

# Seminario regional: Introducción a la evaluación de impacto de políticas y programas públicos

**Dra. Laura H. Atuesta Becerra**

**[laura.atuesta@cide.edu](mailto:laura.atuesta@cide.edu)**

# Algunos comentarios generales antes de empezar...



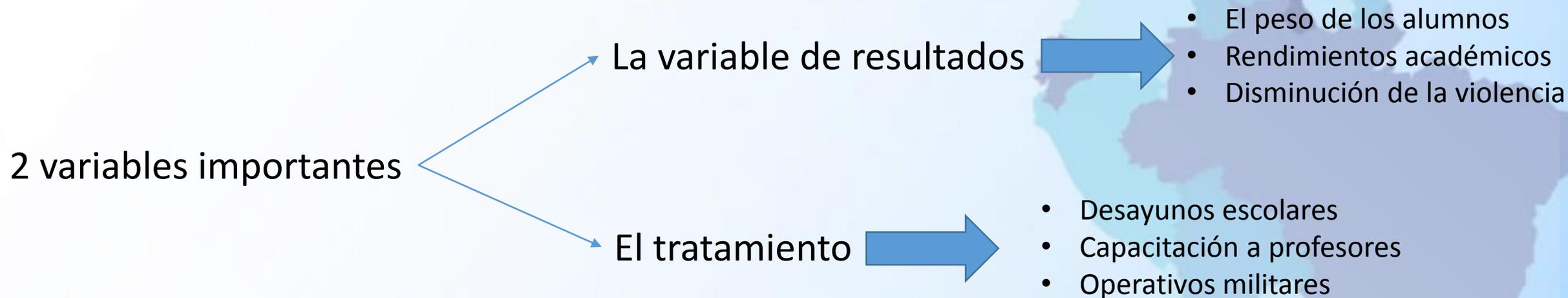
- Preguntas de evaluación: ¿Cuál es el impacto de **D** en **Y**?
  - D: intervención.
  - Y: variable de resultados o variable de impacto.
- ¿Qué podemos medir con la evaluación de impacto y qué no podemos medir?
  - Medición de impacto: diferencia en el promedio de la variable de resultados entre el grupo de tratamiento y el grupo de control.
- ¿Cuáles son las desventajas de la evaluación de impacto?
  - No nos da ninguna explicación sobre el impacto.
  - No le importa el dinero!!!
  - Es un método estadístico, por lo que siempre está sujeta a las limitantes estadísticas.

# Algunos comentarios generales antes de empezar...



- Aunque la aleatorización no es un método costoso, la implementación del protocolo de aleatorización es complicada.
  - ¿Cómo aleatorizar en programas sociales?
  - ¿Qué hacemos cuando alguien seleccionado para participar no quiere hacerlo?
  - Costos políticos de la aleatorización.
- Algo de diseño muestral
  - Importante saber qué nivel de análisis estamos manejando (nivel individual, nivel escuela, nivel barrio).
  - La aleatorización debe corresponder al nivel de análisis que estemos usando.
  - Aleatorización por conglomerados siempre es más barato que aleatorización simple.
  - Aleatorización por estratificación siempre es más costosa que la aleatorización simple.

# Elementos de la evaluación de impacto

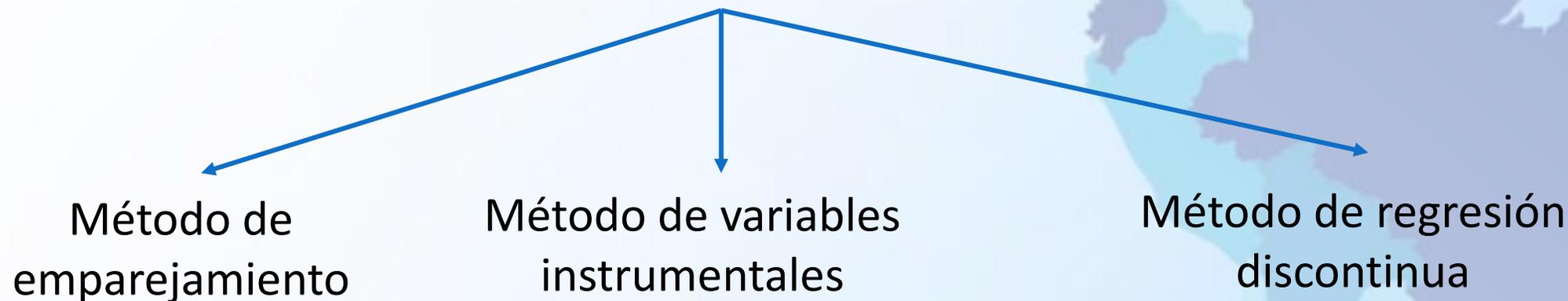


Normalmente los resultados deben estar incluidos en el marco lógico. Tener claro qué queremos lograr con el programa.

¿Qué pasa cuando no tenemos aleatoriedad?  
(que son la mayoría de los casos...)



## Estudios no experimentales



¿Cuándo usamos uno u otro método?

Depende de los datos que tengamos...

# Evaluación de impacto

## Diseños no experimentales

**Dra. Laura H. Atuesta Becerra**

**[laura.atuesta@cide.edu](mailto:laura.atuesta@cide.edu)**

# Método de emparejamiento

# Método de emparejamiento



- El método de emparejamiento “empareja”.
- Asegura que el grupo de tratamiento y el grupo de control son parecidos y comparables.
- ¿Cómo? Buscamos “clones” y comparamos a individuos parecidos.

# ¿Cuál es la lógica del emparejamiento? Ejemplo...



Pregunta: Queremos saber cuál es el efecto de estudiar en Harvard en el nivel de ingresos de los alumnos graduados.

Tratamiento → Haber estudiado la carrera en Harvard.

Variable de resultados → Nivel de ingresos de los recién graduados.

# Ejemplo

## GRUPO DE TRATAMIENTO: Graduados de Harvard



### Problemas?

- No tenemos datos de ellos antes de entrar a Harvard (la mayoría no trabajaba).
- La selección de alumnos no fue aleatoria (exámenes, entrevistas, aplicaciones).



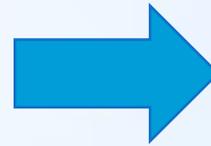
¿Qué hacemos?  
¿cómo conformamos el grupo de control?

Tenemos que buscar un grupo con características similares a los estudiantes de Harvard para comparar sus ingresos!

# Opciones...



Escogemos graduados de MIT...



Y los comparamos con los graduados de Harvard...



