Pontificia Universidad Católica de Chile

Facultad de Ciencias Sociales - Instituto de Sociología

Inferencia Causal (SOL3063)

Profesor: Luis Maldonado

Ayudantes: Camila Ortiz



**Ejercicio 3: Variable instrumental y mecanismos**

El ejercicio debe ser resuelto usando esta pauta de preguntas. La entrega debe hacerse el lunes 3 de julio en el mail [causalidadsol3063@gmail.com](mailto:causalidadsol3063@gmail.com) antes de las 19:00 hrs. El atraso en las entregas causará un descuento de 0,5 puntos por día (hábil y no hábil) de atraso, desde el primer minuto del día y hora de entrega. El nombre del archivo debe comenzar con el apellido del/a estudiante. Se puede trabajar en pareja.

Este ejercicio busca poner en práctica los contenidos y habilidades asociadas a los contenidos de variable instrumental y mecanismos explicativos. Para desarrollar este ejercicio, replicaremos algunos de los resultados de *Bautista et al. (2023). The Geography of Repression and Opposition to Autocracy. American Journal of Political Science, 67(1),* p101-118.

La base de datos es FinalDatasetForReplication.dtadisponible en Canvas. Para efectos de este ejercicio considere las siguientes variables (para más detalles sobre la medición de las variables ver Supporting Info, disponible en Canvas).

**Variable dependiente**

Rate of voter registration: Share\_reg70\_w2.

**Variables independientes**

Victims/10,000 inh: shVictims\_70.

Ln distance closest military base: LnDistMilitaryBase.

Province ID: IDProv.

Vote share Salvador Allende in 1970: share\_allende70.

Ln distance to regional capital: lnDistRegCapital.

Pop. Share rural 1970: sh\_rural\_70.

Population in 1970 (in 100,000s): Pop70\_pthousands.

**Variables para administrar los datos**

Main estimation sample: MainSample.

Population in 1970: Pop70.

**Actividades**

1. En base al modelo (2) de la Table 3 de Bautista et al. (2023), realice las siguientes actividades:
   1. Formalice (exprese en ecuaciones) los modelos del panel A y B. En su formalización, represente las variables de control con una X. **Puntaje: 0,5.**
   2. Represente su ecuación para el modelo del panel B con un DAG. **Puntaje: 0,5.**
   3. Mencione y explique 2 supuestos representados en su DAG que permitirían estimar el efecto del tratamiento en cuestión sin sesgo. **Puntaje: 0,5.**
2. Ahora vamos a hacer estimaciones de los modelos formalizados en el punto 1. Al respecto:

* En base a FinalDatasetForReplication.dta, realice un filtro de modo de trabajar con los datos solo si MainSample = 1.
* Los modelos estimados deben usar la variable Pop70 como ponderador.
* Las covariables son share\_allende70, lnDistRegCapital, sh\_rural\_70, y Pop70\_pthousands.
* Use efectos fijos para provincias (IDProv) en todos sus modelos.

Con esta información, haremos una serie de estimaciones. Presente sus estimaciones en una tabla de calidad como la Table 3 del paper bajo análisis. No es necesario que ésta presente ni Kleibergen Paap F-stat ni Avg. Dependent variable. Específicamente:

* 1. Estime su formalización para el panel A del modelo 2 en la Table 3 de Bautista et al. (2023) en dos versiones: 1) controlando por todas las covariables, salvo Pop70\_pthousands; 2) controlando por todas las covariables. **Puntaje: 0,5.**
  2. Respecto de las estimaciones en a), interprete el coeficiente de LnDistMilitaryBase y su significación estadística para las versiones 1) y 2). Además, discuta estos resultados. **Puntaje: 0,5.**
  3. Estime su formalización para el panel B del modelo 2 en la Table 3 de Bautista et al. (2023) en dos versiones: 1) controlando por todas las covariables, salvo Pop70\_pthousands; 2) controlando por todas las covariables. **Puntaje: 0,5.**
  4. Respecto de las estimaciones en c), interprete el coeficiente de shVictims\_70 y su significación estadística para las versiones 1) y 2) en c). Además, discuta estos resultados. **Puntaje: 0,5.**
  5. Respecto del modelo estimado en a) y c), explique en términos sustantivos: 1) en qué se distinguen los de modelos estimados en a) respecto de los modelos en c); y 2) quienes serían los compliers. **Puntaje: 0,5.**

