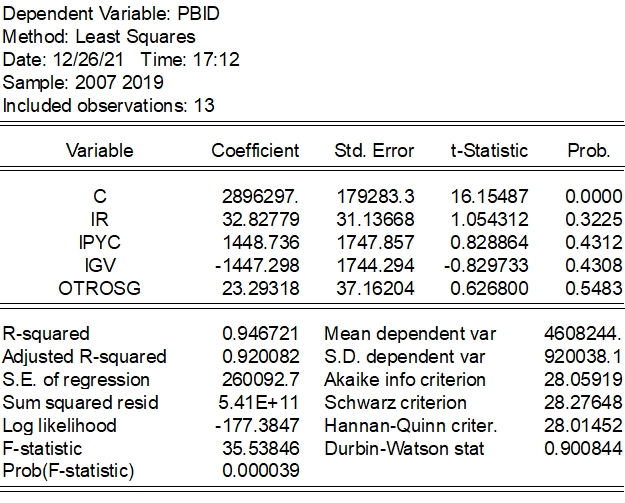
🔵GRUPO EDISON

USO DE LA HOJA **DATA**

# Multicolinealidad (y sus múltiples hallazgos de la validación de auto y su solución)

**data pbid IR IPyC IGV OtrosG**

**ls pbid c IR IPyC IGV OtrosG**



Se observa que las variables dependientes no son significativas

MÉTODOS PARA DETECTAR MULTICOLINEALIDAD

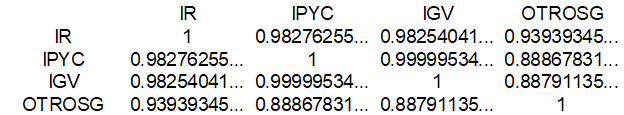
## MÉTODO 1: EXAMEN DE REGRESIÓN

EL R2 es 0.946721 (este R2 debería por coherencia estar acompañado con estimadores que son estadísticamente significativos), entonces nos fijamos que los estimadores NO SON SIGNIFICATIVO, nos indica que hay presencia de multicolinealidad.

## MÉTODO 2: MATRIZ DE CORRELAIÓN SIMPLE

Hacemos la correlación entre las variables independientes

**cor IR IPyC IGV OtrosG**



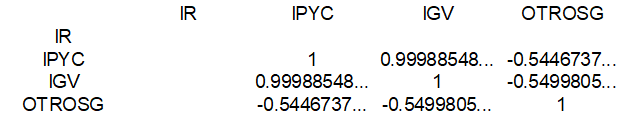
Se supone que entre las variables independientes debería existir baja asociación lineal o baja correlación.

Entonces, observando la matriz de correlación, todas las variables del modelo de regresión están asociadas linealmente, entonces no se cumple con el supuesto. HAY EXISTENCIA DE MULTICOLINEALIDAD

## MÉTODO 3: MATRIZ DE CORRELACIÓN PARCIAL

### MATRIZ DE CORRELACIÓN DE ORDEN UNO

Mantenemos constante la variable IR, así tenemos la matriz de 3x3

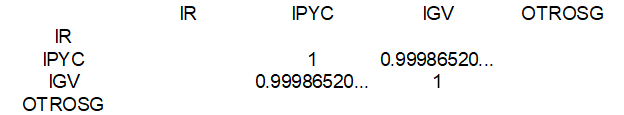


Se observa cómo ha bajado la asociación lineal entre las variables, pero sigue alta asociación las variables IPYC e IGV.

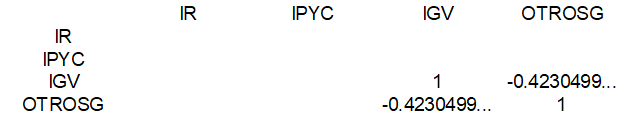
### MATRIZ DE CORRELACIÓN DE ORDEN DOS

Ahora tenemos una matriz de 2x

Manteniendo IR y OTROSG constantes



Manteniendo IR el IPYC constantes, aquí se observa que la asociación entre las variables OTROS e IGV es baja.

