Semana 8. Endogeneidad y Regresión Discontinua Borrosa (RDB)

Equipo Econometría Avanzada

Universidad de los Andes

30 de septiembre de 2022

Contenido

Endogeneidad

En estimaciones de MCO la endogeneidad puede ser causada por:

1. Simultaneidad o doble causalidad.
    - Agotamiento laboral de maestras y desarrollo cognitivo de niños.
    - Número de policías en un municipio y tasa de crimen.
    - **Dirección del sesgo:** depende de la magnitud y dirección de los efectos (de X sobre Y y de Y sobre X) (Base, D. 2015).
2. Variable omitida relevante.
    - y: Desarrollo cognitivo, x: Asistencia efectiva al centro de cuidado, x^*: Conocimiento sobre desarrollo temprano.
    - y: Desempeño en la clase, x: Horas de estudio, x^*: Motivación/afinidad.
    - **Dirección del sesgo:** la dirección de la interacción entre (i) el efecto $VO \rightarrow y$, y (ii) la covarianza entre la VO y x.

Endogeneidad

3. Error de medición en la variable independiente.
   - Violencia.
   - Uso de métodos anticonceptivos.
   - **Dirección del sesgo:** Si se supone error de medición clásico (aleatorio y x exógena) $Plim\left(\widehat{\beta}\right) < \beta$.
4. Sesgo de (auto) selección.
   - y salario, x educación, selección en habilidades.
   - y salario, solo se observa para quienes trabajan.
   - **Dirección del sesgo:** depende del contexto y la forma de modelarlo (ver última semana de clase).

- ¿El estimador de MCO es inconsistente cuando hay error de medición clásico en la y?
- ¿Las pruebas de hipótesis de MCO son inválidas cuando hay error de medición clásico en la y?

Contenido

Contexto y pregunta

Fernanda Leite y Renata Rizzi (2014) utilizan datos a nivel individuo de una encuesta que aplicaron en Brasil el día después de las elecciones presidenciales del 2010, para evaluar la siguiente hipótesis:

- **Hipótesis de la ignorancia racional** (Downs, 1957): en elecciones grandes los individuos reconocen que su voto tiene una probabilidad considerablemente baja de ser decisivo. Por tanto, sin importar cuán interesados estén en los resultados de las elecciones, no van a incurrir en costos asociados a informar su voto.

- El voto es obligatorio entre los 18-69 años, y voluntario entre los 16-17.

Pregunta de investigación

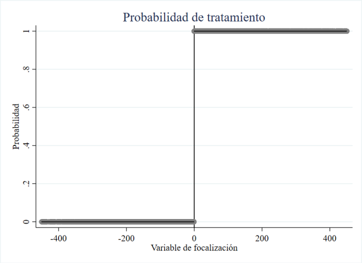
¿Cuál es el efecto de votar sobre el conocimiento político?

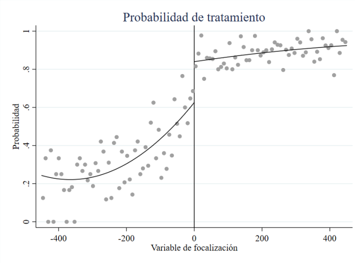
$$test_i = \beta_0 + \beta_1 vota_i + u_i$$

- ¿A qué personas se compara en esta regresión?
- ¿Sería $\widehat{\beta}_1^{OLS}$ un estimador consistente del efecto de interés?.

Regresión Discontinua Borrosa

En este caso, la variable de asignación no resulta en un salto determinístico en la probabilidad de tratamiento de 0 a 1 (o 1 a 0), pero sí en un salto discontinuo en esta.





- Outcome: $test_i$
- Tratamiento: $vota_i$
- Elegibilidad: $voto\_obligatorio_i$
- Focalización: $edad_i$

Sean X_i la variable de focalización, $\bar{x}$ el corte, $T_i = \mathbb{1}\{X_i \geq \bar{x}\}$ la asignación al tratamiento y $D_i = D_i(1)T_i + D_i(0)(1 - T_i)$ el estatus de tratamiento efectivo.

En esta situación, podemos analizar distintos efectos:

1. El efecto de la **asignación** al tratamiento sobre la variable de tratamiento:

$$\begin{aligned} \tau_{PE} &\equiv \lim_{x \downarrow \bar{x}} \mathbb{E}[D_i | X_i = x] - \lim_{x \uparrow \bar{x}} \mathbb{E}[D_i | X_i = x] \\ &= \mathbb{E}[D_i(1) - D_i(0) | X_i = \bar{x}] \end{aligned} \tag{1}$$

2. El efecto de la **asignación** al tratamiento sobre la variable de resultado (ITT):

$$\begin{aligned} \tau_{ITT} &\equiv \lim_{x \downarrow \bar{x}} \mathbb{E}[Y_i | X_i = x] - \lim_{x \uparrow \bar{x}} \mathbb{E}[Y_i | X_i = x] \\ &= \mathbb{E}[(D_i(1) - D_i(0))(Y_i(1) - Y_i(0)) | X_i = \bar{x}] \end{aligned} \tag{2}$$

Estos dos efectos pueden ser recuperados por Regresión Discontinua Nítida.

- ¿Qué pasa si $D_i(1) - D_i(0) = 1$?

No obstante, lo que nos interesa al final es el efecto del **tratamiento efectivo** sobre la variable de resultado.

