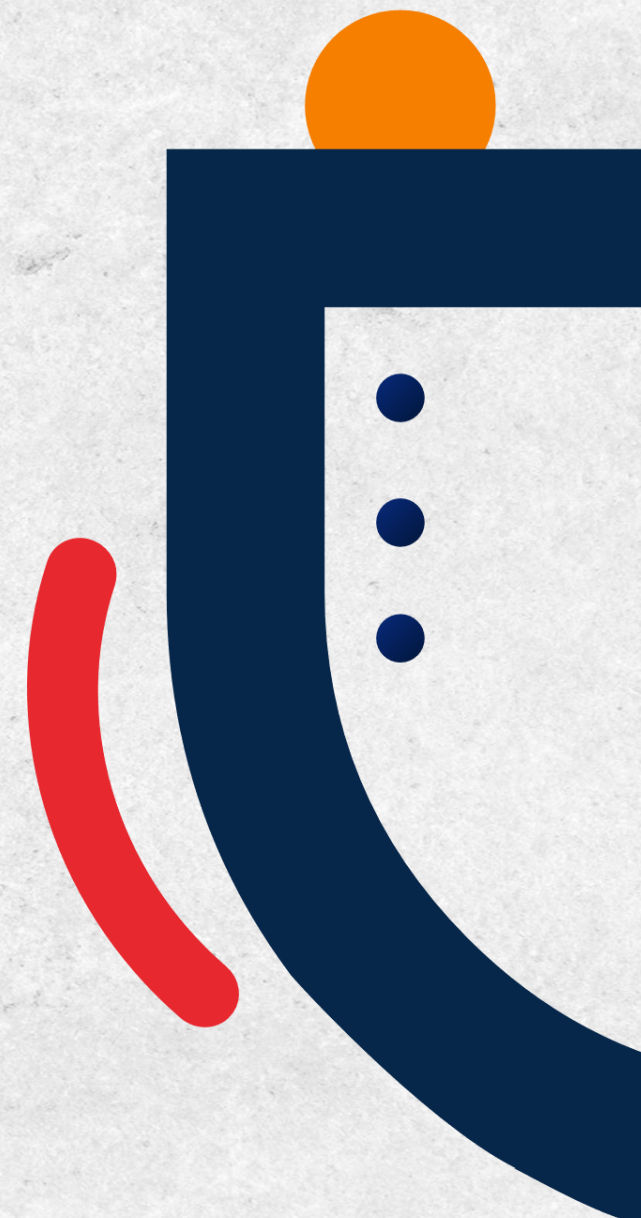




Evaluación de Impacto: Aprendizajes prácticos desde la academia

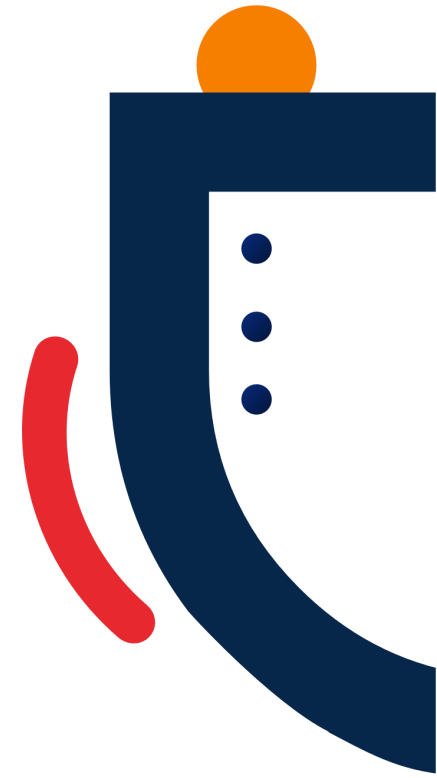
Catherine Rodríguez
Facultad de Economía y CEDE - Universidad de los Andes

Santo Domingo, Octubre 19 2023



Agenda

1. Objetivo e importancia de las evaluaciones de impacto
2. Principales retos técnicos en la implementación de evaluaciones de impacto
3. Aprendizajes de distintas evaluaciones de impacto en América Latina





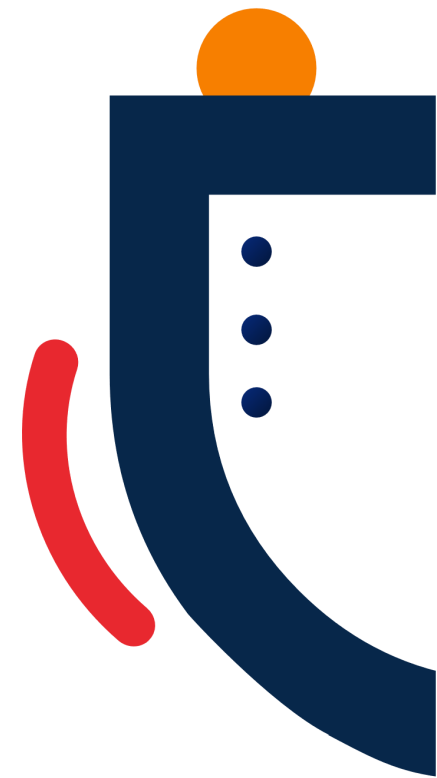
FIED 2023

Primer Foro Internacional de
Evaluación para el Desarrollo

Objetivo e importancia de las evaluaciones de impacto



-
- Las evaluaciones de impacto buscan responder una pregunta:
 - ❖ ¿Cuál es el **impacto** de un programa sobre sus beneficiarios?
 - ¿ Que es el impacto?
 - ❖ Es la **diferencia** entre los resultados que obtienen los beneficiarios de un programa dado vs. lo que **esos mismos beneficiarios** hubiesen obtenido en ausencia de ese mismo programa (*contrafactual*).



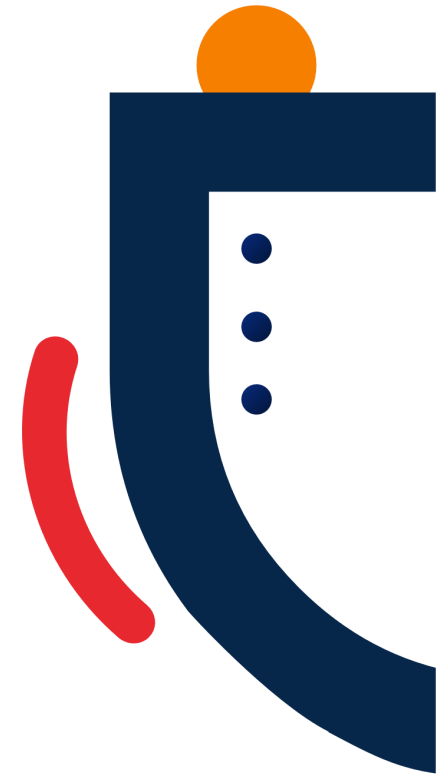
- ¿ Por qué son importantes?

- ❖ La mayoría de los programas se implementan con la **intención genuina** de mejorar un resultado de interés de la población beneficiaria.
- ❖ Sin embargo, algunos programas no funcionan como se esperaba, algunos pueden ser mejorados y otros pueden incluso escalarse.



Importancia

- Las evaluaciones contribuyen a un uso más eficiente de los recursos públicos:
 - ❖ Permiten basar decisiones en la evidencia y no en juicios impresionistas o ideologías.
 - ❖ Permiten guiar las decisiones de eliminación, cambios o escalamiento de programas.