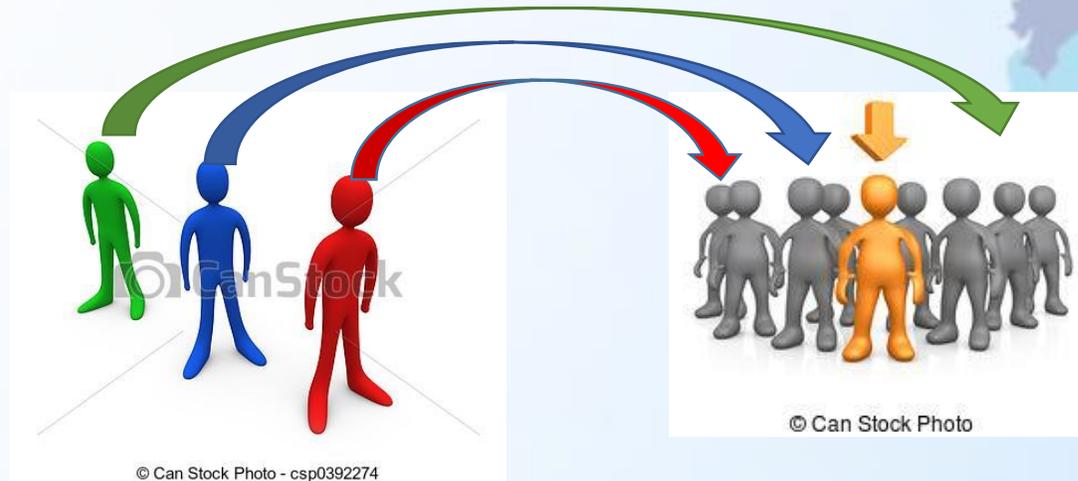
¿Qué problemas podemos tener?

# Supongamos que sabemos las características que identifican a los estudiantes de Harvard...



## Estudiantes de Harvard

$X_1$ : nivel económico de la familia  
 $X_2$ : composición familiar  
 $X_3$ : resultado de los exámenes  
 $X_4$ : alumnos recién graduados  
 $X_5$ : sexo

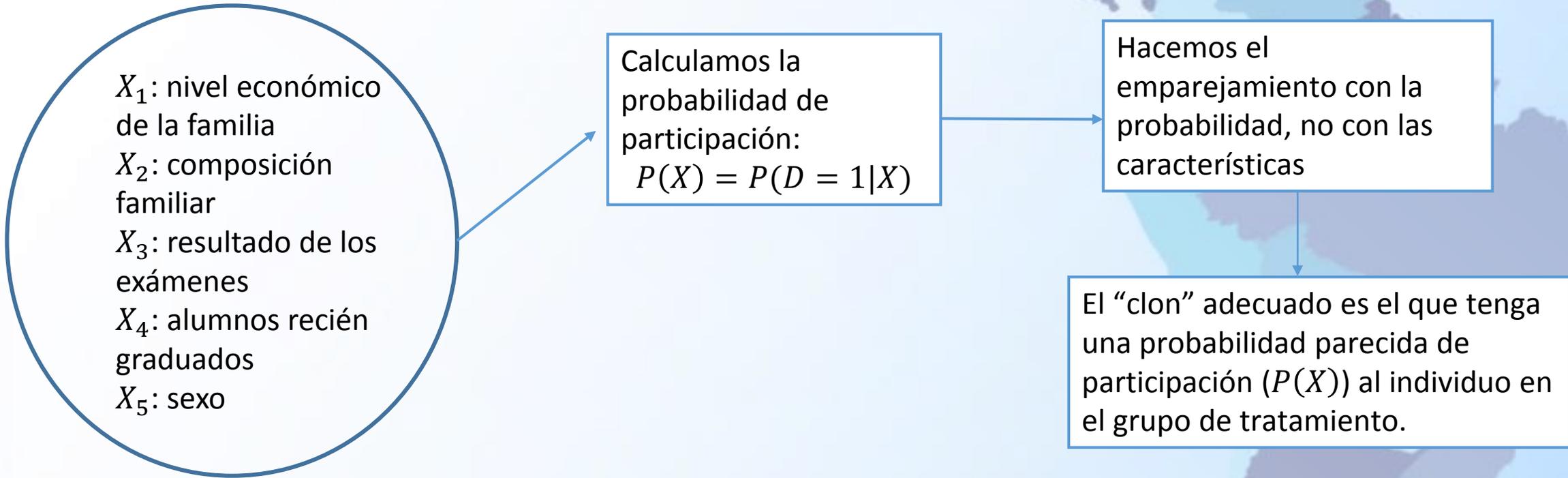


## Estudiantes de MIT y otras universidades

$X_1$ : nivel económico de la familia  
 $X_2$ : composición familiar  
 $X_3$ : resultado de los exámenes  
 $X_4$ : alumnos recién graduados  
 $X_5$ : sexo

Entre más características más difícil encontrar un clon!! “Maldición de la dimensionalidad”

# ¿Cómo armamos el grupo de control?



¿Por qué hacemos la comparación con la probabilidad y no con cada una de las características?

# Ventajas y desventajas del PSM



## VENTAJAS

- Podemos tener un solo periodo de tiempo sin línea base.
- Se usa frecuentemente porque se puede estimar cuando ningún otro método experimental se puede usar.

