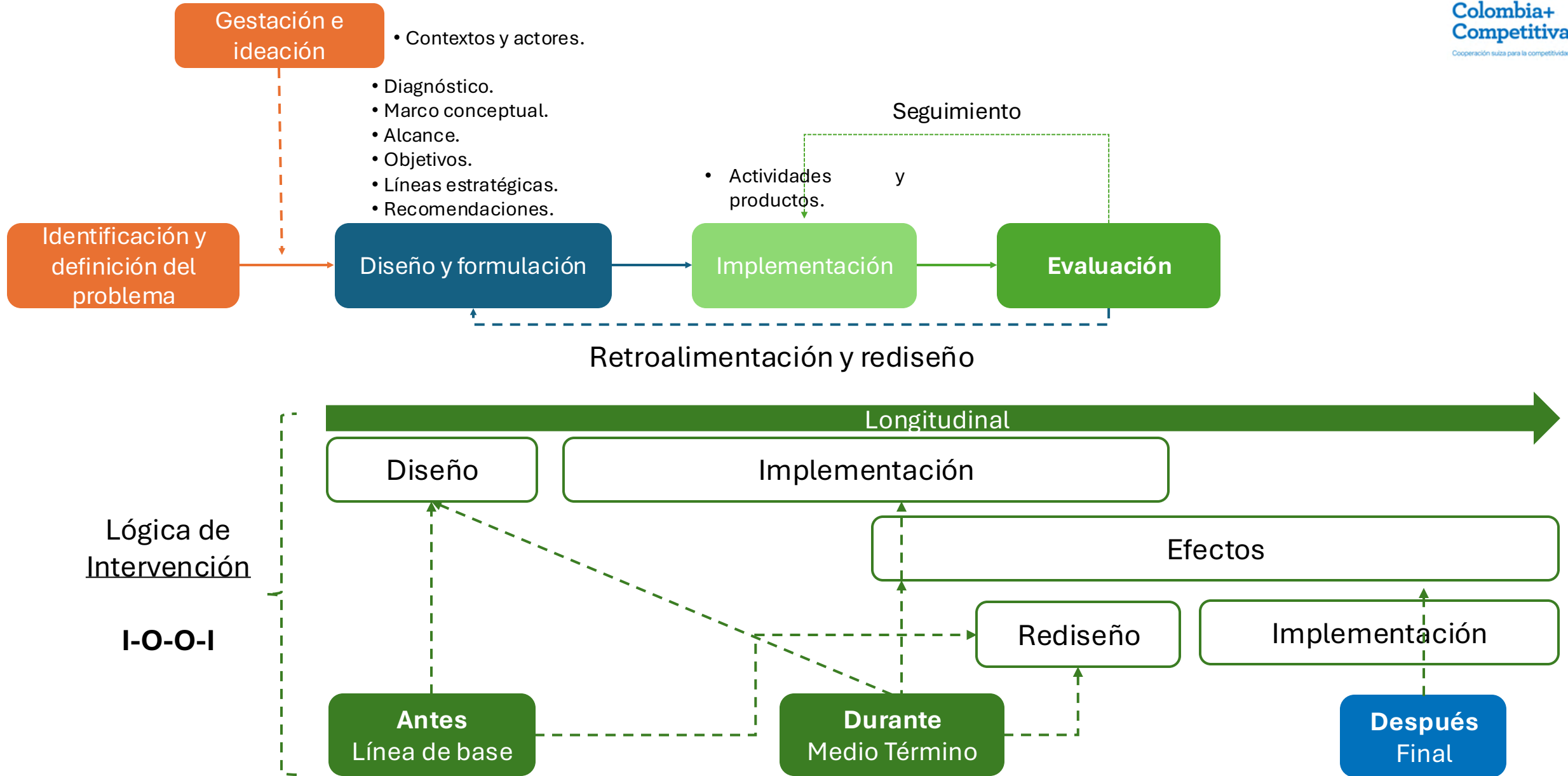
# ¿Evaluar o no evaluar?

NO es recomendable evaluar si:

- Lo beneficios no son claramente cuantificables.
- El programa no está claramente definido.
- El programa fue “descontinuado”.
- El costo-beneficio de la evaluación es cuestionable.

Experimentación – Flexibilidad - Provisionalidad

# Ciclo de las políticas públicas



Regiones +  
Competitivas




Colombia+  
Competitiva

Cooperación suiza para la competitividad

Programa de fortalecimiento de capacidades  
Módulo 9: “Evaluación de impacto” - Conceptos básicos

  
swisscontact

 Schweizerische Eidgenossenschaft  
Confédération suisse  
Confederazione Svizzera  
Confederaziun svizra  
  
Embajada de Suiza en Colombia  
Cooperación Económica y Desarrollo (SECO)

technopolis  
group

 Competitiveness

 OCYT  
OBSERVATORIO COLOMBIANO DE  
CIENCIA Y TECNOLOGÍA

 COLOMBIA  
POTENCIA DE LA  
VIDA

 Departamento Nacional  
de Planeación - DNP