FIED 2023

Primer Foro Internacional de
Evaluación para el Desarrollo

Principales retos técnicos en la implementación de evaluaciones de impacto

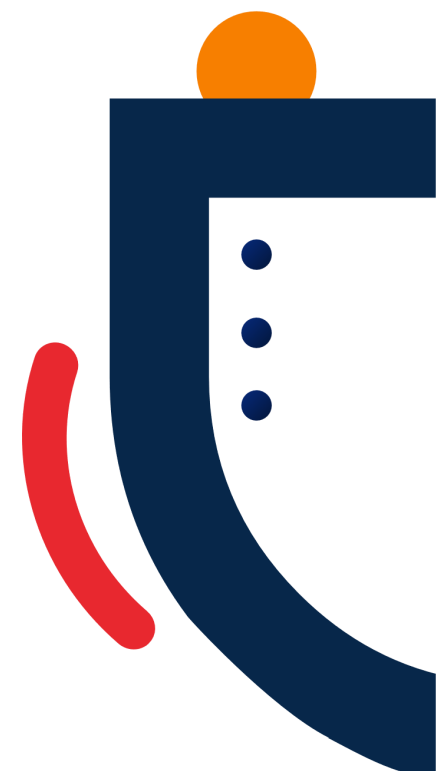
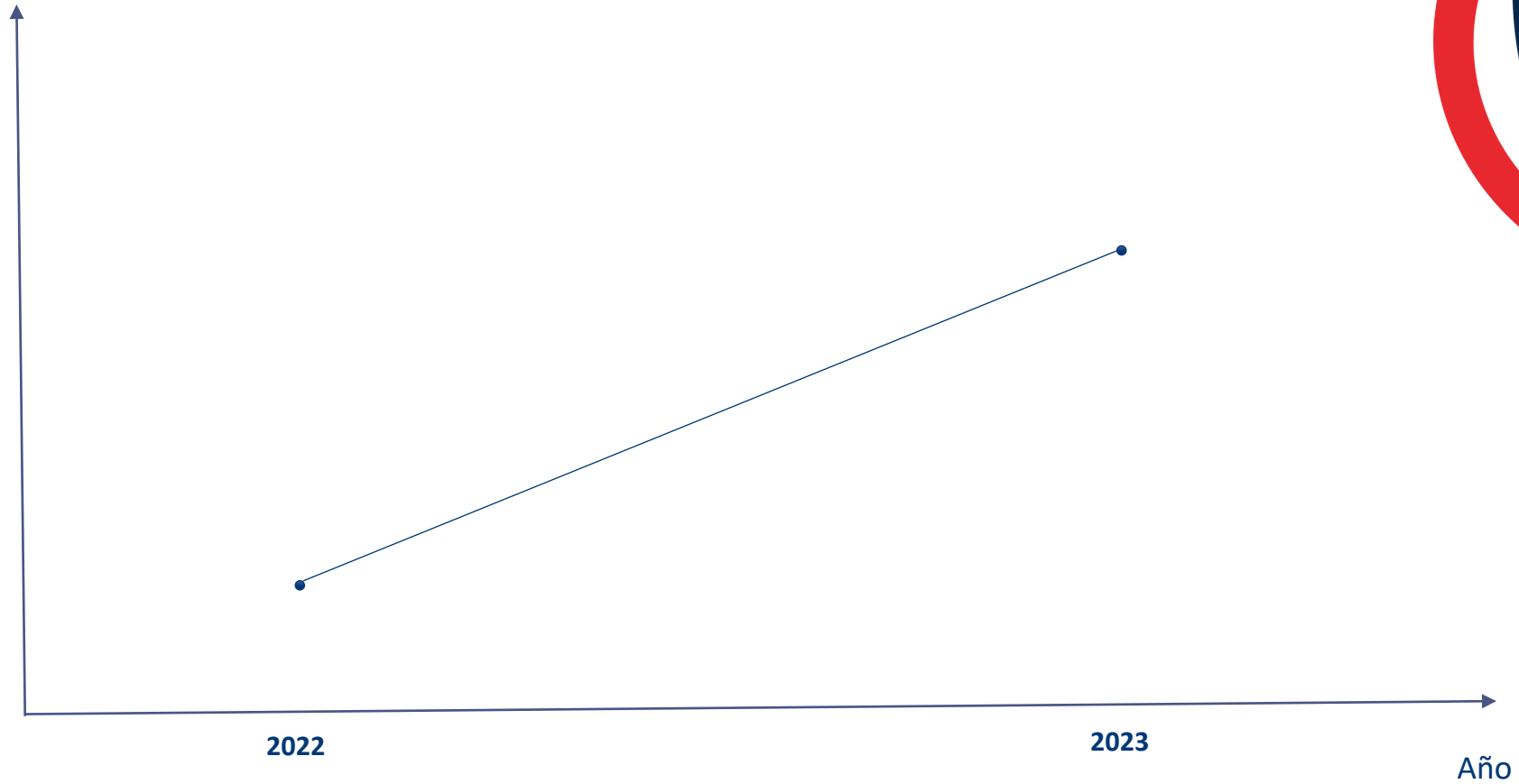


-
- El principal reto técnico: construcción de un buen **contrafactual**.
 - Contrafactual: Lo que les hubiese sucedido a los beneficiarios del programa en ausencia de este.
 - Mayor problema del *contrafactual*: **NO EXISTE, NO OBSERVABLE**
 - Los investigadores buscamos **construir artificialmente** este contrafactual de la mejor manera posible para asegurar que medimos el **impacto causal** de un programa.



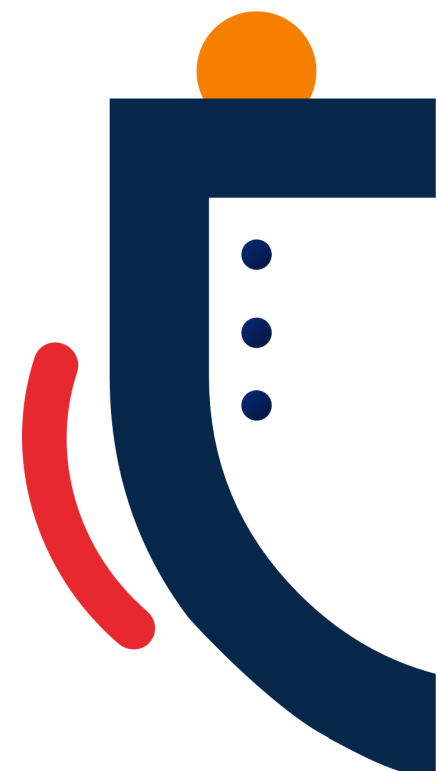
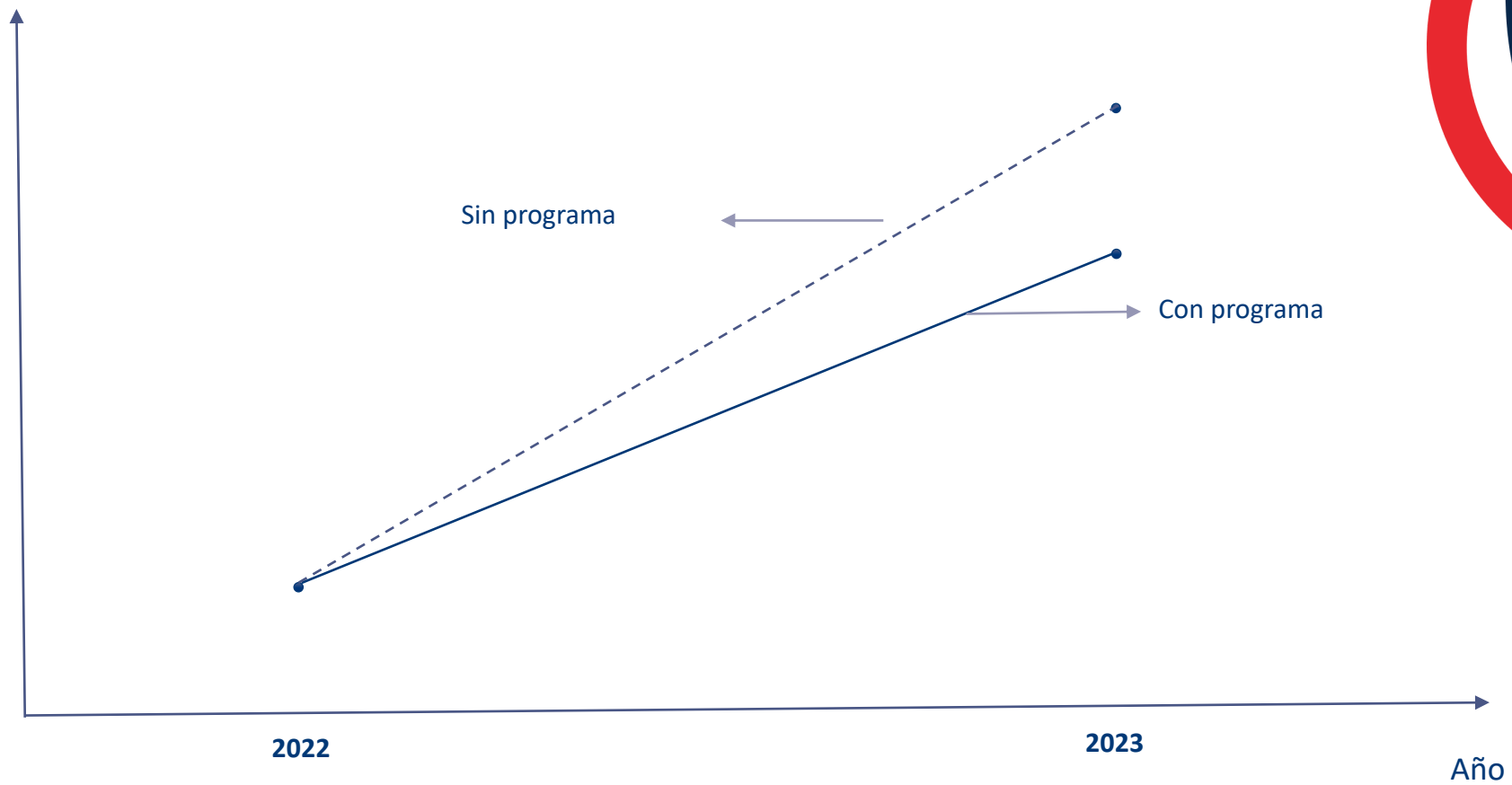
¿Qué pasa si tenemos dos observaciones del grupo de tratamiento? ¿Podemos medir impacto?