## DESVENTAJAS

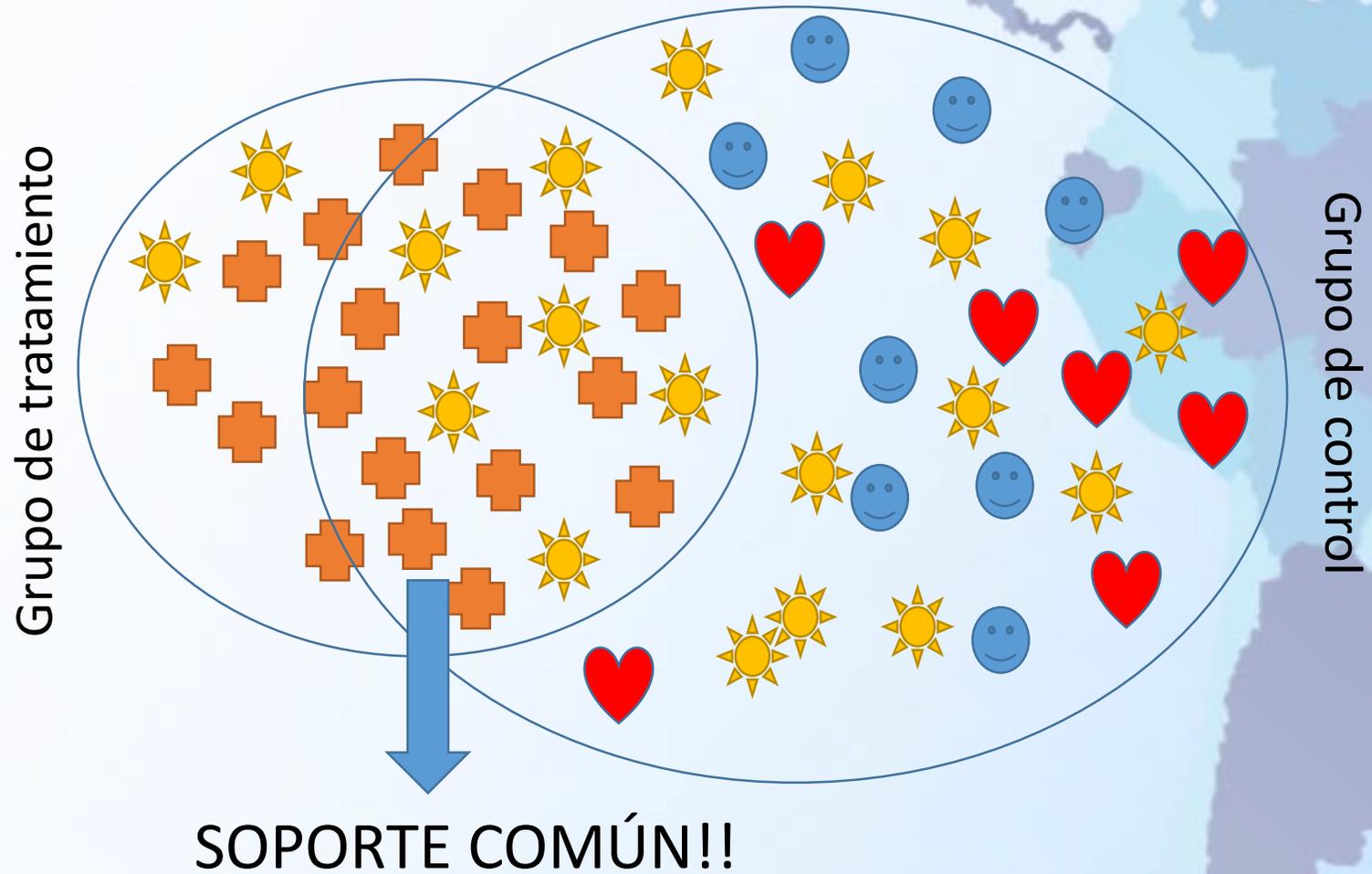
- Los supuestos son teóricos y no los podemos verificar. Lo que podemos hacer es mirar si los dos grupos son muy diferentes.
- Los resultados son confiables sólo si existen razones para pensar que las variables no observables no son fundamentales.

# Pasos para realizar una evaluación por medio de emparejamiento...

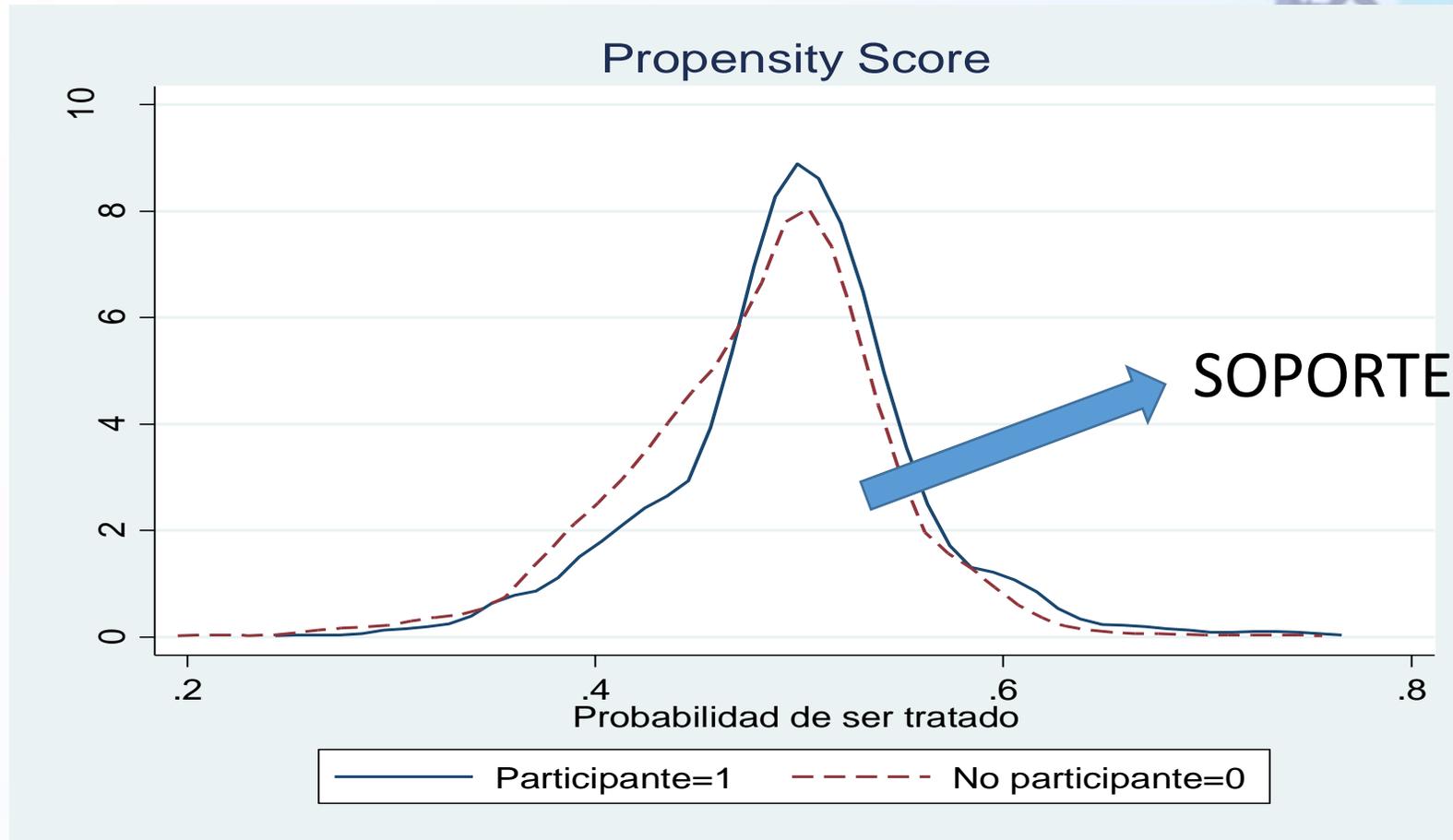


1. Definimos las variables con las cuales vamos a hacer la comparación.
  - ¿Cuáles son las variables que definen quién se gradúa de Harvard?
2. Definimos cuáles individuos vamos a comparar:
  - Submuestra de grupo de tratamiento y submuestra del grupo de control.
3. Escogemos la forma en cómo los vamos a comparar:
  - Si cada uno tiene sólo un clon, o más de un clon, o ponderación por pesos.
4. Estimamos la comparación y nos aseguramos que esté bien hecha!

