

Problem Set 10: Regression discontinuity

Fecha de entrega: 22 de noviembre de 2024

Paola Llamas y Tomás Pacheco

Instrucciones

- Agreguen los nombres y apellidos de los integrantes en el encabezado del archivo que entregan.
- El título de los archivos entregados debe tener la siguiente estructura:
PS10_Garcia_Gomez_Gonzalez_Perez.pdf
- **UN ÚNICO** miembro del grupo debe subir el PDF y *do-file* al Campus Virtual en la sección "Problem Set 10".
- El código lo tendrán que entregar de dos formas distintas. Primero deberán exportar el *do-file* como PDF y unirlo al informe ([ayuda](#)). Luego, deberán entregar el código en formato .do en el lugar indicado en el Campus Virtual.
- El directorio de trabajo debe tener tres carpetas: programs, input y output. El *do-file* debe ser entregado de tal manera que los tutores podamos correr el código **solamente** modificando el global con el *main directory* (**esto es muy importante**).
- Recuerden que siempre deben presentar la información en tablas y gráficos estéticos. Nunca presenten una tabla o figura sin dedicarle unas líneas a comentar lo que se observa.

En Israel, las escuelas tienden a seguir la "Maimonides rule", es decir, intentan no tener clases con más de 40 estudiantes. Esto significa que, cuando más de 40 estudiantes se inscriben en una escuela, se dividen en 2 aulas, mientras que una inscripción menor a 40 llevará a una sola aula. Esto genera una discontinuidad en el tamaño de la clase con una inscripción igual a 40. Disponemos de una base de datos con información sobre calificaciones y otras variables para niños de 5° grado en Israel. Nuestra pregunta principal será el efecto del tamaño de la clase en las calificaciones. Las variables de la base son:

- `enroll`: Inscripción total de la cohorte
 - `classsize`: Tamaño de la clase
 - `avgmath`: Nota promedio en matemática
 - `avgverb`: Nota promedio en lengua
 - `tip_a`: Porcentaje de estudiantes desfavorecidos
- (a) Corra dos regresiones OLS simples que relacionen el tamaño de la clase (`classsize`) con las calificaciones en matemática y lengua. Interprete el signo, la magnitud y la significancia de los coeficientes. ¿Se cumplen los supuestos de OLS?
- (b) Corra las mismas regresiones, agregando los controles disponibles (porcentaje de estudiantes desfavorecidos e inscripción). [Nota: La inscripción es el número total de niños inscritos en ese grado en toda la escuela, y, por lo tanto, es un proxy del tamaño de la escuela y/o de la cohorte]. ¿Cambian mucho los coeficientes de interés?
- (c) Podemos estimar el efecto de interés utilizando regression discontinuity por la regla de Maimónides. Explique por qué. Explique si corresponde a un diseño sharp o fuzzy de acuerdo con los datos observados. **Escriba todas las especificaciones econométricas relevantes.**
- (d) Antes de realizar el análisis de RDD, hay varias pruebas de validez que puede realizar para validar el enfoque. Explique cuáles son, realícelas y reporte los resultados. Escriba y argumente detalladamente si piensa que en este contexto podría haber manipulación de la variable `enroll`.
- (e) Realice el análisis de RDD para ambas calificaciones, matemáticas y lenguaje, utilizando el enfoque de *continuity-based approach*. Para ambos resultados, muestre el gráfico de RD utilizando un ajuste de polinomio de grado 1, 2, 3 y 4. ¿Cuál le parece más adecuado?

- (f) Con el polinomio elegido, corra nuevamente el RDD, seleccionando como *bandwidth* 10, 15 y 7. ¿Qué resultados espera ver? ¿Qué trade-off hay en esta elección?
- (g) (Basado en Angrist (2009)): Suponiendo que se agrega una nueva clase cada vez que la inscripción excede un múltiplo entero de 40, los autores generan la siguiente variable de *predicted class size* previsto para una escuela con inscripción e_s

$$z_s = \frac{e_s}{\text{floor}((e_s - 1)/40) + 1}$$

donde $\text{floor}(x)$ es el mayor entero menor o igual a x . Por ejemplo, $\text{floor}(10,354) = 10$ y $\text{floor}(3) = 3$. Angrist y Lavy (1999) usan z_s como una variable instrumental para *classsize*. Ellos llaman a z_s la “Regla de Maimónides”. Realice esta estimación y muestre como cambian los resultados.