Prueba de Conocimiento



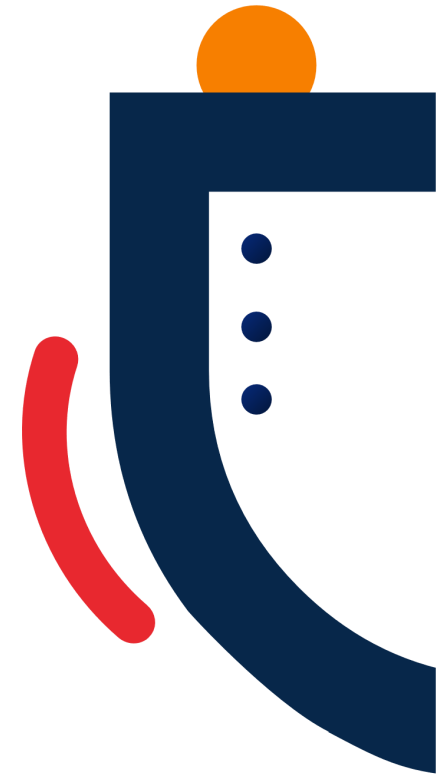
¿Tuvo impacto positivo?

Prueba de Conocimiento



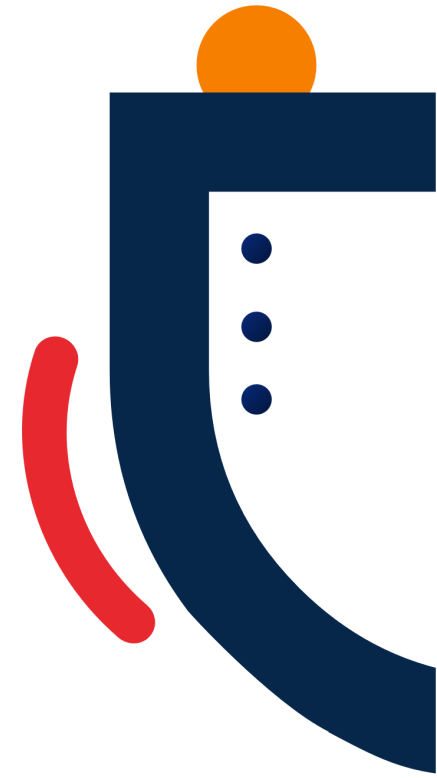
La correcta elección del contrafactual

- Necesitamos el contrafactual para poder establecer que los cambios observados en los beneficiarios se deben de manera causal al tratamiento.
- PERO la correcta elección del contrafactual presenta dos problemas comunes:
 - Los programas se focalizan a una población dada.
 - Participación en el programa es normalmente voluntaria.



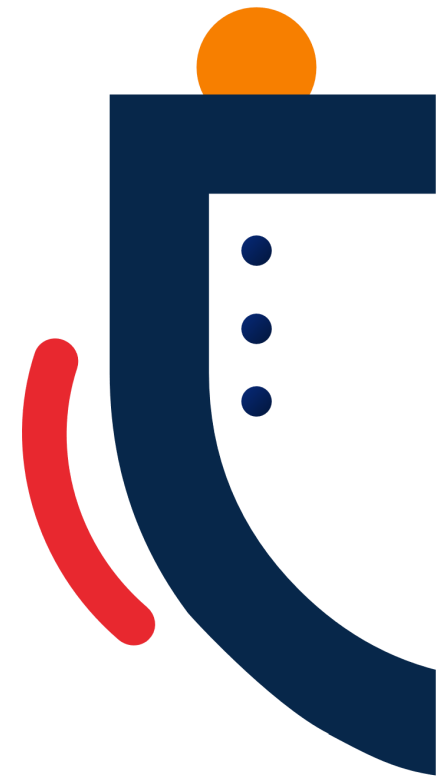
Requisitos de un contrafactual correcto

- Se requiere un grupo de control que sea lo más parecido posible en características **observables y no observables** a los individuos que reciben el programa y que no estén sujetos a contaminación del tratamiento.
- El investigador busca **encontrar clones** de los individuos tratados por el programa que no hayan sido expuestos a él.



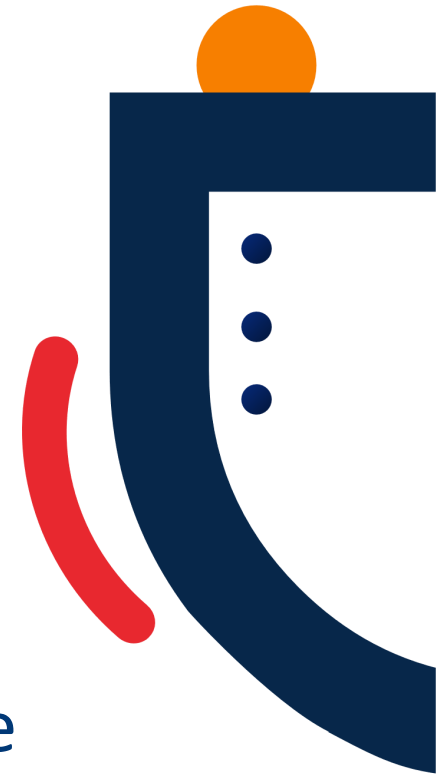
Metodologías para encontrar el contrafactual

- Experimentales
 - Randomized Control Trials (RCTs) – experimentos aleatorios controlados.
- No experimentales
 - Regresión Discontinua (cuasi – experimental)
 - Matching – Emparejamiento estadístico
 - Variables Instrumentales
 - Diferencias en diferencias
 - Combinaciones



¿Cuál de todas las metodologías escoger?