# Estimación del soporte común



# En términos estadísticos...



# Diferentes métodos de emparejamiento



1. Emparejamiento por vecino más cercano.
2. Emparejamiento por 5 vecinos más cercanos.
3. Emparejamiento con y sin reemplazo.
4. Emparejamiento por ponderación de pesos.

# Ejemplos del método de emparejamiento (Attanasio y Gómez, 2004)



- **Familias en Acción:** programa de transferencias condicionadas implementado en Colombia para mejorar la acumulación de capital humano de la población vulnerable en zonas rurales.
- **Tres componentes:**
  1. Nutrición (transferencias a la madre para mejorar nutrición de niños entre 0 y 5 años de edad)
  2. Salud (vacunas y controles de crecimiento y desarrollo)
  3. Educación (transferencia condicionada a asistencia escolar)
- **Participación:** 622 municipios con los siguientes recursos: infraestructura de educación básica y salud, al menos un banco, menos de 100 mil habitantes.
- Todas las familias con niños entre 0 y 17 años de esos municipios eran **beneficiarios potenciales**.

# Ejemplos del método de emparejamiento (Attanasio y Gómez, 2004)



- **Evaluación de impacto:**

- Se escogió una muestra de 50 municipios participantes como **grupo de tratamiento**.
- **Grupo de control:** 50 municipios parecidos a los anteriores en cuanto a localización geográfica y servicios de educación y salud) pero que no fueron escogidos como beneficiarios potenciales.
- Dentro de cada municipio se escogió una muestra de hogares para hacer la evaluación.

- Como no hay aleatoriedad, los autores usan el **método de emparejamiento** para controlar por las variables que diferencian a los dos grupos.

- **Resultados:**

- No hay efectos significativos en la probabilidad de estar matriculado para niños entre 7 y 13 años, pero sí en los niños entre 14 y 17 años.
- El programa tuvo un efecto significativo sobre el peso de los niños que viven en zonas rurales de los municipios seleccionados (tanto en el peso según la edad, como en el peso según la estatura).

# Método de variables instrumentales

# ¿Cuándo usamos el método de variables instrumentales?



- Muchas veces el tratamiento no es aleatorio:
  - Cuando las personas toman la decisión de participar.
  - Cuando hay factores políticos que afectan la selección del programa.
  - Cuando la inscripción al programa es más fácil para unas personas que para otras.
- El Método de Variables Instrumentales nos sirve cuando identificamos este factor que diferencia a los dos grupos... y este factor no está relacionado directamente con la variable de resultados.

# ¿Cuáles pueden ser estos factores? Ejemplos...



Programa de política social para mejorar la nutrición de los niños



**La inscripción está abierta y se lleva a cabo en las oficinas del Ministerio de Desarrollo Social.**



Hay municipios con una oficina, hay otros con 5 oficinas.

¿Quiénes tienen más probabilidad de inscribirse?

- Las personas que viven en municipios con más oficinas.
- Las personas que viven cerca de las oficinas.
- Las madres que no trabajan (horario de las oficinas: 9am-5pm)
- Las personas mejor informadas (flyers en el super, redes sociales, etc)

# Posibles instrumentos...



- Una variable que esté relacionada con el tratamiento pero no con la variable de resultados...

1. Las personas que viven en municipios con más oficinas.
2. Las personas que viven cerca de las oficinas.
3. Las madres que no trabajan (horario de las oficinas: 9am-5pm)
4. Las personas mejor informadas (flyers en el super, redes sociales, etc)

Estas variables nos explican la participación en el programa.

¿Estas variables tienen relación con la variable de resultados (la nutrición de los niños)?

# Método de variables instrumentales...



- La estimación la realizamos en dos etapas:

1. PRIMERA ETAPA: explicamos la participación a través de la variable instrumental.

$$D_i = \beta_0 + \beta_1 Z_i + \beta_2 X_i + \varepsilon_i$$

2. SEGUNDA ETAPA: utilizamos los resultados de la primera etapa para estimar el impacto del programa.

$$Y_i = \gamma_0 + \gamma_1 \hat{D}_i + \gamma_2 X_i + u_i$$

Estimación de un efecto LOCAL porque depende sólo de una fuente de exogeneidad.

# Tipos de instrumentos encontrados en la literatura- Aleatorización imperfecta



Este grupo de personas recibe la invitación para participar en un programa.



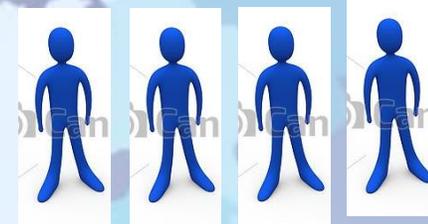
Después de recibir la invitación deciden participar.



Este grupo de personas recibe la invitación para participar en un programa.



Sin embargo no deciden participar (ej. No están tan motivadas como los verdes)



Estas personas no reciben la invitación. No alcanzan los cupos!



Por lo tanto no están invitadas a participar en el programa y no participan.

# Tipos de instrumentos encontrados en la literatura



Cuando se trata de hacer una aleatorización pero no se cumple

Instrumento: la “lotería”

Variable de participación:

