

**Taller #4**  
**Econometría 06169**

**Profesor: Julio César Alonso**

**Notas:**

- Recuerde que sólo dos preguntas, seleccionadas al azar, serán calificadas.
- Este taller es para ser entregado en los primeros 10 minutos de la clase del próximo 16 de Febrero.

**INSTRUCCIONES:**

- Este taller debe ser escrito en computador y entregado en papel.
- La presentación de los resultados debe obedecer a los formatos estudiados en clase.
- Cuando sea posible, debe mostrar el procedimiento efectuado para llegar a sus resultados.

1. Un investigador desea determinar si las transacciones con tarjetas de crédito o débito son afectadas por la nueva medida del Gobierno de devolver el equivalente a 2ptos porcentuales del IVA siempre y cuando las transacciones se realicen con tarjetas crédito o débito. Para esto, el investigador recurre a dos grupos de individuos con iguales características socioeconómicas. El primer grupo es enterado de la medida, mientras que el segundo grupo no tiene ni idea de la medida del gobierno. Para todos los individuos de ambos grupos se recoge la información de las transacciones realizadas con tarjetas de débito y crédito en el mes de Enero (en miles de pesos) (La información se encuentra en el archivo D\_T4.xls)<sup>1</sup>
  - a) Escriba un modelo que permita determinar si existe o no un efecto sobre las transacciones promedio con tarjetas de crédito o débito de la nueva medida. (demuestre que su modelo si sirve para este efecto).
  - b) Estime el modelo y reporte los resultados en una tabla.
  - c) Interprete los coeficientes estimados.
  - d) Determine si de acuerdo a los resultados obtenidos la medida del gobierno afecta las transacciones con estos medios de pago.

Un consultor es contratado para determinar la relación que existe entre el gasto público ( $G$ ) y la Remuneración de los Asalariados residentes en la nación por parte de los empleadores Residentes y no residentes en la nación ( $RAR$ ), la Remuneración a los asalariados residentes en la nación por parte de los empleadores No Residentes en la nación ( $RNR$ ) y la Remuneración a los asalariados Residentes y no residentes en la nación por parte de los empleadores Residentes en la nación ( $RRR$ ). (Todas las variables están medidas en millones de dólares)

<sup>1</sup> Estos datos no son reales.

2. Empleando los datos que se encuentran en el archivo D\_T4.xls, responda las siguientes preguntas:
  - a) Escriba e interprete a priori los coeficientes del modelo.
  - b) Estime el modelo y reporte sus resultados en una tabla.
  - c) Determine la significancia individual y conjunta de los coeficientes estimados.
  - d) ¿Qué tan bueno es el "fit" del modelo?
3. Continuando con el ejercicio anterior, determine si existe o no multicolinealidad. (muestre todo su trabajo y conclusiones)
4. Corrija el problema de multicolinealidad. Explique claramente su decisión y reporte el modelo estimado que usted considere adecuado.
5. Continuando con el ejercicio anterior,
  - a) Constate que el nuevo modelo no posee problemas de multicolinealidad.
  - b) Compare la varianza de los estimadores encontrados en el punto 2 b) y 4. Intuitivamente explique porque ocurre lo que observa.
6. Continuando con el ejercicio anterior, conteste:
  - a) Interprete los coeficientes estimados y discuta la significancia individual y conjunta de los coeficientes.
  - b) Determine si es cierto que el efecto sobre el gasto público de la remuneración de los asalariados residentes en la nación por parte de los empleadores residentes y no residentes en la nación es 10 veces más pequeño y contrario al de remuneración a los asalariados residentes en la nación por parte de los empleadores no residentes en la nación.

**Taller #4**  
**Respuestas Sugeridas**  
**Econometría 06169**

**Profesor: Julio César Alonso**

**Monitora: Natalia Angulo**

**Notas:**

- o Recuerde que sólo dos preguntas, seleccionadas al azar, serán calificadas.
- o Este taller es para ser entregado en los primeros 10 minutos de la clase del próximo 16 de Febrero.

