

Tarea 1 (Puntos totales 100)

Instrucciones generales

- Todos los pasos necesarios para los ejercicios se deberán hacer directamente en Stata, ningún cálculo o limpieza de información se hará en otro software.
- Se debe entregar lo siguiente:
 - El archivo log con el nombre y clave única, es decir “JonDoe_109244_Tarea1RES.log”
 - El do file con el siguiente nombre: “JonDoe_109244_Tarea1COD.do”
- Los insumos para el ejercicio se encuentran en la carpeta “Tarea1” del drive.
- Los archivos se deben enviar por correo electrónico, con asunto: “JonDoe_109244_Tarea1”
- **La fecha de entrega es antes del 2 de septiembre de 2019 a las 11:59pm. Después de esta fecha y hora no se aceptarán los trabajos.**
- El puntaje de cada pregunta se encuentra explícito en el documento.
- **Código de ética: los alumnos tienen prohibido utilizar el mismo código o compartir respuestas, cualquier plagio se penalizará de acuerdo con los lineamientos institucionales del ITAM.**
- **IMPORTANTE:** se debe comentar el código mostrando las secciones y los pasos explícitos de los distintos ejercicios.
- Las respuestas a ciertas preguntas y hallazgos las deberán poner como *Comentarios* en el do file de Stata.

Ejercicios: EL OBJETIVO GENERAL ES DETERMINAR COMO CIERTAS CLAÚSULAS LEGALES AFECTAN EL SPREAD EN LOS BONOS SOBERANOS. En este caso, la variable dependiente de interés es spread, por lo cual el análisis descriptivo y conclusiones se deben anclar a la variable spread.

1. (15 PUNTOS) Abrir la base de datos “Bonos.xlsx” en Stata, posteriormente guardar el archivo como “Bonos.dta”
 - a. Explorar los datos. **¿Qué variables tienen el formato incorrecto (i.e. son string cuando deberían ser numéricas)?** Poner la respuesta comentada en el código. NOTA: en los archivos de Excel las celdas vacías son *missing data*, revisar que una vez que se cargue en Stata, los valores en celdas vacías tienen “.” y no un valor de cero.
 - b. Modificar el formato de las variables para que tengan el formato correcto.
 - c. Mantener en la base de datos únicamente las siguientes variables:
 - i. year: año en el cual un bono se emitió.
 - ii. spread: spread entre el bono soberano y la tasa libre de riesgo.

- iii. aaa: indicadora si el bono es rating en el S&P AAA.
 - iv. aa: indicadora si el bono es rating en el S&P AA.
 - v. a: indicadora si el bono es rating en el S&P A.
 - vi. bb: indicadora si el bono es rating en el S&P BB.
 - vii. b: indicadora si el bono es rating en el S&P B.
 - viii. num_lawyers2: # de firmas de abogados en el contrato.
 - ix. rep_ic_size: indicadora sobre el tamaño de la emisión.
 - x. curr: moneda en la que fue emitido el bono.
 - xi. Law: ley para promover juicios en caso de default
 - xii. rep_ic_top2: indicadora sobre el tamaño de la emisión.
 - xiii. rep_ic_top2_ny: indicadora sobre el tamaño de la emisión de bonos bajo NY law.
 - xiv. rep_ic_top2_eng: indicadora sobre el tamaño de la emisión de bonos bajo English law.
 - xv. lnamt: logaritmo del monto de la emisión
 - xvi. years: años para la madurez del bono.
 - xvii. highrate2: indicadora si el bono es de alta calidad.
 - xviii. hh_index_ib: índice de aglomeración.
 - xix. hh_index_ib_eng: índice aglomeración de bonos bajo English law.
 - xx. hh_index_ib_ny: índice aglomeración de bonos bajo NY law.
- d. Mantener las observaciones que son posteriores (incluyendo) al año 1947 (variable year).
- e. Realizar un análisis descriptivo y **resumir los principales hallazgos.**
2. (15 PUNTOS) Realizar un *merge* entre “Bonos.dta” y “Bonos2.xlsx” para tener la información en una única base de datos. NOTA: en los archivos de Excel las celdas vacías son *missing data*, revisar que una vez que se cargue en Stata, los valores en celdas vacías tienen “.” y no un valor de cero. Las variables son indicadores económicos a nivel nacional por año.
- a. Explorar los datos. ¿Qué variables tienen el formato incorrecto (i.e. son string cuando deberían ser numéricas)? Poner la respuesta comentada en el código.
 - b. Modificar el formato de las variables para que tengan el formato correcto.
 - c. Realizar un análisis descriptivo y **resumir los principales hallazgos de las nuevas variables.**
3. (10 PUNTOS) A través de gráficos (a elección, al menos 3 distintos tipos) realizar un análisis descriptivo, que nos permita determinar relaciones. **Contestar: ¿Cuáles son las principales conclusiones?**
4. (25 PUNTOS) Regresión simple:
- a. Generar las siguientes variables:

- i. num_lawyers_2: es 1 si num_lawyers2==2 y es 0 cuando num_lawyers2==1 (el resto de los casos es cero. HINT: cuidado con las variables omitidas).
 - ii. num_lawyers_0: es 1 si num_lawyers2==0 y es 0 cuando num_lawyers2==1 o num_lawyers2=2 (el resto de los casos es cero. HINT: cuidado con las variables omitidas).
 - iii. Reemplazar rep_ic_size con cero si num_lawyers2=0.
 - b. Resultados: correr las siguientes regresiones
 - i. Variable dependiente: spread; Variables independientes: aaa aa a bb b num_lawyers_2 num_lawyers_0
 - ii. Variable dependiente: spread; Variables independientes: num_lawyers_2 num_lawyers_0
 - c. Tests: realizar las siguientes pruebas:
 - i. Linealidad
 - ii. Homocedasticidad
 - iii. Variables omitidas
 - iv. Especificación
 - v. Multicolinealidad
 - vi. Outliers
 - vii. Normalidad en los errores
 - d. Describir lo siguiente (se debe incorporar la significancia y distintos aspectos vistos en clase):
 - i. Resultados intuitivos de las regresiones
 - ii. Resultados intuitivos de las pruebas
- 5. (25 PUNTOS) Regresión con cambios: Replicar el ejercicio 4 pero añadir lo siguiente y explicar los principales cambios (se debe incorporar la significancia y distintos aspectos vistos en clase):
 - a. Añadir errores estándar robustos y variables dummy para *year* y *curr*
 - b. Considerando el punto (a), correr la regresión para dos submuestras: law==NY y law==English, respectivamente.
 - c. Explicar los principales cambios.
- 6. (10 PUNTOS). Añadir información: Replicar 5a, incluyendo las variables extras de archivo Bonos2 (previamente incluidas en el ejercicio 2) como regresores. **Explicar la principal conclusión al añadir estas variables.**
- 7. (10 PUNTOS ADICIONALES) ADICIONAL (opcional): Correr un Probit con Variable dependiente: rep_ibc_top2; Variables independientes: lnamt years highrate2 hh_index_ib num_lawyers2 (incluir dummies para la variable *year*)

- a. Generar la variable λ con la siguiente forma: $\text{normalden}(xb)/\text{normal}(xb)$ (HINT: ver *predict*).
- b. Replicar 5a con la variable λ como regresor, explicar la intuición de incluir esta variable y los cambios en los resultados.