1= si la persona decidió inscribirse en el programa

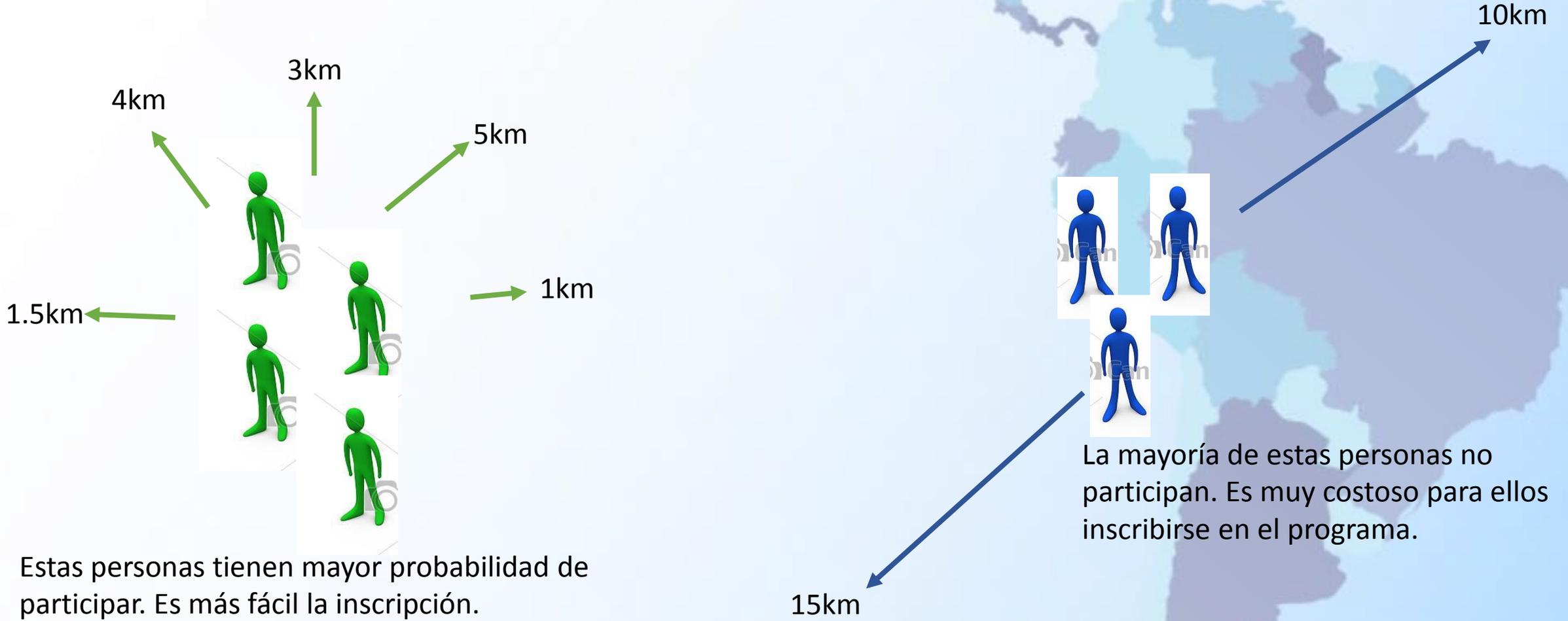
0= si la persona no se inscribió en el programa

Variable de resultados: impacto del programa (ej. Impacto en educación, nutrición, salud, etc)

Para que un instrumento sea válido debe explicar la participación pero no puede tener relación con la variable de resultados.

Estas condiciones se llaman **relevancia y exogeneidad**.

# Tipos de instrumentos encontrados en la literatura- Ubicación geográfica



Estas personas tienen mayor probabilidad de participar. Es más fácil la inscripción.

# Tipos de instrumentos encontrados en la literatura



La distancia a las oficinas de inscripción

Instrumento: distancia del hogar a la oficina de inscripción

Variable de participación:

1= si la persona decidió inscribirse en el programa

0= si la persona no se inscribió en el programa

Variable de resultados: impacto del programa

(ej. Impacto en educación, nutrición, salud, etc)

¿Por qué la distancia a las oficinas de inscripción es un instrumento válido?  
¿Podemos decir que se cumplen los supuestos de relevancia y exogeneidad?

# Problemas de usar un mal instrumento...



- El sesgo de selección es mayor si usamos un instrumento débil al que obtendríamos si estimamos el impacto usando un simple modelo de diferencias.
- No es fácil encontrar buenos instrumentos:
  - A pesar de que es fácil comprobar la “relevancia”, no es fácil comprobar la “exogeneidad”.
- Los supuestos son teóricos, por lo que no los podemos comprobar con los datos.
  - La única prueba que tenemos para probar “exogeneidad” requiere de dos instrumentos!

# Ejemplo. Hogares comunitarios de Bienestar (HCB) en Colombia (Attanasio y Vera-Hernández, 2004)



- **Pregunta de investigación:** ¿cuál es el impacto del programa HCB en indicadores antropométricos y en educación de los niños?
- **Hogares comunitarios de bienestar en Colombia:** centros de cuidado infantil para atención de niños entre 0 y 5 años de familias pobres. Los niños son atendidos por una madre de la comunidad en su propio hogar.
- **Datos:** hogares rurales residentes en municipios pequeños.
  - Es un programa que lleva más de 20 años en funcionamiento. No hay aleatorización.
  - No existen datos longitudinales por lo que no se puede realizar un dif-en-dif.

# Ejemplo. Hogares comunitarios de Bienestar (HCB) en Colombia (Attanasio y Vera-Hernández, 2004)



- **Instrumento:** distancia entre el hogar de residencia y el Hogar Comunitario más cercano.
  - La distancia afecta la participación.
  - La distancia no afecta las medidas de nutrición ni la educación.
  - Se controla también por la distancia a otros locales que prestan servicios sociales (hospitales, colegios, etc).
- **Resultados:**
  1. La distancia afecta la participación en el programa.
  2. Los hogares comunitarios tienen sobre la estatura según la edad y el peso según la edad en los niños participantes (0.4 y 0.2 desviaciones estándares por encima del promedio del grupo de control)
  3. Se incrementa la participación laboral de las madres.
  4. Se mejora la asistencia escolar de los niños entre 8 y 13 años.

# Método de regresión discontinua

# Regresión discontinua



- Observamos una discontinuidad en la participación (o en la probabilidad de participar).
- Identificamos el umbral donde vemos esta discontinuidad:
  - Los individuos a un lado del umbral participan.
  - Los individuos al otro lado del umbral no participan.
- Comparamos a los individuos a los dos lados del umbral.

# ¿Cuándo usamos este método?



Cuando el programa nos da la oportunidad de tener el grupo de comparación PERFECTO!

Programa: A las personas mayores de 70 años se les va a otorgar asistencia médica y despensa.



Población objetivo: personas mayores de 70 años con bajos ingresos económicos.

Grupo de comparación: Personas que aun no han cumplido 70 años pero que están próximos a cumplirlos!

La única diferencia es que unos no han cumplido los 70 años y los otros sí. Es CASI aleatorio...

# ¿Cuándo usamos este método?



Cuando el programa nos da la oportunidad de tener el grupo de comparación PERFECTO!