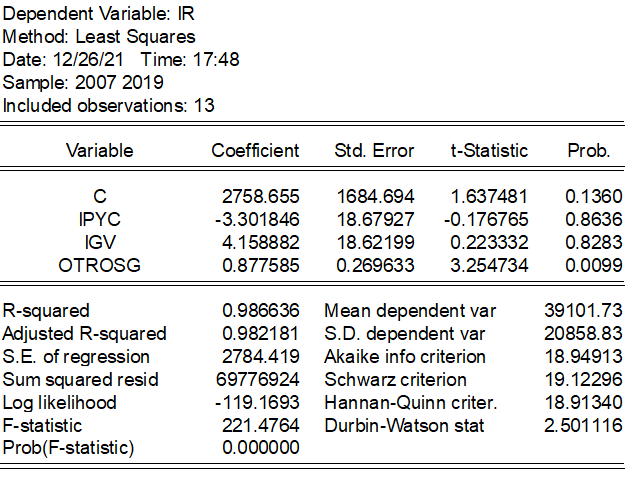


## MÉTODO 4: REGRESIÓN AUXILIAR

Hacemos la regresión de todas las variables independientes sin tener en cuenta la variable dependiente.

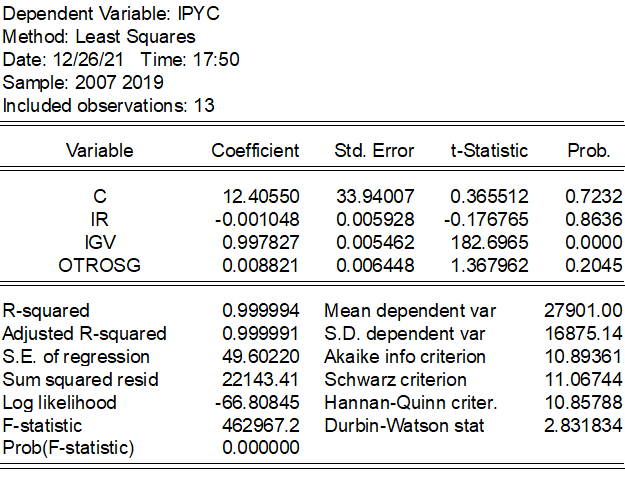
### AUXILIAR 1

**ls IR c IPYC IGV OTROSG**



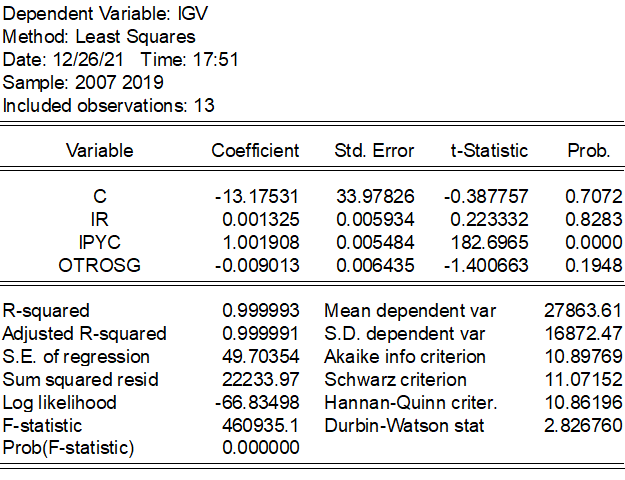
### AUXILIAR 2

ls IPYC c IR IGV OTROSG



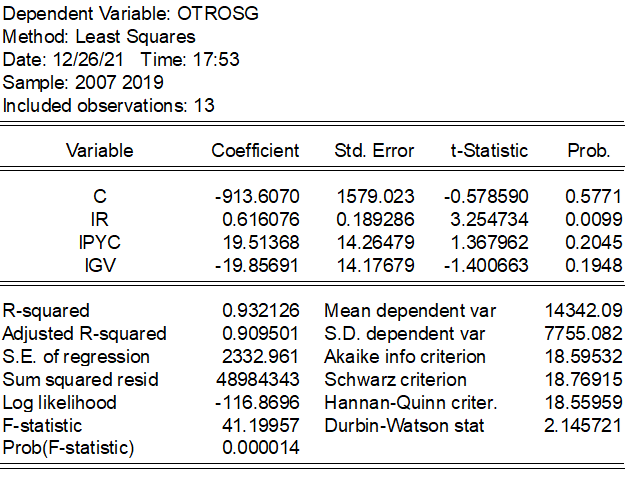
### AUXILIAR 3

ls IGV c IR IPYC OTROSG



### AUXILIAR 4

ls OTROSG c IR IPYC IGV



### Aplicamos la REGLA DE KLEIN

Nosotros ya hemos analizado el R2 de la regresión original, ahora determinamos si algunas de las 4 regresiones auxiliares tienen un R2 más alto que la regresión original.