- La elección de la mejor metodología dependerá de una combinación de factores:
 1. Pregunta que se busca responder.
 2. Proceso de selección de los individuos o grupos que son elegibles/ beneficiarios.
 3. Tiempo de madurez del programa que se busca evaluar.
 4. Costos de la evaluación.
 5. Temas éticos de la evaluación.





FIED 2023

Primer Foro Internacional de
Evaluación para el Desarrollo

Aprendizajes de distintas evaluaciones de impacto en América Latina



Cuatro ejemplos muy distintos

- Evaluación de impacto que permitió la mejora de un programa.

A



- Evaluación de impacto que a pesar de demostrar los beneficios del programa fue ignorada por temas ideológicos.

B



- Evaluación de impacto que permitió aprendizajes mutuos entre academia y gobierno.

C





FIED 2023

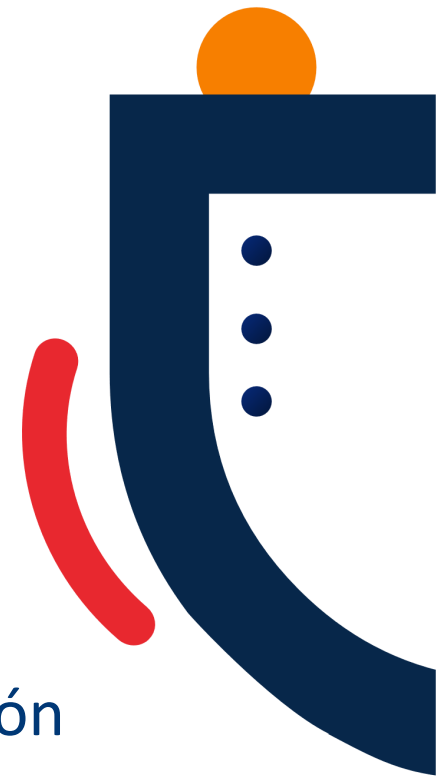
Primer Foro Internacional de
Evaluación para el Desarrollo

Ejemplo A:
Evaluación de impacto que
permitió la mejora de un
programa



Ejemplo A: PTA

- Aspectos técnicos:
 - Elección del correcto contrafactual. Fue necesario implementar distintas metodologías de impacto: RD, matching.
 - Uso de múltiples fuentes de información: diseño e implementación de encuestas e instrumentos de observación en clase.
- Aspectos de política:
 - Evaluación permitió la mejora en la implementación de una política y en el aprendizaje de estudiantes vulnerables en Colombia.



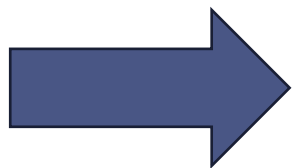
Ejemplo A: PTA

- PTA: estrategia integral para mejorar aprendizajes de un grupo de escuelas en Colombia a través de la **formación docente y entrega de materiales**:
 - Beneficiar a 2.300.000 estudiantes de básica primaria en lenguaje y matemáticas.
 - Formar y acompañar a 70.000 educadores entre docentes de aula y directivos.
- La evaluación de impacto usó estrategias múltiples de evaluación (RD y DID matching) con datos cuantitativos administrativos y primarios y **datos cualitativos**.

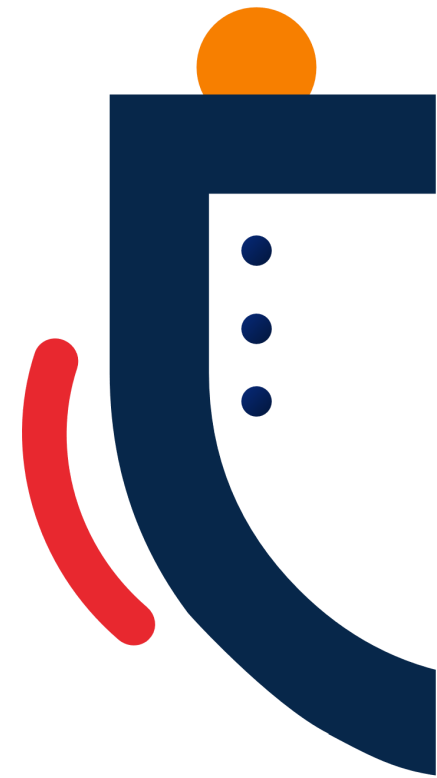


Ejemplo A: PTA

- Impactos positivos encontrados:
 - Aumentó disponibilidad y uso de materiales educativos.
 - Mayor planeación de clase y la creación de comunidades de aprendizaje.
- Impactos negativos encontrados:
 - Menos prácticas activas: Menor tiempo efectivo de enseñanza y menor interacción docente-estudiante.

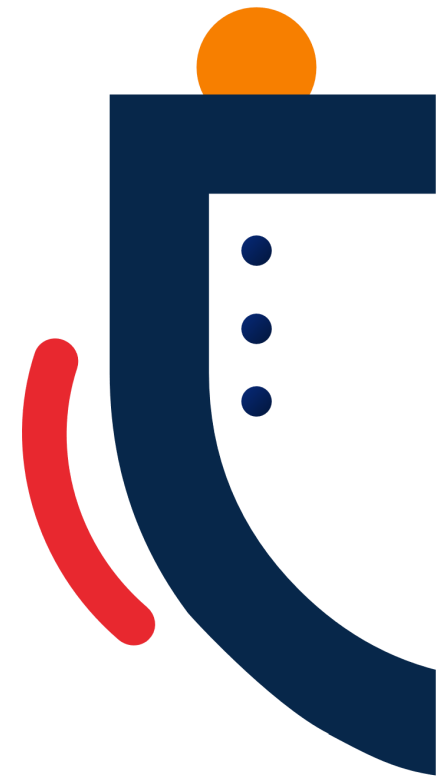


No se encontraron impactos ni en repitencia, deserción ni en resultados en las pruebas de conocimiento estandarizadas.



Ejemplo A: PTA

- La evaluación dejó **recomendaciones específicas** en términos de implementación y diseño del programa.
- Las recomendaciones fueron **tenidas en cuenta** por el Gobierno y se creó la versión mejorada del programa Todos a Aprender 2.0.
- Una evaluación de impacto **posterior** permitió demostrar que los aprendizajes de estudiantes en escuelas tratadas bajo el nuevo modelo mejoraron.





todos a aprender 2.0

PROGRAMA PARA LA EXCELENCIA DOCENTE Y ACADÉMICA



todos a aprender

PROGRAMA PARA LA TRANSFORMACIÓN DE LA CALIDAD EDUCATIVA

La evaluación de impacto permitió la mejora de un programa existente.



FIED 2023

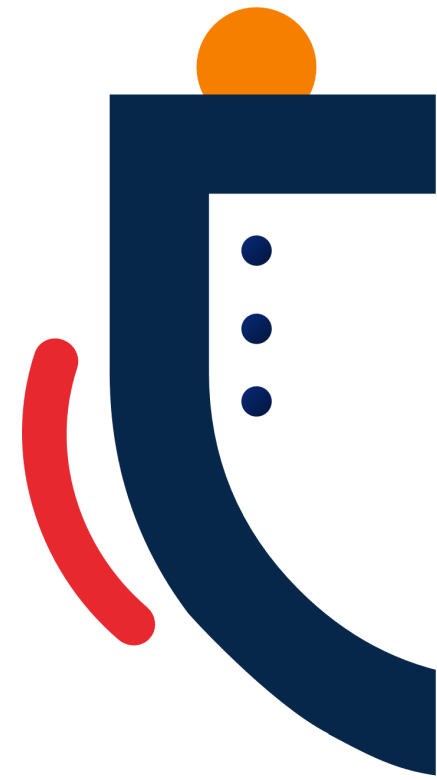
Primer Foro Internacional de
Evaluación para el Desarrollo

Ejemplo B:
Evaluación de impacto que, a
pesar de demostrar los
beneficios del programa, fue
ignorada por temas ideológicos.