Programa: El programa está abierto a las personas que pertenecen al estatus socioeconómico más bajo (SISBEN=11)

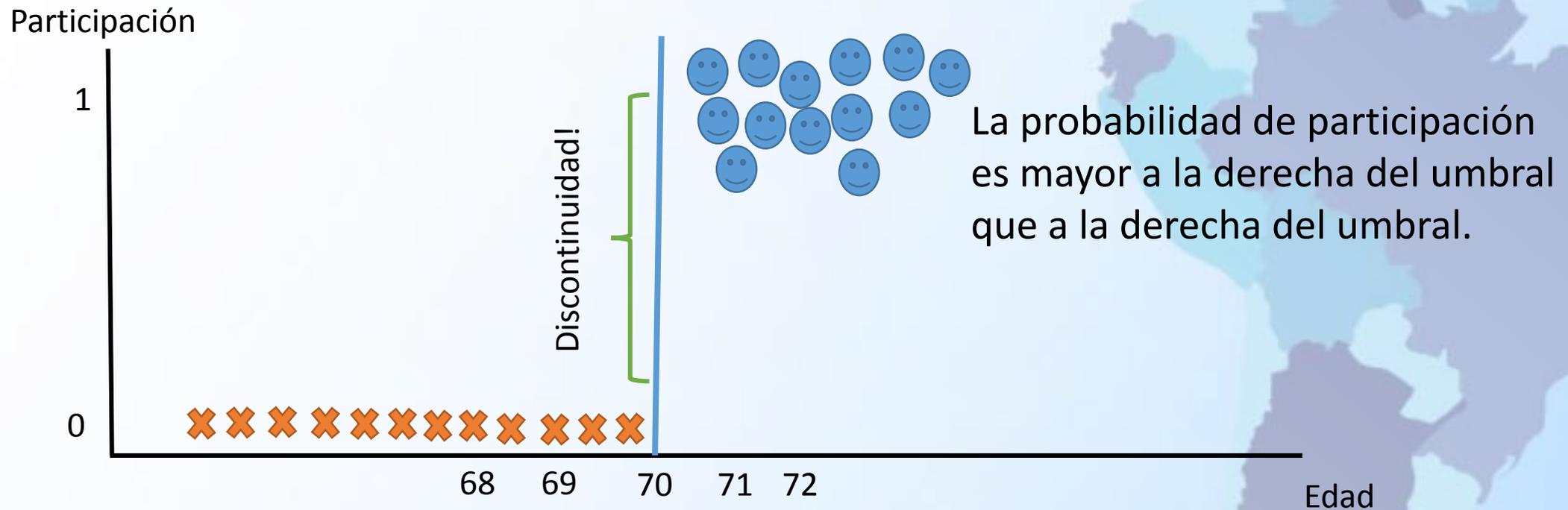


Población objetivo: personas que tienen un SISBEN menor o igual a 11.

Grupo de comparación: Personas que tienen un SISBEN ligeramente superior a 11.

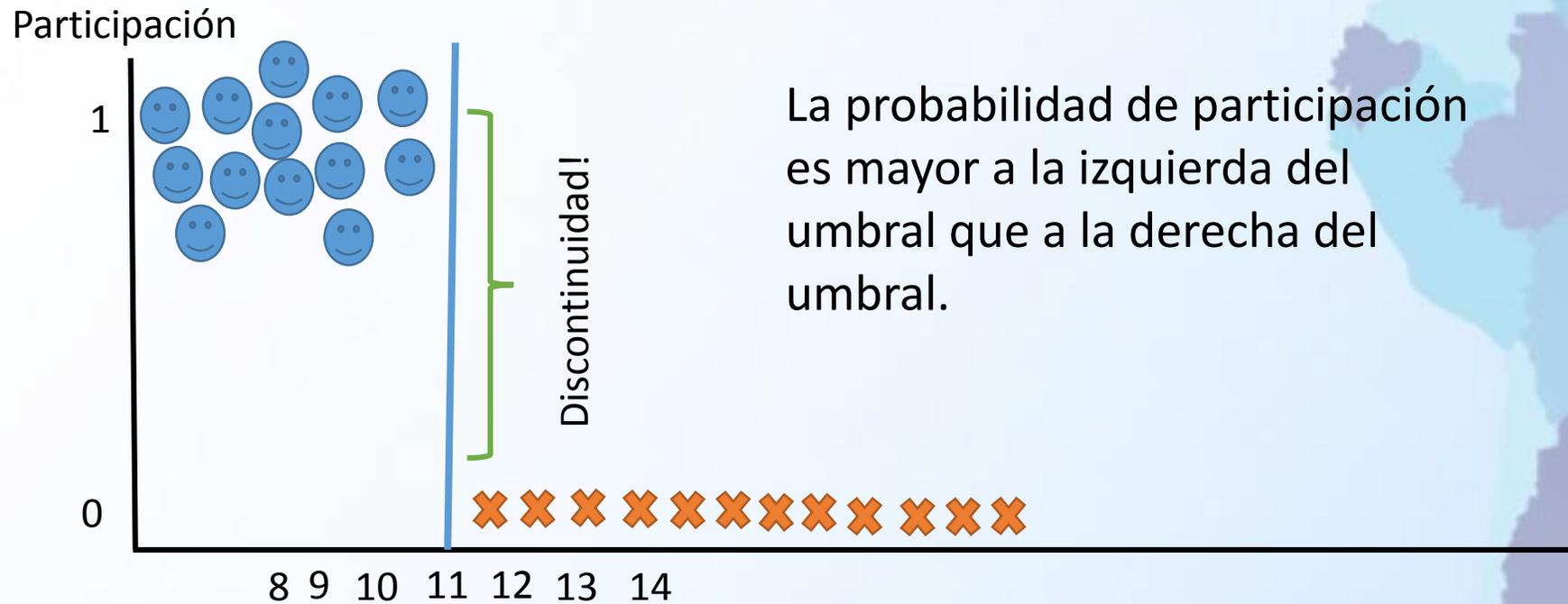
Las diferencias son mínimas entre un individuo con SISBEN=10.9 y SISBEN=11.1  
Son casi iguales!

# ¿Cuándo observamos la discontinuidad?



Comparamos los individuos a los dos lados del umbral para estimar el impacto del programa.

# ¿Cuándo observamos la discontinuidad?



Comparamos los individuos a los dos lados del umbral para estimar el impacto del programa.

# ¿Qué necesitamos?



- Sólo necesitamos un umbral:
  - Que sea independiente de la variable de resultados.
  - Que determine la participación.

