

# Desarrollo Económico y Política Pública

Módulo 4

Clase 3

**Dario Tortarolo**

University of Nottingham

## Calendario de clases virtuales

1. ~~Lunes 6 de Junio de 10 am a 12 pm~~
2. ~~Miércoles 8 de Junio de 10 am a 12 pm~~
3. **Miércoles 15 de Junio de 10 am a 12 pm**
4. Lunes 20 de Junio de 10 am a 12 pm

## Novedades

Ya se encuentran disponibles en la página del curso ([Click acá](#)):

- ▶ Slides de la clase 2
- ▶ Ejercicio de la clase 2
- ▶ Video de la clase 2

- ▶ Nos interesa medir la relación entre un **resultado**  $Y$  y una **política**  $P$  de forma **causal**

## Por ejemplo

- ▶ ¿Los programas de formación profesional aumentan los ingresos de los participantes?
- ▶ ¿La mejora de los caminos aumenta el acceso a los mercados laborales e incrementa el ingreso de los hogares y, en caso afirmativo, en qué medida?
- ▶ ¿Las reducciones de cargas sociales aumentan la demanda de empleo y las horas/esfuerzo laboral?
- ▶ ¿Las fiscalizaciones a contribuyentes ayudan a reducir la evasión?  
¿Cuán persistentes son los efectos?

# Métodos de Inferencia Causal

## Evaluaciones rigurosas de políticas públicas:

- ▶ Procuran medir el efecto de un programa comparando los resultados de un **grupo de tratamiento** (T) con la estimación de un **contrafactual** obtenido de un **grupo de comparación válido** (C)
- ▶ Existen diversos métodos para estimar el **efecto causal** de las políticas:
  1. Asignación Aleatoria
  2. Variables Instrumentales
  3. Diseño de Regresión Discontinua
  4. Diferencias en Diferencias y Event Studies
  5. Control Sintético

- ▶ La clase pasada discutimos acerca de:
  1. Asignación Aleatoria
  2. Variables Instrumentales
  
- ▶ El plan para hoy es:
  1. Variables Instrumentales (LATE)
  2. Diseño de Regresión Discontinua

**¿Preguntas?**

# Diseño de Regresión Discontinua

(aka RDD)



## Umbrales y discontinuidades

- ▶ Muchas políticas a menudo utilizan **umbrales** para decidir quién será afectado por un programa y quién no, si el gasto o los impuestos son mas altos o bajos, etc
- ▶ Por ej., becas/admisiones escolares basadas en puntajes superiores a un umbral
- ▶ El **diseño de regresión discontinua** (RDD) es un método econométrico que permite estimar el efecto causal de una política al comparar unidades *elegibles* y *no elegibles* a un lado y al otro de cierto umbral pre-determinado

## Diseño de Regresión Discontinua (RDD) - condiciones

Para aplicar un DRD, deben cumplirse las siguientes **condiciones**:

1. Elegibilidad definida en función de una variable o índice continuo (ej., puntaje en pruebas estandarizadas, edad, ingresos pre-determinados, etc.)
2. Un umbral por debajo o por encima del cual se clasifica a la población como elegible para el programa (ej., mujeres mayores de 60 años elegibles para una jubilación)
3. El umbral debe ser único para el programa de interés. No debería haber otras políticas que utilicen el mismo umbral
4. La variable o índice que determina la elegibilidad no puede ser manipulada por los beneficiarios potenciales, los administradores del programa o los políticos

## Diseño de Regresión Discontinua (RDD) - estimación

El RDD estima el efecto de una política alrededor del umbral límite de elegibilidad:

1. Calcular el resultado promedio de unidades del lado tratado de la variable límite de elegibilidad (T)
2. Calcular el resultado promedio de unidades en el lado no tratado (comparación) de la variable límite (C)
3. Efecto causal: la diferencia entre el resultado promedio a la izquierda y a la derecha del umbral

- ▶ Pero, **¿Por qué podemos afirmar que el efecto estimado es causal?**

- ▶ Pero, **¿Por qué podemos afirmar que el efecto estimado es causal?**
- ▶ La clave, es que en un entorno muy cercano al umbral, más allá de la política, las unidades tienen atributos similares y por lo tanto son comparables
- ▶ Es decir, nos da un grupo de comparación válido; el contrafactual de “que hubiese pasado si la persona afectada por la política hubiese quedado del lado *no elegible*”
- ▶ ¿A qué les recuerda?