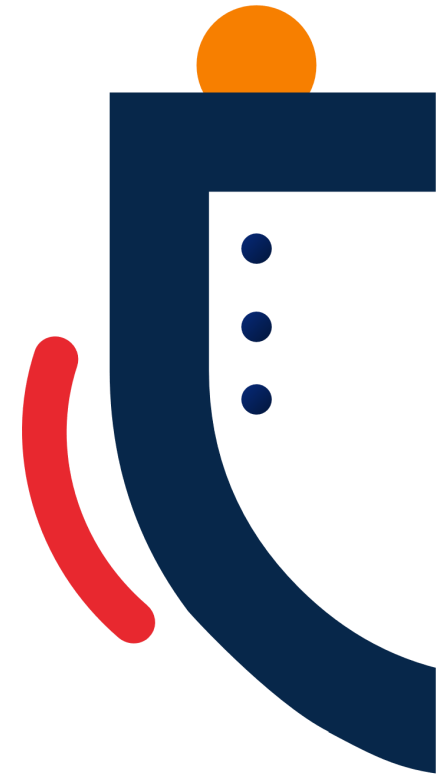
Ejemplo B: Ser Pilo Paga

- Aspectos técnicos:
 - Facilidad en la elección de metodología y selección del correcto contrafactual: RD.
 - Uso de información administrativa – bajo costo.
- Aspectos políticos:
 - A pesar de los impactos evidentes primaron las opiniones ideológicas.

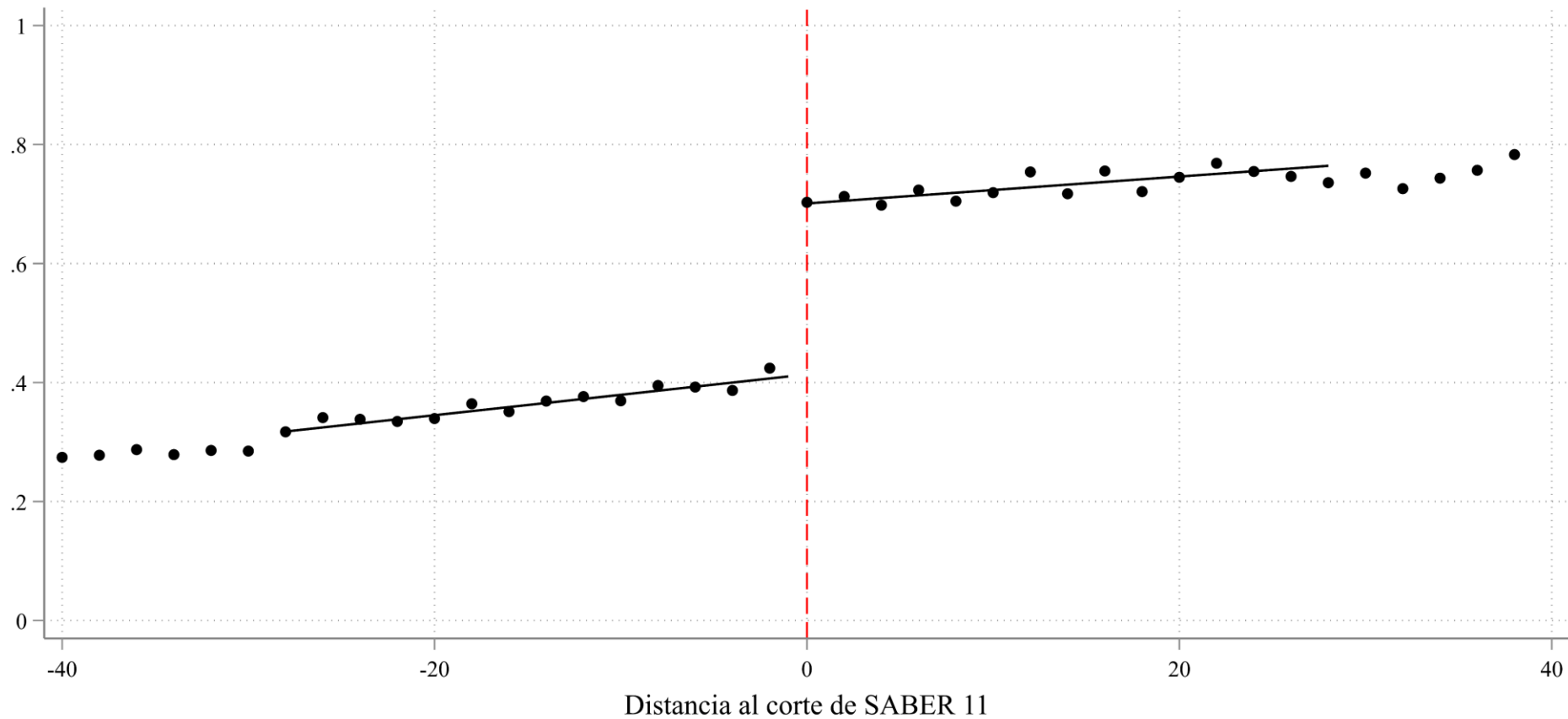


Ejemplo B: Ser Pilo Paga

- Ser Pilo Paga fue un crédito 100% condonable ofrecido a 40,000 jóvenes por el Gobierno Colombiano que cubre el valor de la matrícula en un programa profesional en las mejores universidades del país.
- Para ser elegible el joven debía cumplir requisitos de mérito y elegibilidad.
- La evaluación de impacto uso estrategia de *RD* con datos cuantitativos administrativos y primarios. Se complemento con datos cualitativos.



Impactos SPP en acceso a educación superior



Muestra restringida a elegibles por SISBEN



Ejemplo B: Ser Pilo Paga

- Impactos cuantitativos en movilidad social, equidad, eficiencia en acceso, graduación, aprendizaje y mercado laboral.
- Programa altamente costo-eficiente: por cada dólar invertido se generó casi dos dólares de retornos privados.
- A pesar de esto, por temas más ideológicos que técnicos, el programa fue eliminado.





FIED 2023

Primer Foro Internacional de
Evaluación para el Desarrollo

Ejemplo C:
Evaluación de impacto que
permitió aprendizajes mutuos
entre academia y gobierno.



Ejemplo C: PLIMyR

- Aspectos técnicos:
 - Diseño de estrategias novedosas a problemas comunes y conocidos.
- Aspectos políticos:
 - Colaboración cercana entre gobierno y academia.
 - Aprendizajes mutuos en implementación, recolección de evidencia y evaluación de impacto.

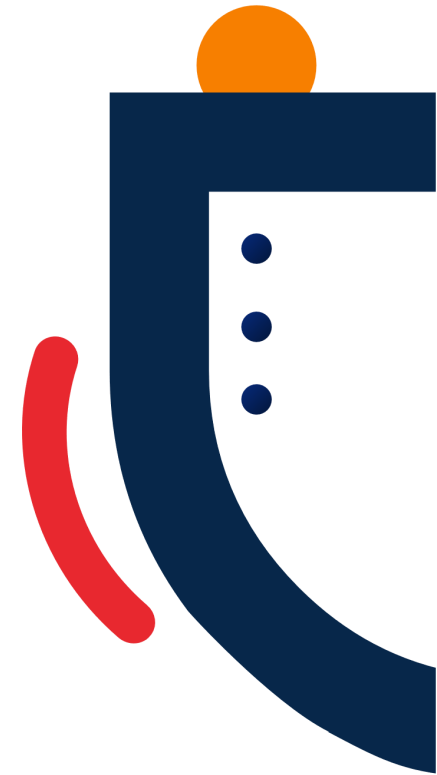


Ejemplo C: PLIMyR

- En República Dominicana se realizan **evaluaciones estandarizadas** en los niveles de Primaria y Secundaria desde hace **tres décadas**.