**INSTRUCCIONES:**

- Este taller debe ser escrito en computador y entregado en papel.
- La presentación de los resultados debe obedecer a los formatos estudiados en clase.
- Cuando sea posible, debe mostrar el procedimiento efectuado para llegar a sus resultados.

1. Un investigador desea determinar si las transacciones con tarjetas de crédito o debito son afectadas por la nueva medida del Gobierno de devolver el equivalente a 2ptos porcentuales del IVA siempre y cuando las transacciones se realicen con tarjetas crédito o débito. Para esto, el investigador recurre a dos grupos de individuos con iguales características socioeconómicas. El primer grupo es enterado de la medida, mientras que el segundo grupo no tiene ni idea de la medida del gobierno. Para todos los individuos de ambos grupos se recoge la información de las transacciones realizadas con tarjetas debitos y crédito en el mes de Enero (en miles de pesos) (La información se encuentra en el archivo D\_T4.xls)<sup>1</sup>

a) Escriba un modelo que permita determinar si existe o no un efecto sobre las transacciones promedio con tarjetas de crédito o debito de la nueva medida. (demuestre que su modelo si sirve para este efecto).

**Un modelo que capture esta hipótesis esta dado por:**

$$Y_i = \alpha + \beta D_i + u_i \tag{1}$$

$$i = 1, \dots, N_1, N_1 + 1, \dots, N_1 + N_2$$

**Donde  $Y_i$  corresponde a las transacciones promedio con tarjetas débito y crédito del individuo  $i$  y  $D_i$  es una variable ficticia definida de la siguiente manera**

$$D_i = \begin{cases} 1, & i \in \text{Grupo1} : i = 1, \dots, N_1 \\ 0, & i \in \text{Grupo2} : i = N_1 + 1, \dots, N_1 + N_2 \end{cases}$$

Para entender la interpretación de los parámetros  $\alpha$  y  $\beta$ , tenga en cuenta que:

<sup>1</sup> Estos datos no son reales.

$$E[Y_i | D_i = 1] = E[\alpha + \beta D_i + u_i | D_i = 1] = \alpha + \beta$$

$$E[Y_i | D_i = 0] = E[\alpha + \beta D_i + u_i | D_i = 0] = \alpha$$

El término independiente, por tanto, puede interpretarse como el valor medio de las transacciones con tarjeta de débito o crédito del segundo grupo, mientras que el coeficiente  $\beta$  de la regresión se interpreta como el diferencial existente entre las medias poblacionales de ambos grupos.

b) Estime el modelo y reporte los resultados en una tabla.

La estimación del modelo (1) se presenta en la TABLA 1.

**TABLA 1. Estimación del Modelo (1)**

	VARIABLE DEPENDIENTE: $Y_i$ Estadísticos t entre paréntesis
	Ecuación 1 MCO
constante	211.36147 (6.04) ***
$D_i$	78.84339 (1.60)
$R^2$	0.124068
$R^2$ Ajustado	0.075405
F	2.550
# de Obs.	20

(\*) nivel de significancia: 10%

(\*\*) nivel de significancia: 5%

(\*\*\*) nivel de significancia: 1%

MCO: Mínimos Cuadrados Ordinarios

c) Interprete los coeficientes estimados.

La estimación de  $\alpha$  ( $\hat{\alpha}$ ) debe interpretarse como la estimación muestral de la media poblacional de las transacciones con tarjeta debito o crédito (Y) del segundo grupo, mientras que la suma de las estimaciones de  $\alpha$  y  $\beta$  proporciona una estimación de la media poblacional del primer grupo.

El valor medio de la variable Y para el grupo 1 es igual a 290,20486 miles de pesos, mientras que para el grupo 2, el valor medio es igual a 211,36147 miles de pesos.

Así  $\hat{\beta}$ , 78,843 pesos corresponde a la diferencia en las transacciones realizadas por el grupo 1 y 2.

- d) Determine si de acuerdo a los resultados obtenidos la medida del gobierno afecta las transacciones con estos medios de pago.

