



PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE
DEPARTAMENTO DE INGENIERÍA INDUSTRIAL Y DE SISTEMAS
ICS2563 - ECONOMETRÍA APLICADA
SECCIÓN 1

Tarea 4

9 de mayo de 2022

2022-1 - Profesor Patricio Domínguez

Aspectos generales

- La tarea puede ser desarrollada en forma individual o en parejas (2 estudiantes).
- La fecha de entrega de la tarea es el **19 de Mayo a las 23:59 hrs** en el portal del curso en Canvas.
- La entrega debe incluir el script de R (o el dofile de STATA) y un informe de análisis en formato pdf.
- Sobre el **informe**:
 - El informe debe incorporar las respuestas a todas las preguntas, incorporando las figuras y/o tablas que estime conveniente.
- Sobre el **código**:
 - Cada entrega debe incluir el script de R (o el dofile de STATA) desde donde se corre todo el análisis.
 - Es importante que el código esté adecuadamente comentado para facilitar su corrección. Por ejemplo, indicar qué parte del informe/preguntas se desarrollan en cada sección del código.
 - **Reproducibilidad:** El código debiera estar escrito de manera tal que cualquier persona lo pueda correr, y reproducir los resultados desde su computador. En el caso de crear datos aleatorios recomendamos fijar una semilla que permita reproducir los resultados tal cual queden especificados en el informe de reporte.
- Parte de la tarea es que se vean enfrentados a tomar decisiones. Si deben aplicar criterio, háganlo y justifiquen sus elecciones.

- Ante dudas o preguntas, se recomienda fuertemente la utilización del foro de preguntas del curso. Las respuestas pueden servir a otros compañeros, que se enfrentan a las mismas dudas. No se permite publicar respuestas.
- La tarea tiene un total de **85 puntos**, en donde la presentación del informe posee 5 puntos (El informe debe incorporar las respuestas a todas las preguntas, y las figuras y/o tablas que estime necesarias), y la presentación del script 5 puntos (todos los cálculos realizados).
- **Bonificación:** Como un incentivo al uso del procesador de texto LATEX, se entregarán 5 puntos adicionales a quienes lo usen para escribir su informe.

Descripción de la tarea

En esta tarea usted deberá desplegar sus conocimientos en evaluación de programas y su capacidad de analista utilizando sus conocimientos de Econometría Aplicada. Usted deberá desarrollar distintas tareas orientadas a entender la conveniencia de diferentes diseños para preguntas puntuales, pensando como siempre en los problemas de sesgo y las condiciones que se debieran cumplir para entender una relación observada como causal.

Preliminar

Utilice la BDD **earthquake_T4.dta**, la cual corresponde a datos de sección cruzada agrupados (pooled cross-section). La BDD contiene información de individuos de 3 ciudades en tres momentos del tiempo. Las ciudades están ubicadas en el norte, centro y sur del país. En la discusión y justificación de sus respuestas aproveche de explorar la base de datos y las relaciones de las variables disponibles.

La información de la BDD describe las características de los habitantes de tres ciudades en tres momentos del tiempo. Usted deberá pensar cómo hacer uso de los datos disponibles para abordar de la mejor manera las preguntas propuestas. Recuerde que puede haber más de una respuesta correcta en algunas preguntas y se evaluará especialmente su capacidad de análisis al respecto.

Fíjese que la BDD tiene información de las siguientes variables asociadas a diferentes individuos en distintos momentos del tiempo:

- **city:** ciudad de residencia.
- **post:** tiempo 0 si **post=0**, 1 si **post=1**, 2 si **post=2**.
- **female:** 1 si el individuo es mujer, 0 si no.
- **immig:** 1 si es extranjero, 0 si no.
- **age:** edad.

- `educ`: años de educación formal completados por cada individuo.
- `exp`: experiencia laboral, construida como `age-educ-6`.
- `exp2`: experiencia laboral al cuadrado.
- `log_income`: ingresos promedio por hora recibido en el último mes.

Para algunas preguntas específicas usted también podría utilizar la variable `id` (identificador individual). Note que al utilizar la variable `id`, usted cuenta con un panel a nivel individual con información en tres momentos del tiempo. Utilice `log_income` como la principal variable a analizar (*outcome*) e interprete sus resultados según corresponda.