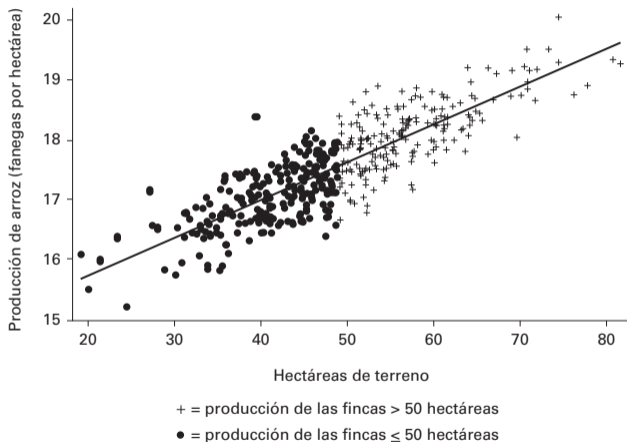
- ▶ Pero, **¿Por qué podemos afirmar que el efecto estimado es causal?**
- ▶ La clave, es que en un entorno muy cercano al umbral, más allá de la política, las unidades tienen atributos similares y por lo tanto son comparables
- ▶ Es decir, nos da un grupo de comparación válido; el contrafactual de “que hubiese pasado si la persona afectada por la política hubiese quedado del lado *no elegible*”
- ▶ ¿A qué les recuerda? Estadísticamente, en el límite, es una suerte de **Asignación Aleatoria**. Es más, el efecto estimado es *local* a las unidades cerca del umbral

## Ejemplo: Subsidio para fertilizantes

- ▶ Supongan un programa agrícola cuyo fin es aumentar la producción de arroz subsidiando la compra de fertilizantes
- ▶ El programa está destinado a fincas con 50 hectáreas o menos
- ▶ ¿Cuál es la variable de elegibilidad?
- ▶ ¿Cuál es el umbral límite?

## Ejemplo: Subsidio para fertilizantes

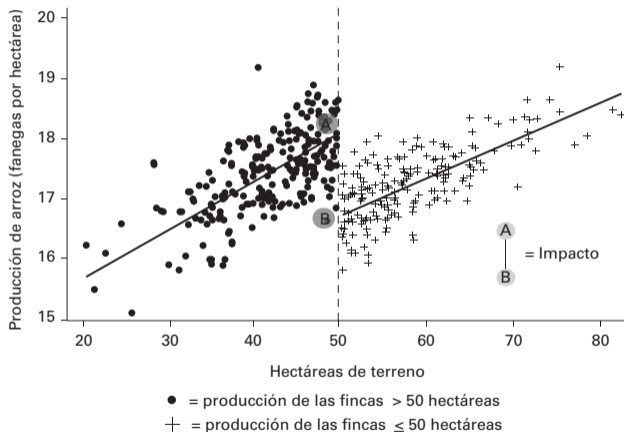
- Situación **pre-subsidio**
- Las fincas de 48, 49, 50 ha son elegibles. Las fincas con 50.1, 50.2, 51 ha no
- Es probable que las fincas con 49.9 ha sean muy similares a las de 50.1 ha en todos los aspectos, salvo que un grupo recibió el subsidio para fertilizantes y el otro no





## Ejemplo: Subsidio para fertilizantes

- Situación **post-subsidio**
- Efecto RDD = producción de arroz promedio en punto A – punto B
- Otros factores (clima, shocks de precios, otras políticas), se cancelan en esa resta



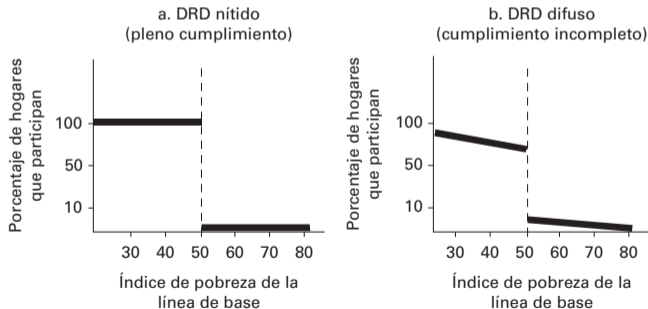
## Ejemplo: Subsidio para fertilizantes

### Validez interna y validez externa

- ▶ El método RDD es muy potente y creíble
- ▶ En la jerga se dice que goza de una elevada *validez interna*
- ▶ Sin embargo, la *validez externa* (o extrapolación) es limitada
- ▶ Esto se debe a que el efecto está estimado usando fincas en un entorno de 50 ha
- ▶ El efecto causal es válido “a nivel local”. De esta manera, se obtiene una estimación de un *efecto local promedio del tratamiento* (LATE)
- ▶ En este ejemplo, poco podemos decir del efecto de fertilizantes en fincas de 10 o 20 ha