Internacionales - Pruebas Muestrales	Nacionales – Pruebas censales
UNESCO (PERCE, SERCE, TERCE)	Pruebas Nacionales (desde 1991)
PISA (2017, 2019)	Pruebas Diagnósticas (desde 2017 en distintos grados)

- El Ministerio de Educación genera reportes e informes a nivel nacional, regional, distrital y de centros educativos, estableciendo una **política de difusión** a partir de sus resultados, los cuales deben ser utilizados para la mejora de la calidad educativa de acuerdo con la Ordenanza 1-2016



Problemas asociados a la difusión de resultados:

Bajo conocimiento de las estrategias de difusión

- Falta de evidencia sobre:
 - **Difusión** efectiva,
 - **Comprensión** de los reportes,
 - **Uso** , y
 - **Mejora educativa** gracias a la difusión
- La política de difusión no es **estandarizada**
- **No existe evidencia** oficial del cumplimiento de esta política

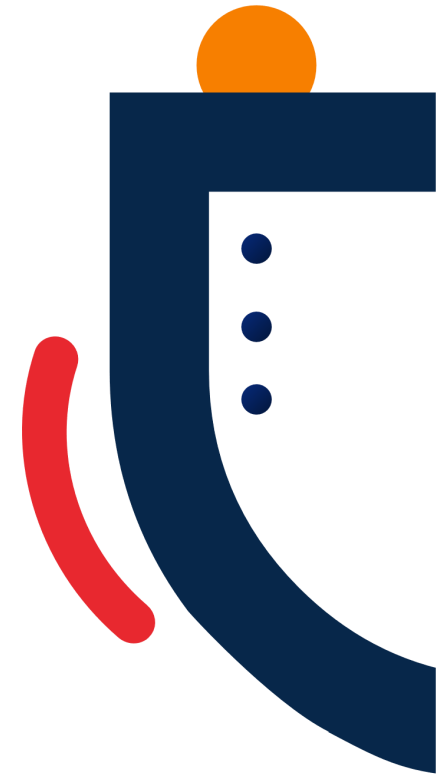
Bajo nivel de adopción de políticas de bajo riesgo

- Evaluaciones experimentales han estimado **bajos niveles de adopción** de políticas de bajo riesgo (*low stake tasks*):
 - Programa de escalamiento la campaña AVE-RD sobre el valor de la educación: 30% en el grupo de control.
 - Estrategia de difusión de las Evaluaciones Diagnósticas de primaria: 10% en el grupo de control.



Ejemplo C: PLIMyR

- Investigadores y *Policy Makers* del MINERD trabajaron en conjunto para:
 - **Estandarizar el proceso** de difusión de resultados de la ED-2019.
 - Descubrir formas efectivas de **aumentar el cumplimiento de políticas de bajo riesgo.**
 - **Medir el nivel de adopción** y efectividad de la estrategia.



Ejemplo C: PLIMyR

Estrategias innovadoras para alcanzar objetivos

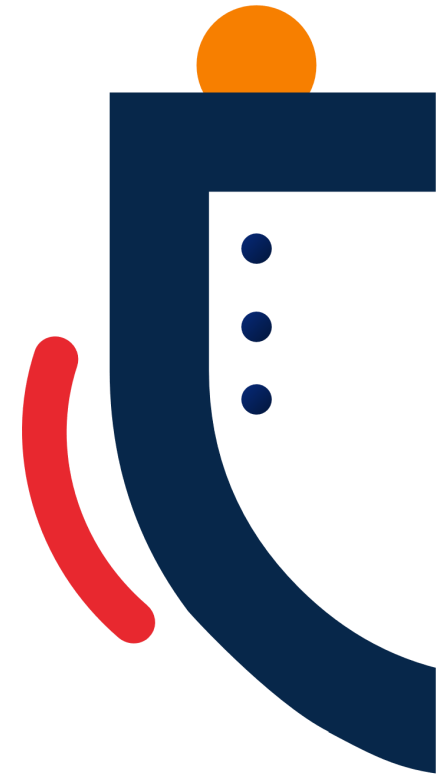
1. Creación de una App de monitoreo:

- Plataforma Interactiva de Monitoreo y Respuesta (**PLIMyR**) que permite monitorear el cumplimiento de políticas educativas y “tareas” en tiempo real.

2. Campañas de comunicación y capacitación Regionales-Distritales.

3. Implementación de estrategias innovadoras para aumentar niveles de adopción:

- i. RCT basado en estrategias de motivación e instrucción.
- ii. Monitoreo intensivo
- iii. Ultima milla



Ejemplo C: PLIMyR

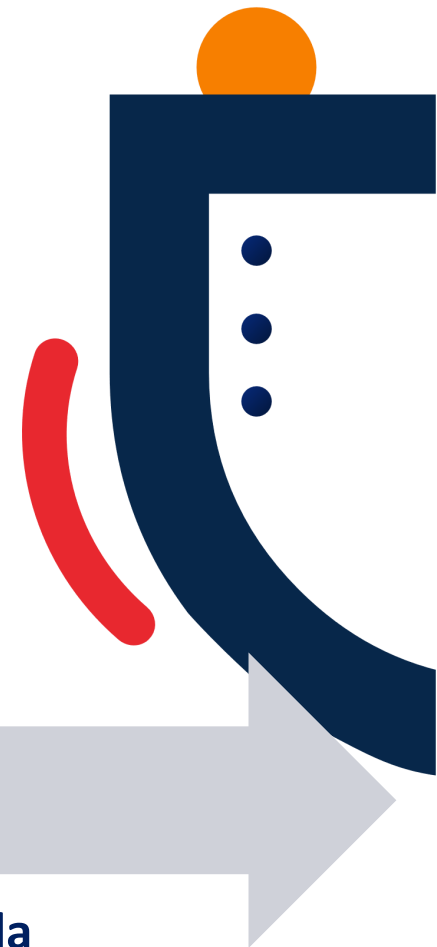
Campañas de
Capacitación
(febrero –abril)