El contraste de la hipótesis de igualdad de las medias poblacionales,  $\mu_1 = \mu_2$ , implica verificar:

$$H_0 : E[Y_i | D_i = 1] = E[Y_i | D_i = 0]$$

O, en términos de los parámetros de la regresión:

$$H_0 : \beta = 0$$

De acuerdo a los resultados presentados en la tabla 1, la anterior hipótesis nula no se puede rechazar; por lo que no puede rechazarse la igualdad de medias poblacionales entre los dos grupos y por tanto no existe evidencia que la medida cambia el comportamiento en las transacciones de los individuos.

Un consultor es contratado para determinar la relación que existe entre el gasto público ( $G$ ) y la Remuneración de los Asalariados residentes en la nación por parte de los empleadores Residentes y no residentes en la nación ( $RAR$ ), la Remuneración a los asalariados residentes en la nación por parte de los empleadores No Residentes en la nación ( $RNR$ ) y la Remuneración a los asalariados Residentes y no residentes en la nación por parte de los empleadores Residentes en la nación ( $RRR$ ). (Todas las variables están medidas en millones de dólares)

2. Empleando los datos que se encuentran en el archivo D\_T4.xls, responda las siguientes preguntas:

- a) Escriba e interprete a priori los coeficientes del modelo.

El modelo está dado por:

$$G_i = \beta_1 + \beta_2 RAR_i + \beta_3 RNR_i + \beta_4 RRR_i + \varepsilon_i \quad (2)$$

Se espera que el signo de  $\beta_1$  sea positivo, pues a pesar de que las variables explicatorias sean iguales a cero, puede esperarse que el gasto público ( $G_i$ ) sea positivo.

El signo de los coeficientes  $\beta_2$ ,  $\beta_3$  y  $\beta_4$  es incierto, pues no se puede concluir con certeza el efecto que tienen las variables asociadas a estos parámetros en el Gasto Público.

- b) Estime el modelo y reporte sus resultados en una tabla.

Los resultados se presentan en la TABLA 2. Estimación del Modelo (2) y (3)

TABLA 2. Estimación del Modelo (2) y (3)

	VARIABLE DEPENDIENTE: $G_i$ Estadísticos t entre paréntesis	
	Ecuación 2 1987-1995 MCO	Ecuación 3 1987-1995 MCO
constante	3,406,812.70 (8.51) ***	3,566,715.209 (16.84) ***
$RAR_t$	1.853 (-2.02) *	0.129 (41.56) ***
$RNR$	-12.319 (-2.02) *	-13.484 (-2.57) **
$RRR_t$	-1.722 (-0.48)	--
$R^2$	0.99670	0.99654
$R^2$ Ajustado	0.9947	0.9954
F	503.01 ***	864.73 ***
# de Obs.	9	9

(\*) nivel de significancia: 10%

(\*\*) nivel de significancia: 5%

(\*\*\*) nivel de significancia: 1%

MCO: Mínimos Cuadrados Ordinarios

c) Determine la significancia individual y conjunta de los coeficientes estimados. De acuerdo a la estimación del modelo, el intercepto es significativo individualmente con un nivel de significancia del 1%. Los coeficientes asociados a las variables ( $RAR$ ) y ( $RNR$ ) son significativos individualmente a un nivel de significancia del 10%. El otro coeficiente no es significativo en forma individual.

Para la significancia conjunta, se debe probar la hipótesis:  $H_0 = \beta_2 = \beta_3 = \beta_4 = 0$ . Para esto, se compara el  $F_{calculado} = 503.01$  con  $F_{(0,01),(3,6)} = 9.779$ , dado que  $F_C > F_{tabla}$ , se rechaza la hipótesis nula. Los coeficientes asociados a las pendientes son significativos conjuntamente a un nivel de significancia del 1%.

¿Qué tan bueno es el "fit" del modelo?

Noten que el modelo tiene un alto  $R^2$ , aparentemente el fit del modelo sería muy bueno, pero los todos los síntomas de multicolinealidad están presentes, así el  $R^2$  puede ser alto por la presencia de este problema econométrico.