Hay dos clases de umbrales

Los que determinan perfectamente la participación



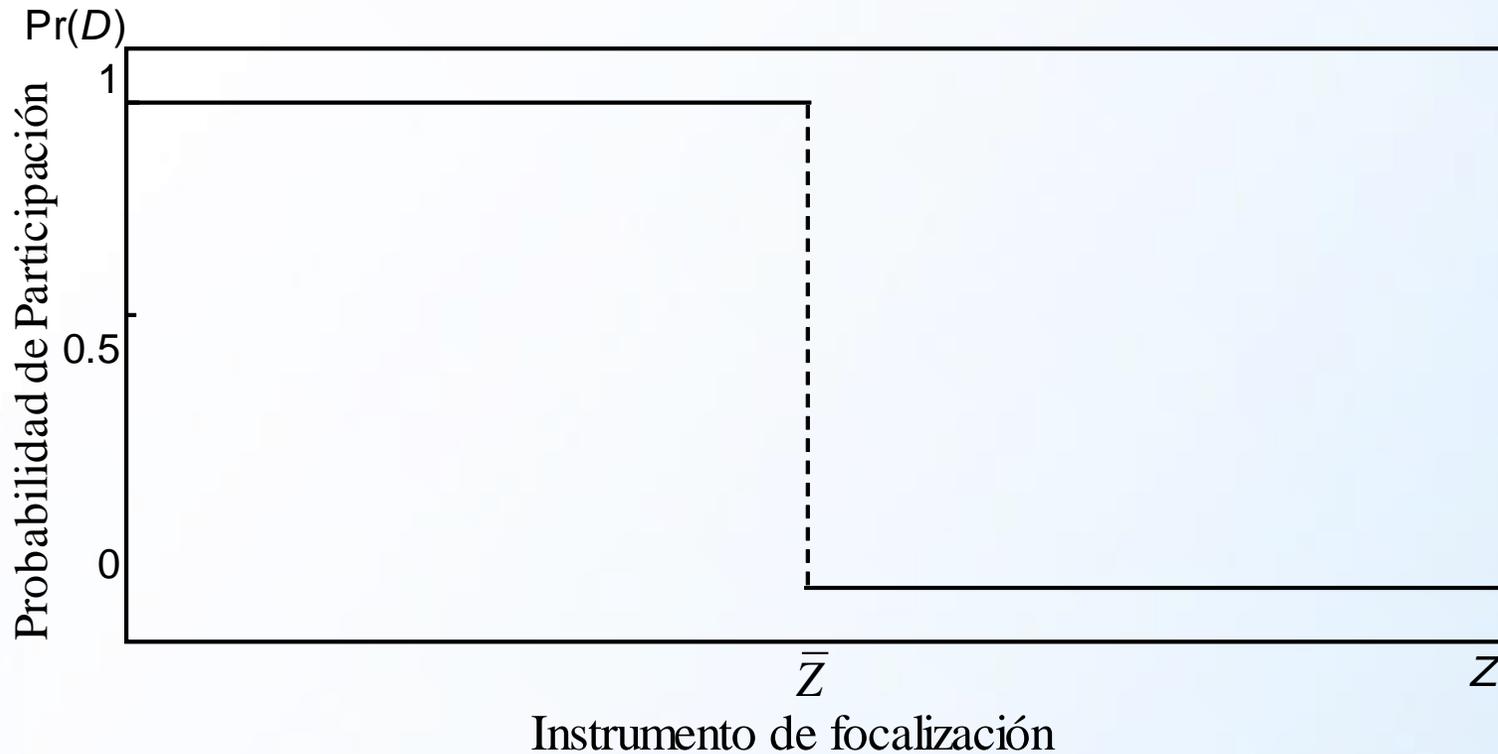
Regresión discontinua nítida

Los que determinan la *probabilidad* de participar



Regresión discontinua borrosa

# Regresión discontinua nítida (RDN)



## Familias en Acción

El umbral es igual a 11.

Las familias que tienen un  $SISBEN \leq 11$  participan en el programa.

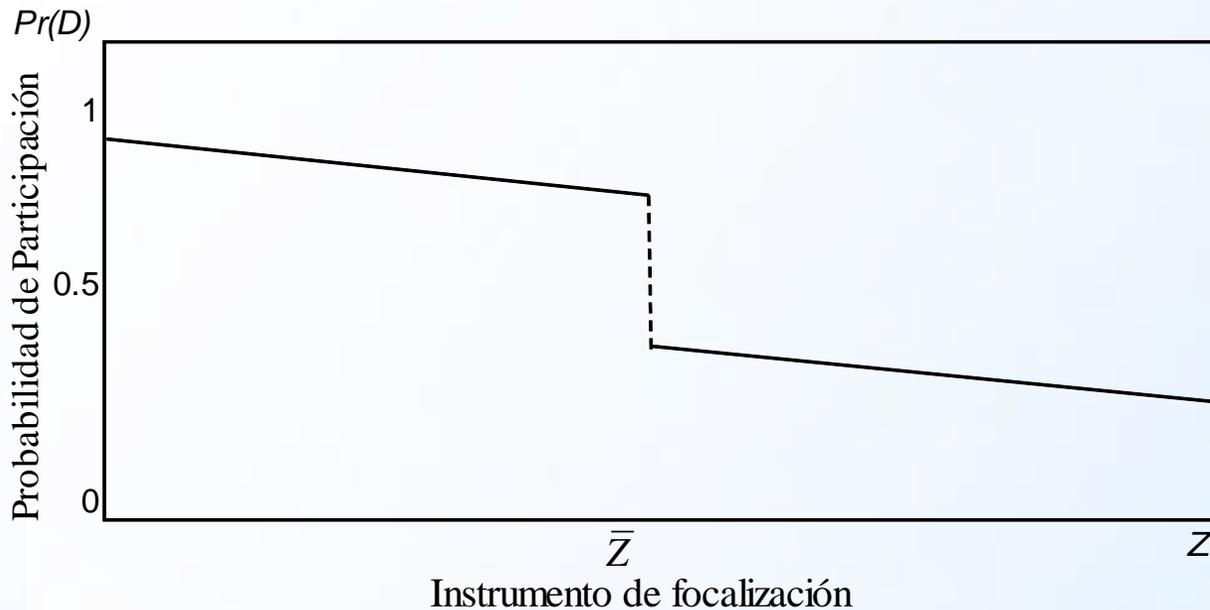
Las familias que tienen un  $SISBEN \geq 11$  no son elegibles de participación.

SISBEN: Sistema de Identificación de Posibles Beneficiarios de Programas Sociales

# Regresión discontinua borrosa (RDB)



- Se asimila más a la realidad que la RDN.
- Lo que cambia de lado a lado es la probabilidad de participación (es decir, hay individuos a los dos lados que no se comportan como el evaluador esperaría).



## Familias en Acción

El umbral es igual a 11.

Se le da prioridad a las familias que tienen un  $SISBEN \leq 11$  participan en el programa.

Si sobran recursos, se les otorgarán a las familias con un  $SISBEN \geq 11$ .

La probabilidad es positiva a los dos lados del umbral.

# Ejemplos de RDN y RDB



## REGRESIÓN DISCONTINUA NÍTIDA

- La participación es obligatoria para un segmento de la población.
- Todos los niños que asisten a escuelas específicas deben cumplir con los requerimientos del programa.
- Implementación del voto electrónico para ciudades con una población mayor a un umbral.

## REGRESIÓN DISCONTINUA BORROSA

- Participación en programas sociales focalizados en la población de más bajos ingresos.
- Sistema de pensiones para la población de adultos mayores.
- Transferencias condicionadas a asistencia escolar.