Este efecto podemos recuperarlo a través de Regresión Discontinua Borrosa si se cumplen los siguientes supuestos:

1. **Monotonicidad Local:** $D_i(x+\varepsilon) \geq D_i(x)$ (o al revés) con $\varepsilon > 0$ para todo i cuando $x \to \bar{x}$.
2. **Continuidad local:** $\lim_{x \downarrow \bar{x}} \mathbb{E}[Y_i(d)|X_i = x] = \lim_{x \uparrow \bar{x}} \mathbb{E}[Y_i(d)|X_i = x]$ con $d \in \{0, 1\}$
3. **Discontinuidad en la probabilidad de tratamiento:**

$$\lim_{x \downarrow \bar{x}} \mathbb{P}(D_i = 1|X_i = x) \neq \lim_{x \uparrow \bar{x}} \mathbb{P}(D_i = 1|X_i = x)$$

Bajo los anteriores tres supuestos, el efecto de interés está dado por

$$\tau_{RDB} \equiv \frac{\lim_{x\downarrow\bar{x}} \mathbb{E}[Y_i|X_i = x] - \lim_{x\uparrow\bar{x}} \mathbb{E}[Y_i|X_i = x]}{\lim_{x\downarrow\bar{x}} \mathbb{E}[D_i|X_i = x] - \lim_{x\uparrow\bar{x}} \mathbb{E}[D_i|X_i = x]}$$

$$= \mathbb{E}[Y_i(1) - Y_i(0)|X_i = \bar{x}, D_i(1) > D_i(0)] \quad (3)$$

Es decir, es un efecto local (para $X_i = \bar{x}$) LATE (para los i tales que $D_i(1) > D_i(0)$, los *compliers*).

Nótese que que el efecto está dado por la razón entre τ_{ITT} y τ_{PE}.
¡Regresión Discontinua Borrosa es IV LATE!

- ¿Cómo se comparan los supuestos de RDB con los de IV-LATE?

Estimación paramétrica

Primera etapa:

$$vota_{ic} = \alpha_0 + \alpha_1 voto\_obligatorio_{ic} + f(edad\_ajust_{ic}) + \theta_c + x_{ic}\alpha_2 + e_{ic}$$

Segunda etapa:

$$test_{ic} = \gamma_0 + \gamma_1 \widehat{vota}_{ic} + g(edad\_ajust_{ic}) + \theta_c + x_{ic}\gamma_2 + u_{ic}$$

- ¿Qué ecuación permite estimar el ITT?

Estimación no paramétrica: Polinomios Locales

Primera etapa:

Para un determinado kernel $K(\cdot)$, un ancho de banda h y un polinomio de grado j, estimar a la derecha del corte:

$$\min_{\hat{\mu}_+,\hat{\mu}_{+,j}} \sum_{i=1}^{n} K\left(\frac{X_i-\bar{x}}{h}\right)\left[D_i-\hat{\mu}_+-\sum_{j=1}^{p}\hat{\mu}_{+,j}(X_i-\bar{x})^j\right]^2 \mathbb{1}[X_i\in(\bar{x},\bar{x}+h)]$$

y a la izquierda:

$$\min_{\hat{\mu}_-,\hat{\mu}_{-,j}} \sum_{i=1}^{n} K\left(\frac{X_i-\bar{x}}{h}\right)\left[D_i-\hat{\mu}_--\sum_{j=1}^{p}\hat{\mu}_{-,j}(X_i-\bar{x})^j\right]^2 \mathbb{1}[X_i\in(\bar{x}-h,\bar{x})]$$

y obtener

$$\hat{\tau}_{PE}=\hat{\mu}_+-\hat{\mu}_-$$

Segunda etapa:

Para los mismos $K(\cdot)$, h y j estimar a la derecha del corte:

$$\min_{\hat{\delta}_+, \hat{\delta}_{+,j}} \sum_{i=1}^{n} K\left(\frac{X_i - \bar{x}}{h}\right)\left[y_i - \hat{\delta}_+ - \sum_{j=1}^{p} \hat{\delta}_{+,j}(X_i - \bar{x})^j\right]^2 \mathbb{1}[X_i \in (\bar{x}, \bar{x} + h)]$$

y a la izquierda:

$$\min_{\hat{\delta}_-, \hat{\delta}_{-,j}} \sum_{i=1}^{n} K\left(\frac{X_i - \bar{x}}{h}\right)\left[y_i - \hat{\delta}_- - \sum_{j=1}^{p} \hat{\delta}_{-,j}(X_i - \bar{x})^j\right]^2 \mathbb{1}[X_i \in (\bar{x} - h, \bar{x})]$$

y obtener

$$\hat{\tau}_{ITT} = \hat{\delta}_+ - \hat{\delta}_-$$

A partir de $\hat{\tau}_{PE}$ y $\hat{\tau}_{ITT}$ obtener $\hat{\tau}_{RDB} = \frac{\hat{\tau}_{ITT}}{\hat{\tau}_{PE}}$. Nuevamente, este debe ser corregido por el sesgo que presenta.

Conclusión

- RDB utiliza una **discontinuidad no-determinística** en la probabilidad del tratamiento.
- RDB es un IV en un vecindario cercano al punto de corte.
- RDB instrumenta localmente el tratamiento efectivo con la asignación al tratamiento.
- RDB brinda un estimador del LATE que es **local en dos dimensiones.**
  - Para los individuos en la vecindad cercana al punto de corte.
  - Para los *compliers*.
- Los supuestos de continuidad local y discontinuidad en la probabilidad de tratamiento pueden ser verificados empíricamente.
  - El supuesto de monotonicidad no.

Gracias!