## RDD nítido (sharp) vs difuso (fuzzy)

- ▶ **Nítido:** Cuando todas las unidades cumplen con la asignación que les corresponde según la variable de elegibilidad
- ▶ **Difuso:** Cuando hay incumplimiento. Algunas unidades elegibles deciden no participar, o unidades no elegibles encuentran un modo de participar



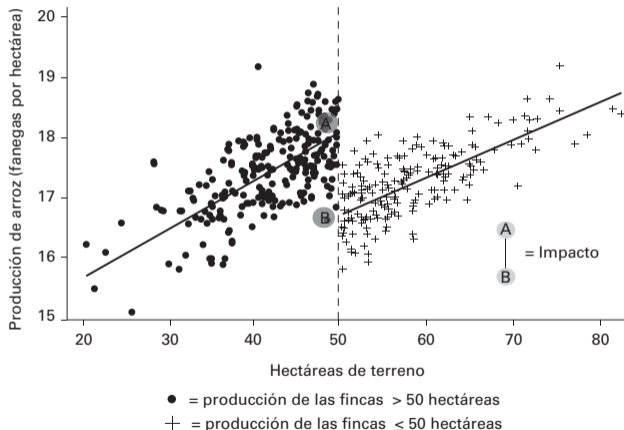
## RDD difuso + Variables Instrumentales

- ▶ Si el RDD es difuso, se puede usar el enfoque de **variables instrumentales** para “corregir” el incumplimiento
- ▶ Recuerden la clase 2: asignación aleatoria con incumplimiento. Utilizamos la asignación aleatoria como instrumento
- ▶ En RDD podemos usar la asignación original basada en la variable de elegibilidad como instrumento (esta sería la **1a etapa** de VI)
- ▶ Volvamos al ejemplo del subsidio para fertilizantes

# RDD difuso + Variables Instrumentales

## Ejemplo: Subsidio para fertilizantes

- Situación post-subsidio
- 1a etapa: cómo cambia la probabilidad de solicitar el subsidio en el umbral
- ITT = producción de arroz promedio en A – B
- 2a etapa (LATE) = Efecto RDD = ITT / 1a etapa



## Verificación de la validez del RDD

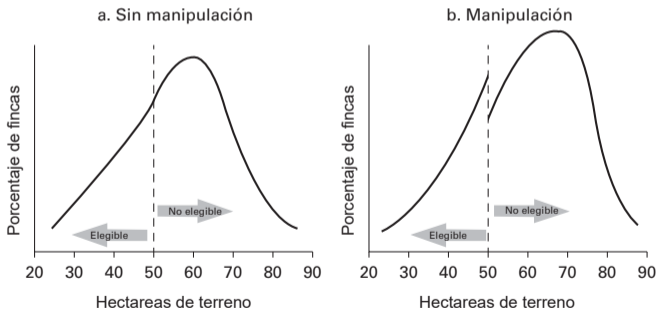
En la práctica, el método de RDD requiere pasar dos tests:

1. **No manipulación de la variable de elegibilidad:** es crucial que en la cercanía del umbral los individuos no puedan cambiar su condición de tratamiento o control (ej., agricultores con 50.3 ha venden 0.5 ha)
2. **Balance de atributos:** cerca del umbral, las unidades deberían ser muy similares, en promedio, en características observadas y no observadas. Esto significa que la *distribución* de esas variables debe ser continua alrededor del umbral

# Test de no manipulación

Distribución de las unidades según la variable que determina elegibilidad

- ✓ La densidad en torno al umbral es continua (panel a)
- ✗ La densidad en torno al umbral presenta “bunching” (panel b)  
i.e., un número mayor de fincas se agrupa justo por debajo del umbral



## Limitaciones del RDD

- ▶ RDD proporciona estimaciones del LATE en torno al umbral donde las unidades de T y C son más similares
- ▶ No es posible generalizar ese efecto a unidades alejadas del umbral
- ▶ Pero, ¿Es esto bueno o malo?