# Requisitos para la estimación de una regresión discontinua nítida



- Debe haber un salto en la participación a ambos lados del umbral.
- El umbral no puede ser modificado por los potenciales beneficiarios ni por los administradores del programa.
- Debe haber suficientes individuos a ambos lados del umbral.
- Todas las otras características de los individuos deben ser continuas a ambos lados del umbral.

# Ejemplo- Tamaño de clase (Angrist & Lavy, 1999)

- En el S.XII el académico rabino Maimonides propuso una clase de máximo 40 estudiantes para que el aprendizaje fuera efectivo. Hoy en día se usa esta regla en las escuelas públicas de Israel.
- ¿Qué pasa cuando hay más de 40 estudiantes por clase? ¿Y cuando hay 39?
- **Datos:** programa para examinar a alumnos de tercero, cuarto y quinto de primaria en 1991 y 1992.
- ¿Cuál es el impacto del tamaño de clases en el desempeño académico de los estudiantes?

# CONCLUSIONES



- Siempre es mejor una evaluación rigurosa que otro tipo de evaluación.
  - Importante seguir los pasos para lograr una evaluación rigurosa.
- Más importante aún, identificar adecuadamente las preguntas de evaluación. ¿Qué efecto queremos estimar?
  - Tener en cuenta que la evaluación tiene que ser acorde con el resto del programa o política pública.
- El evaluador debe ser capaz de escoger la mejor metodología de evaluación con base en el diseño del programa.
  - No todas las metodologías son adecuadas para todos los programas...

# ¿Cuándo utilizamos cada una de las metodologías?



Modelo de diferencias → Cuando tenemos aleatoriedad

Modelo de diferencias en diferencias → Cuando tenemos línea base

Método de emparejamiento → Cuando lo único que importa es lo que observamos

Método de variables instrumentales → Cuando la aleatoriedad no se cumple a la perfección

Método de regresión discontinua → Cuando tenemos un umbral

# EJEMPLOS DE EVALUACIONES DE POLITICA SOCIAL

8 ejemplos: diferentes  
metodologías

# Caso práctico 1: los cigarros y el peso de los niños al nacer



- El bajo peso de los niños al nacer afecta su desarrollo posterior e incrementa el riesgo de mortalidad infantil.
- A pesar de que hay muchas causas del bajo peso de los niños al nacer, una de las principales es cuando la madre fuma cigarros estando embarazada.
- El artículo analiza el efecto de una ley que incrementa el precio de los cigarros en el peso de los niños al nacer asumiendo que a un mayor precio, menos madres van a fumar estando embarazadas.
- Se utilizan datos de nacimientos de 1989 a 1992 del Natality Detail Files.

## Caso práctico 2: parásitos y asistencia escolar



- Este artículo examina el impacto de un programa en donde 65 escuelas primarias rurales de Kenia entraron en un programa de “desparasitación” de una manera aleatoria.
- Este programa es diferente a los anteriores ya que no se hizo la aleatorización a nivel individual sino a nivel de escuela. El programa fue financiado por una ONG holandesa en cooperación con el Ministerio de Salud de Busia (Kenya).
- El proyecto tuvo lugar en la parte sur de Busia, una región de granjeros, pobre y altamente poblada, localizada en el occidente de Kenya. Se escogieron 75 escuelas, casi todas escuelas primarias rurales con una población estudiantil de más de 30,000 estudiantes entre 6 y 18 años.

## Caso práctico 3: las intervenciones militares en México y el incremento de la violencia



- Muchos autores en México han demostrado que la implementación de operativos militares en el país ha incrementado la violencia.
- Sin embargo muchos de estos estudios son descriptivos o no usan buenas técnicas de evaluación de impacto.
- Uno de los problemas es que utilizar estados como unidad de análisis no nos sirve porque sólo tendríamos 32 observaciones. Y municipios tampoco porque habría contaminación entre las muestras.
- Los autores generan unas “zonas de intervención” para medir el impacto de los operativos militares en la violencia.

## Caso práctico 4: voto electrónico en Brasil y participación electoral



- Brasil tenía una de las tasas más altas de Latinoamérica de abstención electoral porque tenían un sistema muy complicado de votación.
- El problema es que el sistema hacía muy complicada la votación para la población analfabeta que normalmente es la población de bajos ingresos.
- Esto incrementaba las desigualdades en términos de oportunidades porque los pobres no estaban representados políticamente.
- En 1998 implementaron el voto electrónico en municipios con una población mayor a 40,500 votantes registrados. Este voto era mucho más fácil que la papeleta.
- Los autores encontraron que el voto efectivo incrementó, principalmente en las poblaciones con alta población analfabeta.

## Caso práctico 5: Experimento STAR en Tennessee



- **Experimento:** los estudiantes y los profesores se asignaron aleatoriamente en clases pequeñas, medianas y medianas con un profesor de medio tiempo durante los primeros cuatro años de escuela.
- **Objetivo:** analizar el efecto de haber asistido a clases pequeñas en los exámenes académicos del ACT o SAT para entrar a universidad.
- **Resultado:** El haber atendido una clase pequeña en los primeros grados de escuela está asociado con un incremento en la probabilidad de haber tomado el examen ACT o SAT para entrar a la universidad, especialmente en los estudiantes pertenecientes a minorías. También se encontró un resultado en mejoras en los resultados de estos exámenes.

## Caso práctico 6: escuela para granjeros en Perú



- CARE incrementó un programa de capacitación dirigido a granjeros de Cajamarca para incrementar la productividad de los cultivos de papa.
- La selección no fue aleatoria y cualquier podía participar, con tal de que asistiera a las clases de capacitación.
- La mayoría de los individuos inscritos ya habían recibido capacitación de CARE. Sin embargo el programa se amplió para otras comunidades en donde no habían tenido capacitaciones anteriormente.
- se utilizó una encuesta que se le hizo a los cultivadores de papa en 1999 en 13 comunidades de la provincia de San Miguel.

# Caso práctico 7: experimento de violencia doméstica en Minneapolis



- La mayoría de los victimarios en casos de violencia doméstica reinciden en el acto. Muchas veces porque la víctima decide retirar la demanda después de un tiempo.
- Los investigadores de este artículo están interesados en saber si las políticas de mano dura (i.e. enviarlos a la cárcel) disminuye la reincidencia.
- Para esto hicieron un experimento en el cual una computadora le dictaba a la policía qué debía hacer con el victimario en casos de violencia doméstica. Las opciones era enviarlo a la cárcel o alguna política más “suave”.
- El experimento no se cumplió a la perfección porque la policía podía decidir enviarlo a la cárcel siempre que el victimario se comportara de manera agresiva o si estaba bajo los efectos del alcohol.