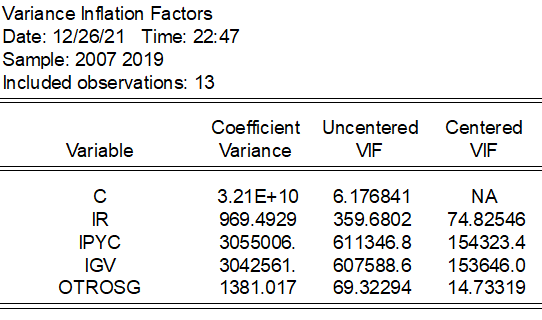
Comparando los R2 de las regresiones auxiliares, con el R2 de la regresión original. Se observa que las R2 de las AUXILIARES son superiores al R2 del ORIGINAL. Entonces se concluye que existen graves problemas de multicolinealidad.

## MÉTODO 5: MEDIDAS CONTEMPORÁNEAS (Medidas propias de Eviews, que nos va a permitir determinar la existencia o no de la multicolinealidad)

### ÍNDICE DE CONDICIÓN

Hacemos uso de los factores de varianza FIV o VIF

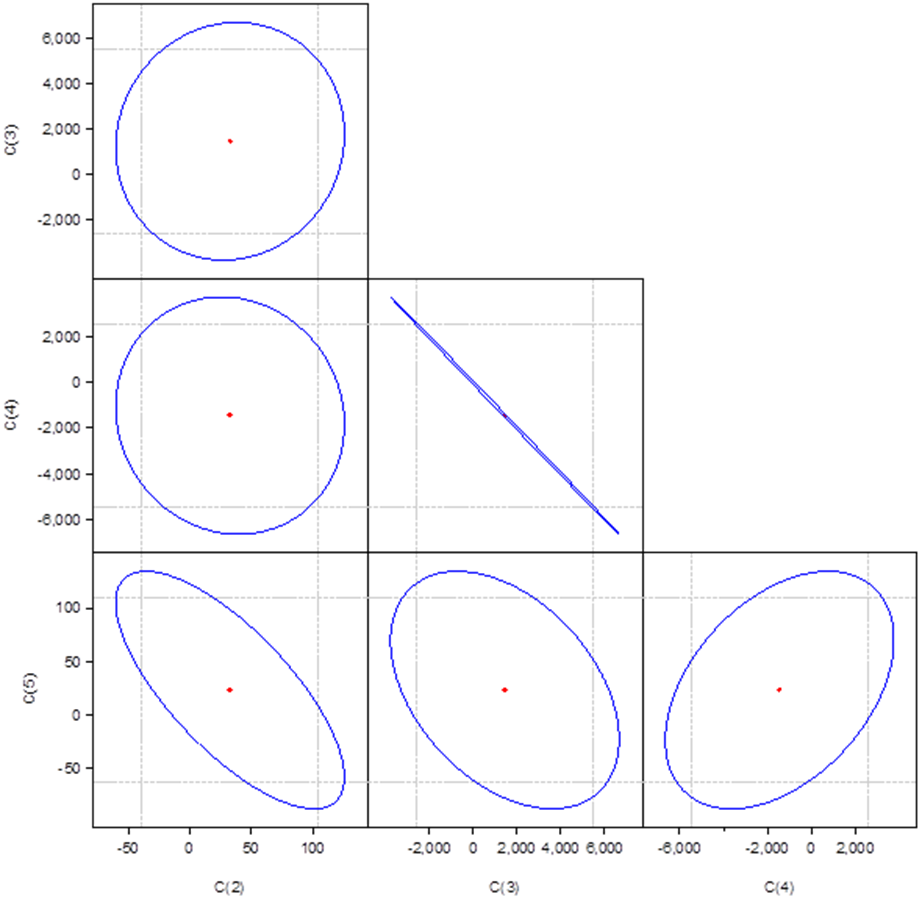
Nos indica que si el FIV es más grande que 10 se tiene un serio problema de multicolinealidad.



En el gráfico se observa que hay un serio problema de multicolinealidad. Ya que los VIF son mayores que 10.

### GRÁFICO DE ELIPSES

Los gráficos muestran la relación entre variables:



Si la relación entre variables es baja se va a dibujar un círculo, por el contrario, se va a asemejar a un elipse.

En los gráficos se puede observar que las variables c(4), c(3) y c(5) y c(2) tienen problemas serios de multicolinealidad.