1. Autores sugieren que controlar por diferencias regionales produce sesgo post-tratamiento. Los modelos arriba estimados controlaron por efectos fijos para provincias. Al respecto, realice las siguientes actividades:
   1. En base a lectura de Pepinski et al. (2023), explique desde un punto de vista teórico bajo qué supuestos los efectos fijos para provincias en sus modelos en el punto 2) no generan sesgo en la estimación del efecto del tratamiento bajo análisis. **Puntaje: 0,5.**
   2. Tal como vimos en clase, podemos estimar efectos directos controlados en presencia de variables post-tratamiento. Al respecto,
      1. En base a su DAG del modelo del panel B y asumiendo a LnDistMilitaryBase como tratamiento, señale en cuál de las siguientes categorías clasifican las variables de su DAG:
         1. Variable dependiente o resultado.
         2. Tratamiento.
         3. Mediador.
         4. Confounders pre-tratamiento.
         5. Confounders post-tratamiento intermedios.

**Puntaje: 0,5.**

* + 1. En base a su clasificación en punto i) y asumiendo que los efectos fijos capturan algún confounder post-tratamiento intermedio, estime el efecto controlado directo del tratamiento *LnDistMilitaryBase* e interprete el coeficiente estimado y su significación estadística. Para su estimación, centre el mediador en su media y use el ponderador utilizado en sus estimaciones previas. **Puntaje: 0,5.**

**Referencias relevantes para resolver el ejercicio 3**

* Clases 12b, 13, 14, y 15.
* Artículos sustantivos
  + Bautista, M. A., González, F., Martínez, L. R., Muñoz, P., & Prem, M. (2023). The geography of repression and opposition to autocracy. American Journal of Political Science, 67(1), 101-118.
  + Pepinsky, T. B., Goodman, S. W., & Ziller, C. (2023). Modeling spatial heterogeneity and historical persistence: Nazi concentration camps and contemporary intolerance. American Political Science Review, 1-10.
* IV 1: idea general
  + Angrist, J. D. y J. Pischke (2015). Mastering ’Metrics: The Path from Cause to Effect. Princeton University Press. Capítulo 3 para IV.
  + Wooldridge, Jeffrey M. (2009). Introducción a la econometría: un enfoque moderno. 4a edición. México: Cengage Learning. Cap. 15.
* IV 2: LATE/CACE y concepto de compliers
  + Gerber, A. S. y D. P. Green (2012). Field Experiments: Design, Analysis, and Interpretation. New York: W. W. Norton. Caps. 5.
  + Sovey, A. J. y D. P. Green (2011). Instrumental Variable Estimation in Political Science: A Readers’ Guide. American Journal of Political Science, Vol.55, N.1:188-2000.
* Efectos directos controlados
  + Acharya, A., Blackwell, M., and Sen, M. (2016). Explaining causal findings without bias: Detecting and assessing direct effects. American Political Science Review, 110(3), 512-529.
  + VignetteACDE.pdf

IMPORTANTE: SE DEBE ENTREGAR SOLO UN ARCHIVO (WORD O PDF) QUE CONTENGA LAS RESPUESTAS Y EL CÓDIGO DE ANÁLISIS COMPLETO EN R. NO OLVIDE PRESENTAR SUS RESULTADOS EN TABLAS Y FIGURAS DE CALIDAD Y REPORTAR SU CODIGO DE R EN EL APENDICE. EL NOMBRE DEL ARCHIVO ENTREGADO DEBE COMENZAR CON EL APELLIDO DEL/A ESTUDIANTE.