## Limitaciones del RDD

- ▶ RDD proporciona estimaciones del LATE en torno al umbral donde las unidades de T y C son más similares
- ▶ No es posible generalizar ese efecto a unidades alejadas del umbral
- ▶ Pero, ¿Es esto bueno o malo? Depende de la pregunta de interés. Por ej:

## Limitaciones del RDD

- ▶ RDD proporciona estimaciones del LATE en torno al umbral donde las unidades de T y C son más similares
- ▶ No es posible generalizar ese efecto a unidades alejadas del umbral
- ▶ Pero, ¿Es esto bueno o malo? Depende de la pregunta de interés. Por ej:
  1. **¿El programa debería existir o no?** El efecto de tratamiento promedio (ATE) para toda la población elegible puede ser el parámetro más relevante
  2. **¿El programa debería suspenderse o ampliarse?** El RDD produce precisamente la estimación local de interés para sustentar esa decisión

## Otras consideraciones prácticas

- ▶ Potencia estadística: intensivo en el número de observaciones alrededor del umbral
- ▶ Ancho de banda (lo más amplio posible, pero que conserve el balance en las características observadas)  
Práctica: mostrar si el efecto es sensible o no al ancho utilizado
- ▶ Forma funcional (la relación entre variables  $Y$  y  $X$  puede ser no lineal)  
Práctica: probar con diversas formas funcionales (lineales, cuadráticas, cúbicas, etc)
- ▶ Múltiples dimensiones: algunas políticas utilizan más de una variable de asignación con diferentes umbrales
- ▶ Se puede combinar con otros métodos (por ej., diff-in-diff)

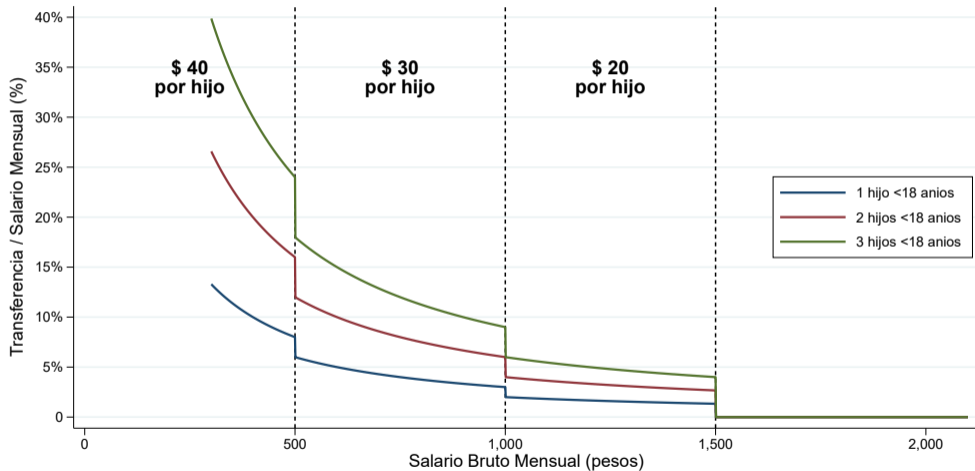
## Ejercicio: Asignaciones Familiares (AAFF)

- ▶ Consideren el programa AAFF para trabajadores formales en Argentina
- ▶ Este subsistema otorga una transferencia mensual por cada hijo menor de 18 años
- ▶ El monto por hijo decrece a medida que aumenta el salario mensual del solicitante
- ▶ Por ejemplo, consideren el esquema vigente entre los años 1996 a 2004

1. ¿Qué tipo de discontinuidad(es) provee este programa?
2. ¿Existe espacio para manipulación?
3. ¿Qué preguntas podrían estudiar con esta variación?

# Asignaciones Familiares 1996–2004

Fuente: Garriga y Tortarolo (2022)



## Ejercicio de la Clase 1

El curso actual cuenta con 35 estudiantes. Supongan que solo 15 asisten a todas las clases y hacen las tareas y lecturas (T). Por razones desconocidas, los restantes 20 no asisten a ninguna (C). Luego del examen final, observan lo siguiente:

1. Las notas promedio son mayores para el grupo T respecto a C. ¡El curso es buenísimo!
2. Las notas promedio son menores para el grupo T respecto a C. ¡El curso es malísimo!

Evalúen estas afirmaciones. ¿Se puede dar una interpretación causal?  
¿Por qué si o por qué no?

[CLICK ACÁ PARA RESPONDER](#)

## REFERENCIAS

Evaluación de Impacto en la Práctica (<http://www.worldbank.org/ieinpractice>)

Gertler, P. J., Martinez, S., Premand, P., Rawlings, L. B., Vermeersch, C. (2017). *Impact Evaluation in Practice*, Second Edition. Washington, DC: Inter-American Development Bank and World Bank ([web](#))