Fase 2 – Monitoreo Intensivo

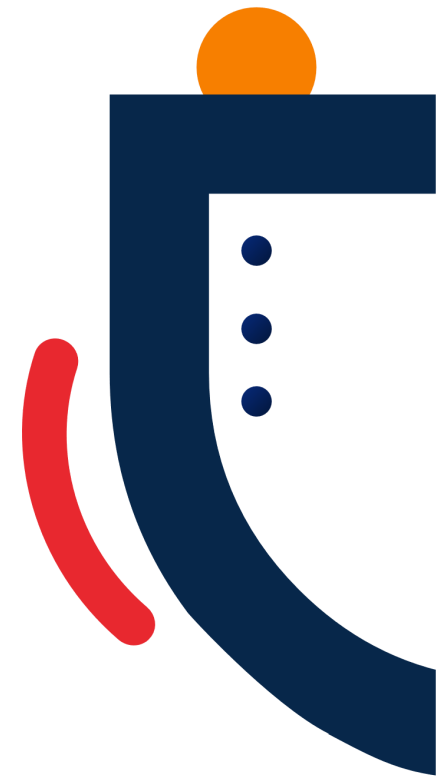
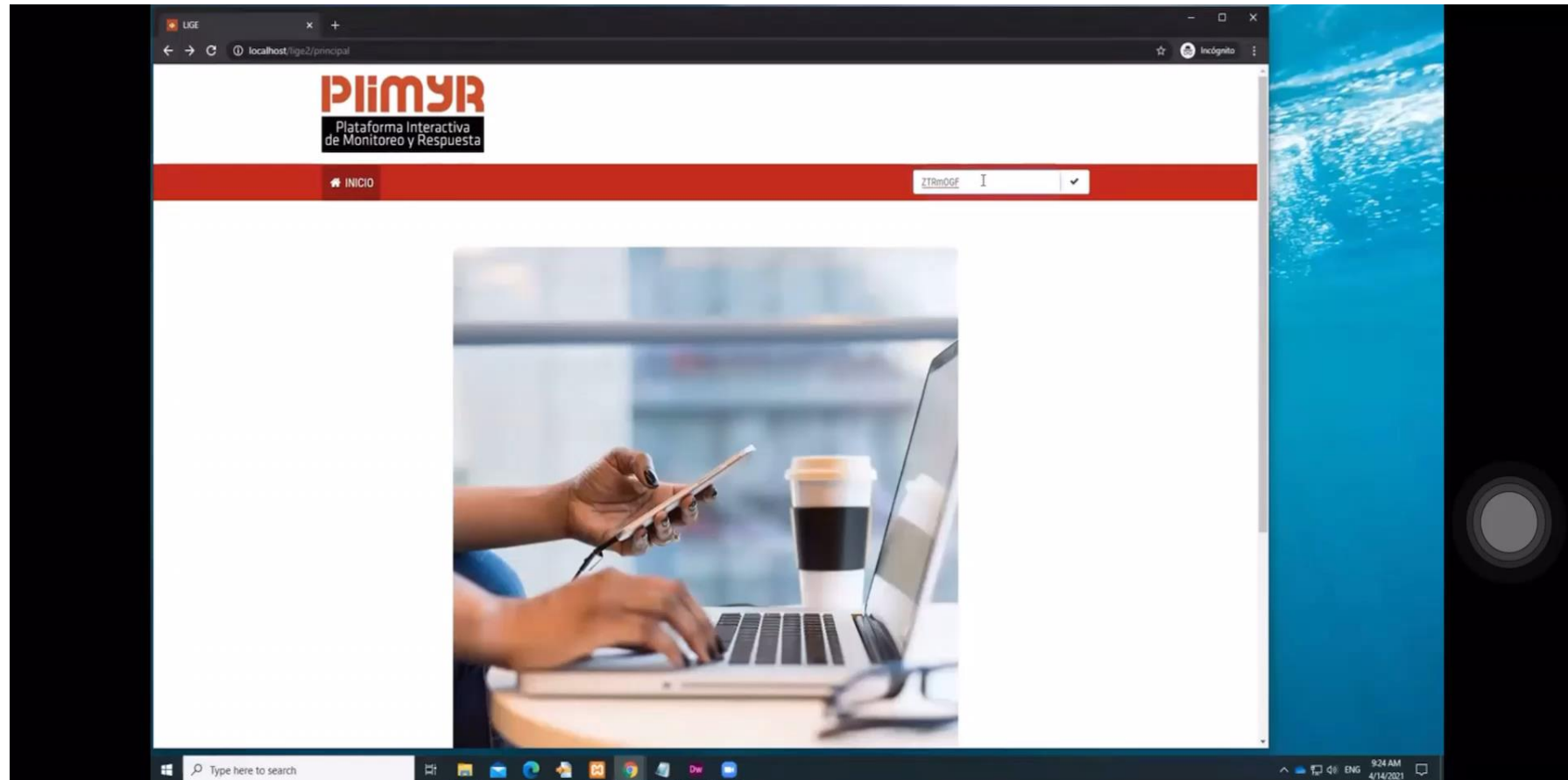
Extensión fecha
límite y reportes de
inteligencia (17-28
de mayo)

**Fase 1 –
Experimental**
motivación vs.
instrucción (16
abril al 14 de
mayo)

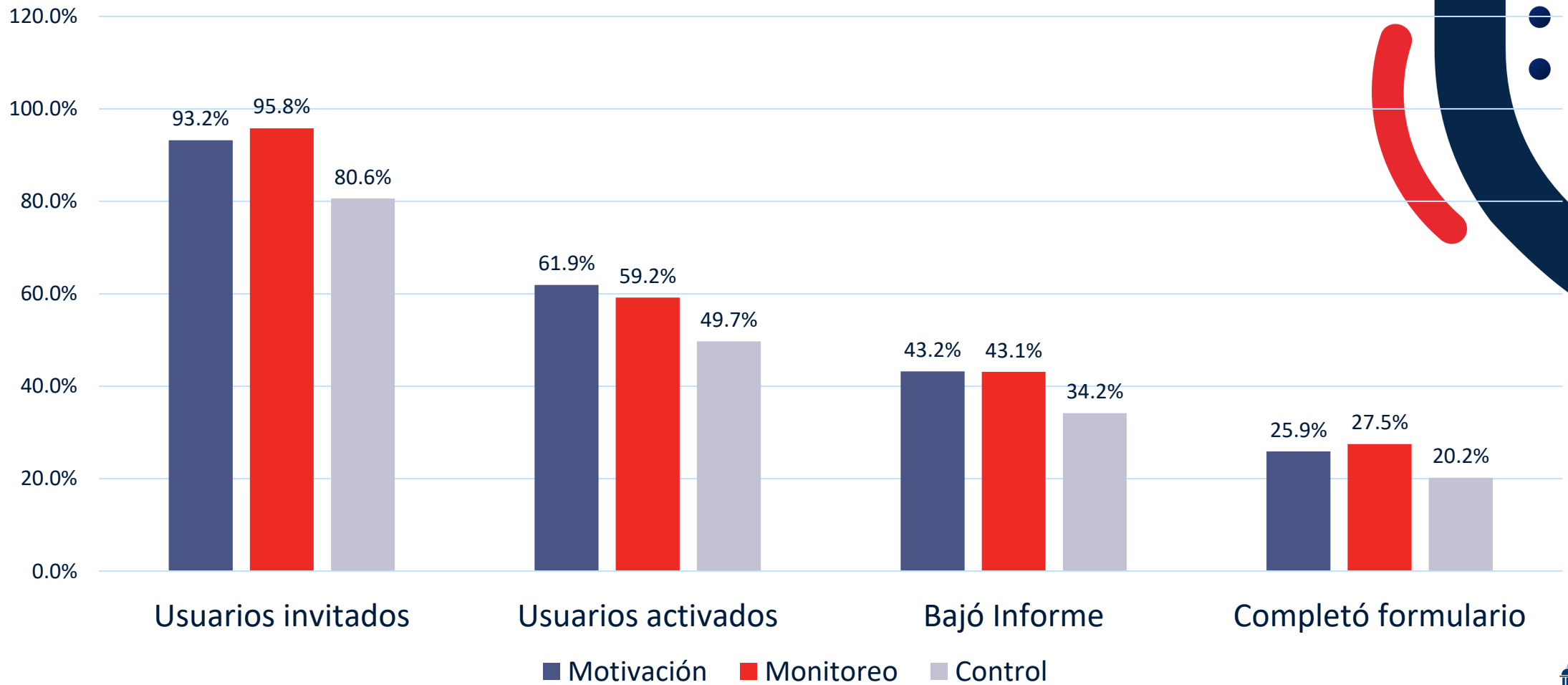
Fase 3- Última milla
Detección cuellos de botella
y extensión final fecha límite
(30 de mayo – 18 de junio)



Ejemplo D: PLIMyR



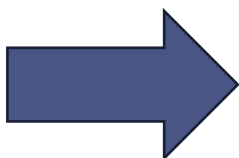
Ejemplo C: PLIMyR



Ejemplo C: PLIMyR

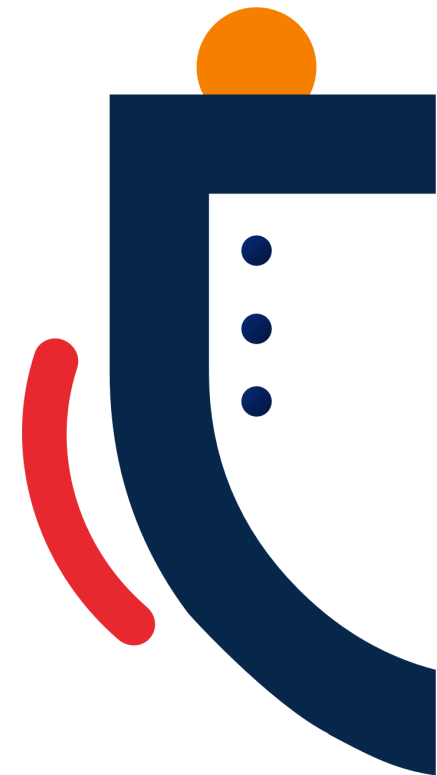
- **Fase 2 – Monitoreo Intensivo**

- Primer informe resultados a Distritos **y Regionales** - 21% cumplimiento
- Extensión del plazo 2 semanas más.
- Envío de reportes de inteligencia de cumplimiento a Distritos & Regionales **cada tres días.**