### GRÁFICO DE DISPERSIÓN ENTRE LAS REGRESORAS

Revisar BLOCK.DOCX

## SOLUCIÓN A LA MULTICOLINEALIDAD

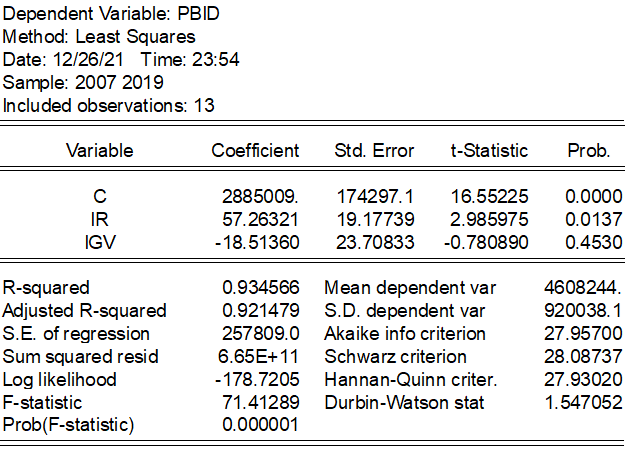
### AMPLIAR LA MUESTRA

### SUPRIMIR ALGUNAS VARIABLES CON JUSTIFICACIÓN ESTADÍSTICA Y ECONÓMICA

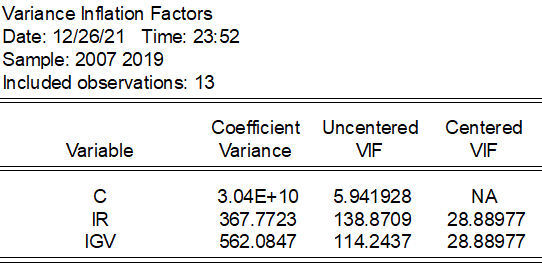
En este caso podemos eliminar la variable IPYC pues sabemos que está relacionado con la variable IGV. Además, el Otro impuesto.

Entonces hacemos la regresión:

ls pbid c IR IGV



**Hacemos la prueba de multicolinealidad**

****

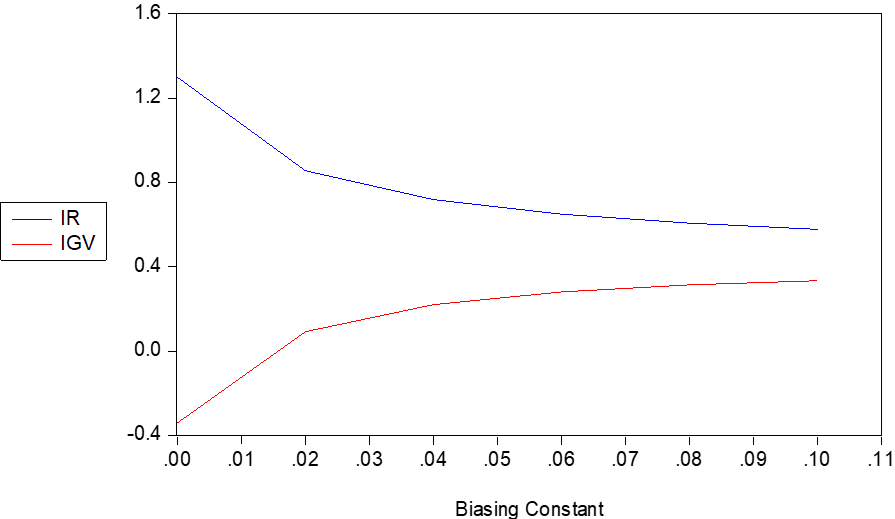
Se observa que sigue existiendo multicolinealidad

### REGRESIÓN EN CADENA

Entonces se procede a usar la regresión en cadena que ofrece como estimadores de los parámetros

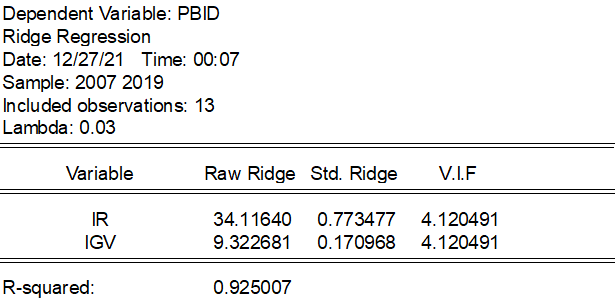


Siendo “c” una constante adecuada. Donde en la vida práctica, suele estar entre 0.01 y 0.1 haciendo que el ajuste sea bueno en cuanto a R2 y significancia individual y conjunta. (permitimos que se incremente un poquito el sesgo a cambio de que se reduzca la varianza alta que provoca la multicolinealidad)



En el gráfico anterior se observa que la varianza no es estable desde 0.00 hasta 0.03, por lo cual si tomamos el valor 0.03 para la c los variables son mucho más estables.Hacemos la regresión en cadena.