3. Continuando con el ejercicio anterior, determine si existe o no multicolinealidad. (muestre todo su trabajo y conclusiones)

Como se mencionó anteriormente, existen síntomas de multicolinealidad, así emplearemos nuestras pruebas para corroborar la presencia o no de este problema.

**Matriz de Correlación de las X's**

Al calcular el determinante de las matrices de correlaciones obtenemos:

$$|R| = 0.9803568 \cdot 0.0000004 \cdot 2.0196427 = 7.92E-7$$

Este determinante es muy cercano a cero, por tanto el problema de multicolinealidad es muy serio.

**Medida de Multicolinealidad de Belsley, Kuck y Welsch (1980)**

En este caso tenemos que:

$$\kappa(X) = \frac{\sqrt{\lambda_{Max}}}{\sqrt{\lambda_{Min}}} = \frac{\sqrt{2.0196427}}{\sqrt{0.0000004}} = 2247.022$$

En este caso como  $\kappa(X)$  es mayor que 30, tenemos un problema de multicolinealidad grande.

**Matriz de Correlaciones de los Coeficientes Estimados.**

En este caso tenemos la siguiente matriz de correlaciones de los valores estimados:

	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_1$
$\hat{\beta}_2$	1.000000	0.394052	-1.000000	-0.824496
$\hat{\beta}_3$		1.000000	-0.394138	-0.806432
$\hat{\beta}_4$			1.000000	0.824350
$\hat{\beta}_1$				1.000000

Noten lo elevado de la correlación en especial entre  $\hat{\beta}_4$  y  $\hat{\beta}_2$ . Así, el problema de multicolinealidad es grave.

**Conclusión.**

Todas nuestras pruebas brindan evidencia en torno a la fuerte multicolinealidad presente en este modelo.

4. Corrija el problema de multicolinealidad. Explique claramente su decisión y reporte el modelo estimado que usted considere adecuado.

Noten que en este caso para solucionar el problema, podemos omitir una variable. Podemos suprimir la variable  $R_{RR}$ , que está fuertemente correlacionada con la variable  $R_{AR}$  y es la variable que que menos información aporta al modelo. Así el modelo será:

$$G_i = \beta_1 + \beta_2 R_{AR} + \beta_3 R_{NR} + \varepsilon_i \quad (3)$$

La estimación del modelo (3) se presenta en la TABLA 2.

5. Continuando con el ejercicio anterior,

a) Constate que el nuevo modelo no posee problemas de multicolinealidad.

**Matriz de Correlación de las X's**

Al calcular el determinante de las matrices de correlaciones obtenemos:

$$|R| = 1.0998745 \cdot 0.9001255 = 0.99$$

Este determinante es muy cercano a uno, por tanto el problema de multicolinealidad ha sido solucionado.

**Medida de Multicolinealidad de Belsley, Kuck y Welsch (1980)**

En este caso tenemos que:

$$\kappa(X) = \frac{\sqrt{\lambda_{Max}}}{\sqrt{\lambda_{Min}}} = \frac{\sqrt{1.0998745}}{\sqrt{0.9001255}} = 1.105$$

En este caso como  $\kappa(X)$  es muy cercano a 1, así no existe un problema de multicolinealidad.

**Matriz de Correlaciones de los Coeficientes Estimados.**

En este caso tenemos la siguiente matriz de correlaciones de los valores estimados:

	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_1$
$\hat{\beta}_2$	1.000000	-0.000059	-0.000164
$\hat{\beta}_3$		1.000000	-0.925551
$\hat{\beta}_1$			1.000000

Noten que ahora las correlaciones son relativamente bajas.

**Conclusión.**

Todas nuestras pruebas brindan evidencia en torno a que la multicolinealidad ha desaparecido.

b) Compare la varianza de los estimadores encontrados en el punto 2 b) y 4. Intuitivamente explique porque ocurre lo que observa.