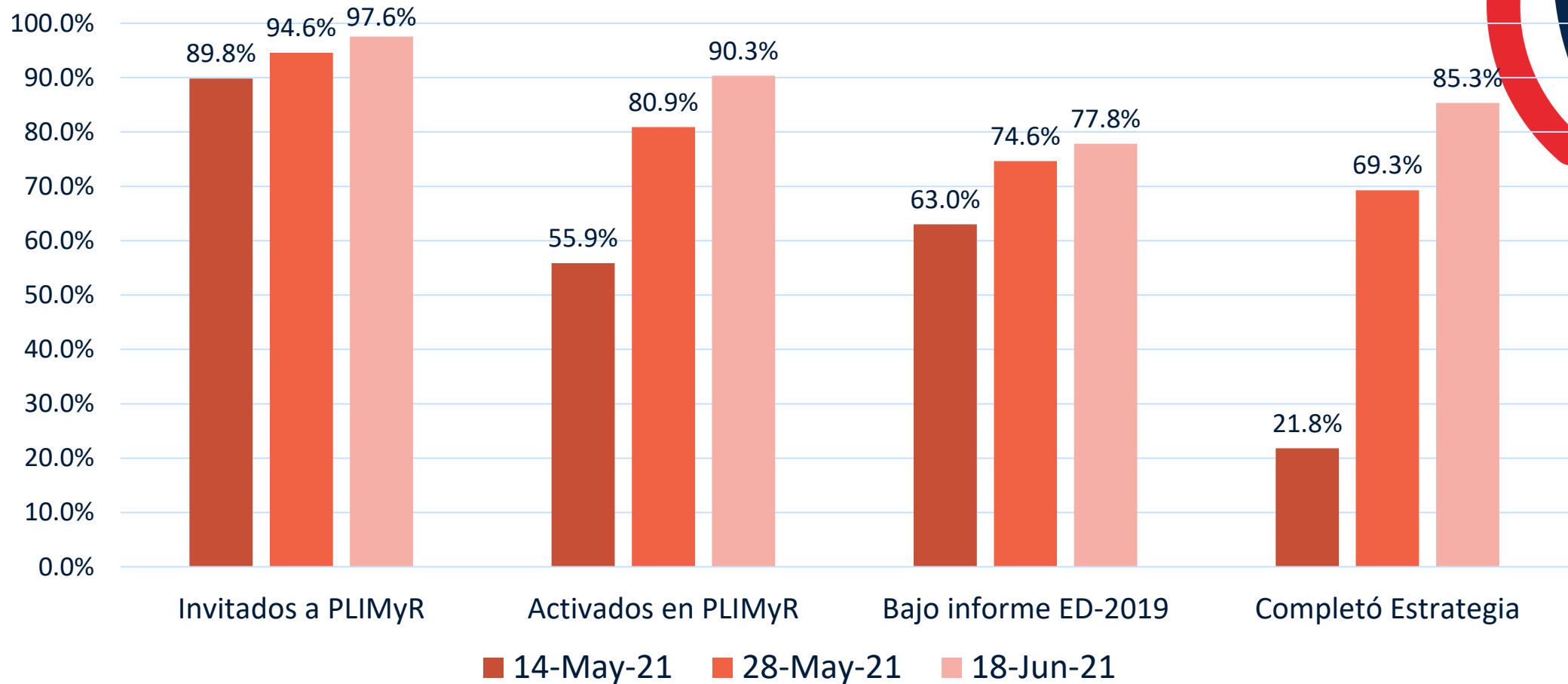
- Monitoreo del MINERD a Distritos y Regionales
- Evidencia que con App PLIMyR podían ellos mismos monitorear también a sus Distritos y Centros.

- **Fase 3 - Última milla/ cuellos de botella**



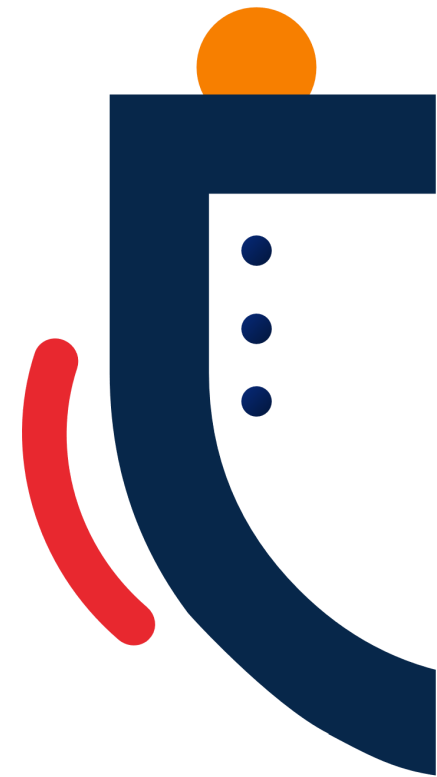
Ejemplo C: PLIMyR

Resultados finales por fase



Ejemplo C: PLIMyR

- Principales aprendizajes:
 - A pesar de toda la información y capacitaciones proporcionadas, hay un **bajo nivel de cumplimiento** de políticas de bajo costo (21.8% en fase 1).
 - La estrategia de monitoreo implementada en fases 2 y 3 aumentó en 63.5 puntos porcentuales los directores que adoptaron la estrategia ED-2019.
 - Es decir, aproximadamente 1,505 Directores/as implementaron los talleres de difusión de la Evaluación Diagnóstica gracias a **estrategias basadas en evidencia, monitoreo creíble, motivación y seguimiento.**





FIED 2023

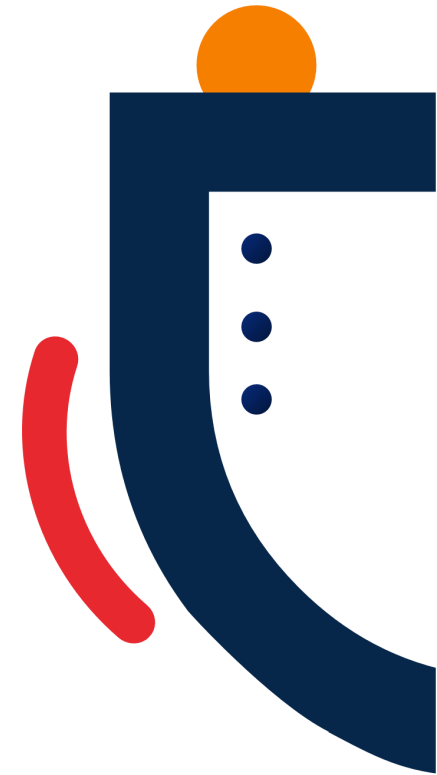
Primer Foro Internacional de
Evaluación para el Desarrollo

Conclusiones



Conclusiones

- La evaluación de impacto es un **proceso de enseñanza y aprendizajes mutuo** para hacedores de política y académicos.
- La evaluación de impacto permite tener **información para hacer un uso eficiente de los recursos** escasos.
- Academia y *hacedores de política* pueden **colaborar mutuamente** para descubrir e investigar la efectividad de innovaciones que mejoren la implementación y resultados de políticas sociales.





Muchas gracias!

Catherine Rodríguez

cathrodr@uniandes.edu.co



<https://issuu.com/> <https://mepyd.gob.do/>