Tenemos la siguiente ilustración:



Donde los VIF son menores a 10 que significa que no hay presencia de multicolinealidad.

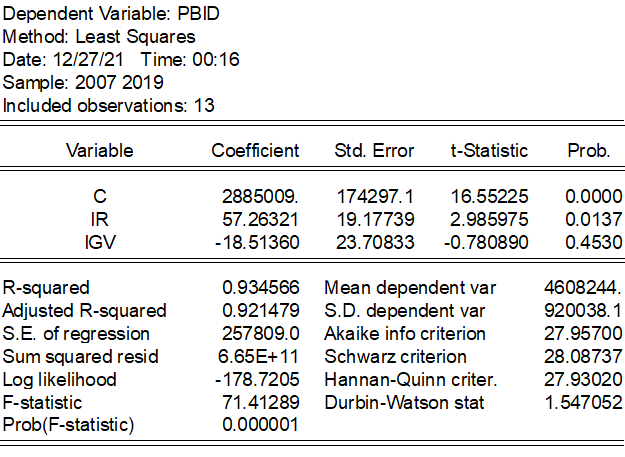
POR LO TANTO, el modelo está corregido de multicolinealidad.

# Autocorrelación (y sus múltiples hallazgos de la validación de auto y su solución)

## PRUEBA DE DW

data pbid IR IGV

ls pbid c IR IGV



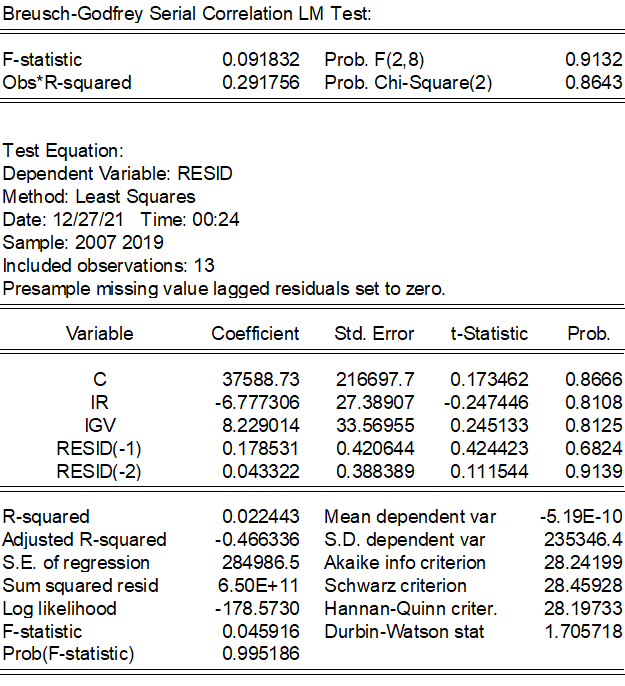
Se tiene la siguiente información:

* d = 2 indica que no hay autocorrelación
* d <2 indica correlación serial positiva
* d > 2 indica correlación serial negativa

En general, si d es menor que 1,5 o mayor que 2,5, existe un problema de autocorrelación potencialmente grave. De lo contrario, si d está entre 1,5 y 2,5, es probable que la autocorrelación no sea motivo de preocupación.

El estadístico Durbin-Watson sta es 1.547052 que es mayor a 1.5, por lo tanto, no hay problemas de autocorrelación.

## SOLUCIÓN PRUEBA DE BREUSCH-GODFREY\*



Las probabilidades menos a 0.05 rechazan la hipótesis nula

H0: (No existe correlación entre los residuos) Que algún error sea igual de 0

H1: (Los residuos están autocorrelacionados) Que algún error sea diferente de 0

En este caso como 0.8643>0.05 entonces aceptamos la hipótesis nula. Entonces podemos inferir que NO HAY AUTOCORRELACIÓN.

\*La prueba de Breusch-Godfrey no es necesaria en este modelo, ya que no hay presencia de autocorrelación.