La matriz de varianzas y covarianzas del modelo (2) corresponde a:

	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_1$
$\hat{\beta}_2$	114.113526	77.1011584	-113.946952	-10584077.1
$\hat{\beta}_3$	77.1011584	335.488209	-77.0054253	-17750159.8
$\hat{\beta}_4$	-113.946952	-77.0054253	113.78072	10566764.1
$\hat{\beta}_1$	-10584077.1	-17750159.8	10566764.1	1,444,082,190,000

Por otro lado la matriz de varianzas y covarianzas del modelo (3) corresponde a:

	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_1$
$\hat{\beta}_2$	27.469	-0.002	-181.760
$\hat{\beta}_3$	-0.002	27.469	-1027386.472
$\hat{\beta}_1$	-181.760	-1027386.472	44,856,885,889

Noten que la matriz de varianzas y covarianzas del último modelo es mucho menor que la del primer modelo. Como lo discutimos en clase, la multicolinealidad provoca que la matriz

$X^T X$  sea muy grande y por tanto la matriz de varianzas y covarianzas estimada  $\left( s^2 (X^T X)^{-1} \right)$  será mucho más grande.

6. Continuando con el ejercicio anterior, conteste:

- a) Interprete los coeficientes estimados y discuta la significancia individual y conjunta de los coeficientes.

$\hat{\beta}_1 = 3566715.209$ . 3566715.209 millones de dólares corresponden al gasto público que no depende de la Remuneración de los Asalariados residentes en la nación por parte de los empleadores Residentes y no residentes en la nación (*RAR*), ni de la Remuneración a los asalariados residentes en la nación por parte de los empleadores No Residentes en la nación (*RNR*).

$\hat{\beta}_2 = 0.129$ . Un aumento de un millón de dólares en la Remuneración de los Asalariados residentes en la nación por parte de los empleadores Residentes y no residentes en la nación (*RAR*) provocará un aumento en el Gasto Público en 0.129 millones de dólares.

$\hat{\beta}_3 = -13.484$ . Un aumento de un millón de dólares en la Remuneración a los asalariados residentes en la nación por parte de los empleadores No Residentes en la nación (*RNR*) provocará una disminución en el Gasto Público en -13.484 millones de dólares.

De acuerdo a la estimación del modelo, el intercepto y el coeficiente asociado a la variable (*RAR*) son significativos individualmente con un nivel de significancia del 1%. El coeficiente asociado a la variable (*RNR*) es significativo individualmente a un nivel de significancia del 5%.

Para la significancia conjunta, se debe probar la hipótesis:  $H_0 = \beta_2 = \beta_3 = 0$ . Se compara el  $F_{calculado} = 864.73$  con  $F_{(0.01),(2,7)} = 9.546$  para tomar una decisión. Como el  $F_C > F_{tabla}$ , se rechaza la hipótesis nula. Los coeficientes son significativos conjuntamente a un nivel de significancia del 1%.

- b) Determine si es cierto que el efecto sobre el gasto público de la remuneración de los asalariados residentes en la nación por parte de los empleadores residentes y no residentes en la nación es 10 veces más pequeño y contrario al de remuneración a los asalariados residentes en la nación por parte de los empleadores no residentes en la nación.

Noten que esta hipótesis implica que  $-10\beta_2 = \beta_3$ , es decir  $-(10\beta_2 + \beta_3) = 0$ . Esta hipótesis se puede escribir de la forma  $R\beta = C$ , donde

$$R = \begin{bmatrix} 0 & 10 & 1 \end{bmatrix} \text{ y } C = \begin{bmatrix} 0 \end{bmatrix}$$

Esto lo podemos probar con una prueba de Wald, en este caso tenemos que el estadístico corresponde a 5.42. AL comparar este estadístico con el valor crítico de la distribución Chi-cuadrado con 1 grado de libertad, tenemos que a un nivel de sinificancia del 5% se puede rechazar la hipótesis nula.

Así, con un nivel de confianza del 95% (y también al 90%) podemos concluir que la remuneración de los asalariados residentes en la nación por parte de los empleadores residentes y no residentes en la nación NO es 10 veces más pequeño y contrario al de remuneración a los asalariados residentes en la nación por parte de los empleadores no residentes en la nación.