Preguntas

1. Descripción de los datos (5 puntos)

Por medio de una tabla de balance, compare las características de los individuos de las diferentes ciudades. ¿Qué puede decir respecto de los salarios en cada una de estas ciudades?

Hint: Para realizar la comparación es conveniente que fije un momento del tiempo. Usted debe escoger cuál.

2. Ecuación de Mincer (10 puntos)

- a) Por medio de una regresión lineal de Mincer¹ muestre la relación entre años de educación (escolaridad) y salarios. Estime al menos tres modelos distintos, incorporando en cada uno las variables de control que crea necesario. Discuta sus resultados y compare los coeficientes obtenidos para la variable escolaridad.
- b) ¿Cómo cambia su estimación si utiliza datos de sección cruzada agrupados (pooled cross-section)? Discuta, y a partir de sus resultados, refiérase a la comparabilidad de los datos entre ciudades y a lo largo del tiempo.

3. Impacto del Terremoto (15 puntos)

Imagine que ocurrió un terremoto que afectó a la ciudad 1 (`city=1`), ubicada al centro del país, entre $t=1$ y $t=2$. Este terremoto no solo causó la destrucción total o parcial de viviendas y daños en infraestructura, sino que también afectó fuertemente la capacidad productiva y los empleos de los habitantes.

Estime el impacto del terremoto en los ingresos de las personas en $t=2$. Discuta la conveniencia de utilizar la información de individuos de todas las ciudades disponibles o solo un subconjunto

¹En honor a Jacob Mincer, un famoso economista laboral que estudió la relación entre educación y salarios utilizando datos de sección cruzada, esto es una muestra de individuos **en un momento del tiempo**.

de ellas, así como el rol de los controles individuales y variables específicas asociadas al período de análisis.

4. Efecto en los migrantes (20 puntos)

Imagine que rápidamente luego del terremoto, el gobierno implementó un plan de asistencia económica a las familias damnificadas. Por dificultades técnicas y de focalización se decidió asignar el beneficio a personas chilenas excluyendo así a los extranjeros. Ello ha desatado una polémica respecto del impacto diferenciado que ha generado el terremoto y la débil asistencia del gobierno a los migrantes.

A usted se le encomienda evaluar el efecto diferenciado sobre los migrantes en $t=2$. Escriba cómo intentaría medir ese efecto y realice su mejor estimación. Discuta.

5. Habilidad (10 puntos)

Para esta pregunta considere que por medio de un convenio con el Ministerio de Desarrollo Social usted consigue los identificadores individuales de las personas encuestadas en cada momento del tiempo (id). A su vez, usted ha leído un artículo que señala que la habilidad es un determinante importante no sólo en el nivel educacional alcanzado por las personas sino además en los ingresos que obtienen en el trabajo. Una clave adicional sobre la noción de habilidad es que sería un atributo relativamente fijo en el tiempo para cada individuo.

- a) Explique por qué la omisión de habilidad podría inducir un sesgo en su estimación de la relación entre educación e ingresos. Discuta la dirección del sesgo.
- b) Explique cómo podría usar los datos de panel individual para resolver o eliminar el sesgo discutido en 5.a.

6. *Movers* (15 puntos)

Imagine que un colaborador suyo le indica que seguir recolectando muestras de individuos que sean representativas de cada ciudad y hacerlo en forma estable en el tiempo resulta demasiado caro. A cambio, le plantea que, en vez de recolectar la extensa base de datos, usted podría conseguir datos administrativos de personas que se hayan movido de una ciudad en el período observado. La ventaja es que cada persona al ejecutar el traslado de ciudad queda consignada como tal y resulta muy fácil conseguir un historial laboral de ellos similar al de los datos recolectados en el panel de esta tarea.

Para examinar la conveniencia de este cambio, usted podría identificar a las personas que se han cambiado de ciudad en el tiempo y analizar si esa (sub)muestra sería conveniente para el análisis desarrollado en esta tarea. Realice las estimaciones pertinentes y discuta la conveniencia de usar esos datos para estimar el impacto del terremoto. ¿Por qué cree que los resultados se parecen/difieren de sus estimaciones preferidas en 3? Justifique su análisis mostrando evidencia que caracterice a la población de estudio.