

## Problem Set 4: Variables instrumentales

---

**Fecha de entrega:** 16 de septiembre de 2022

Gastón García Zavaleta 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:  
PS4\_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 4".
- 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.

Este trabajo está basado en el paper “Following the Poppy trail: Origins and consequences of Mexican drug cartels” (Murphy y Rossi, 2020)<sup>1</sup>. En este artículo usan la migración china a principios del siglo XX como una instrumento de la presencia de un cartel en el presente con el objetivo de estudiar si los carteles mejoraron los *outcomes* socioeconómicos. Lean el *paper* antes de comenzar a responder las preguntas. Utilicen la base *poppy.dta*. Importante: en el artículo se detallan todas las variables que utilizan en cada estimación. Dada la confidencialidad de los datos, no todas las variables están disponibles. Para resolver este trabajo utilicen solo las disponibles y tengan en cuenta que los resultados no van a ser idénticos a los del *paper*.

1. Generen la variable “Chinese presence” tal como lo especifica el *paper*.
2. Hagan estadística descriptiva de las variables relevantes (como en la Tabla 1) con todas las variables que tengan. De ahora en más, no utilicen a Distrito Federal para hacer las estimaciones.
3. Repliquen las regresiones hechas por OLS de las columnas 3 a 6 de la Tabla 5. Todas estas regresiones se estiman *clustereando* los errores (lo vamos a ver más adelante). En vez de utilizar la opción *robust*, corran `cluster(id_estado)`.
4. Reproduzcan la Tabla 7 utilizando la presencia del cartel en 2010 como variable instrumentada. Nota: vamos a llamar a la segunda columna de esta tabla como la “especificación básica”. Además, repliquen la Tabla 8. Expliquen intuitivamente qué muestran los resultados.
5. ¿Pueden testear la exogeneidad del instrumento? Si supieran que el instrumento es exógeno, ¿podrían testear la exogeneidad de la presencia del cartel? Expliquen y lleven a cabo el test. Utilicen la especificación básica.
6. Tomen la especificación básica y agregen la variable “german presence” como instrumento. Lleven a cabo un test de sobreidentificación (Sargan o J-test). ¿A qué conclusión llegan?
7. Critiquen el *paper* en no más de tres párrafos (focalicen su crítica en la exogeneidad).

---

<sup>1</sup>Lean la versión publicada.