## Estimación de un modelo corregido y con rezago y análisis de seguimiento de supuestos Gauss-Márkov

### 1.- Explicación del modelo.

Se trata de un modelo basado en una función Cobb-Douglas (Capital, Trabajo), el cual ha sido corregido al transformar las variables regresoras y la dependiente en forma logarítmica y con un rezado. Se planea utilizar este modelo para explicar el Producto Interno Bruto de un país.

#### 2.- Resultados del modelo.

```
Call: lm(formula = 1PIB ~ lag1StK + lag1L)
```

#### Residuals:

```
Min 1Q Median 3Q Max -0.089699 -0.025698 0.005854 0.028584 0.068884
```

#### Coefficients:

```
Estimate Std. Error t value Pr(>|t|)

(Intercept) 0.60478 0.36849 1.641 0.10488

laglStK 0.12391 0.03767 3.289 0.00152 **

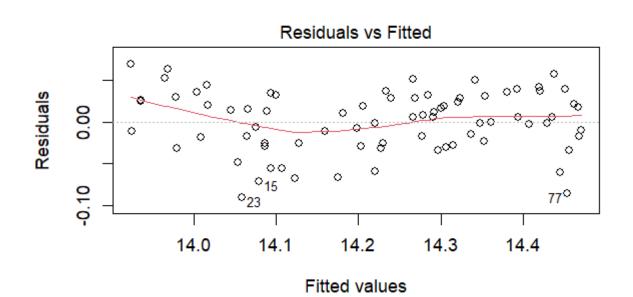
laglL 1.12164 0.07058 15.892 < 2e-16 ***
```

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

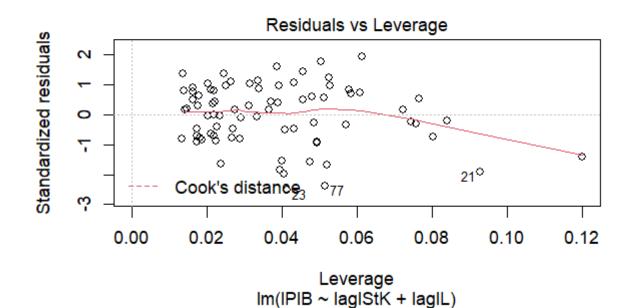
Residual standard error: 0.03654 on 76 degrees of freedom (1 observation deleted due to missingness)

Multiple R-squared: 0.9531, Adjusted R-squared: 0.9518

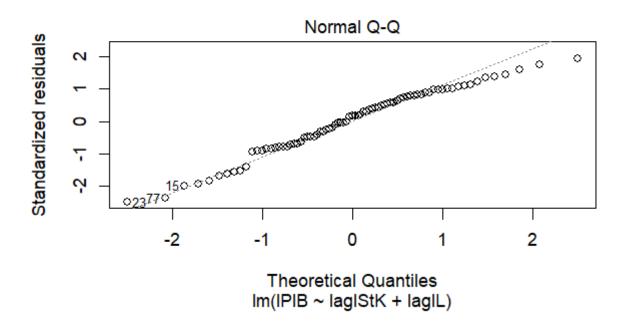
F-statistic: 771.6 on 2 and 76 DF, p-value: < 2.2e-16



Im(IPIB ~ lagIStK + lagIL)

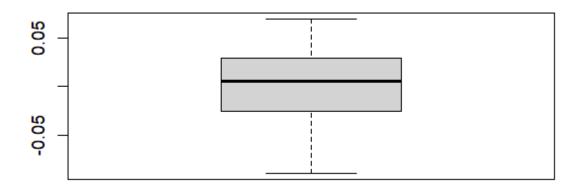


- a) Pruebas de Normalidad de la distribución de los residuos.
  - Prueba Gráfica de distribución residual.

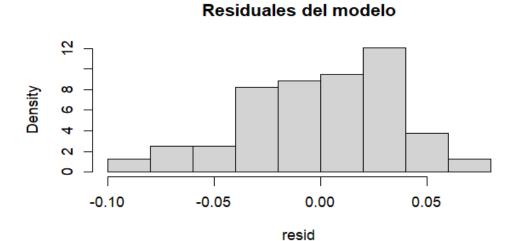


La gráfica muestra que la mayor parte de los residuos estandarizados caen sobre la línea. Sin embargo, una porción notable de éstos, en especial aquellos en el extremo derecho, divergen. A partir de ello, no es tan claro si los residuos del modelo están distribuidos de manera normal.

Diagrama de Caja y Bigotes.



• Histograma.



Los dos gráficos muestran que la mayoría de los residuos se distribuyen en los valores intermedios del gráfico, lo que podría indicar que los residuos tienen una distribución normal.

Prueba Jarque-Bera.

Jarque-Bera test for normality

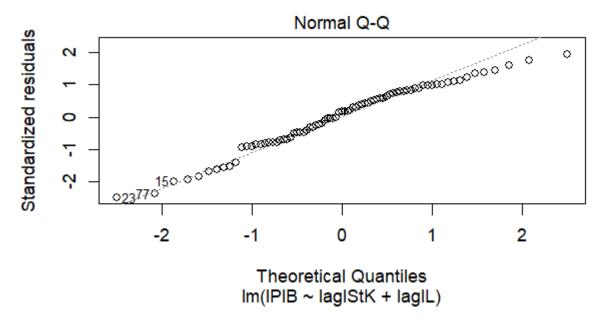
Puesto que la Prueba Jarque-Bera es menor al valor límite de 5.99 y su *valor p* no alcanza una significancia estadística de 5%, la prueba indica que no es posible rechazar la hipótesis nula: los residuos están distribuidos de forma normal.

Prueba Shapiro-Wilk.

Puesto que el valor de la "W" supera el umbral establecido por la prueba (0.96) y posee una significancia estadística menor a 5%, no se puede rechazar la hipótesis nula, la cual propone que los residuos se distribuyen de forma normal.

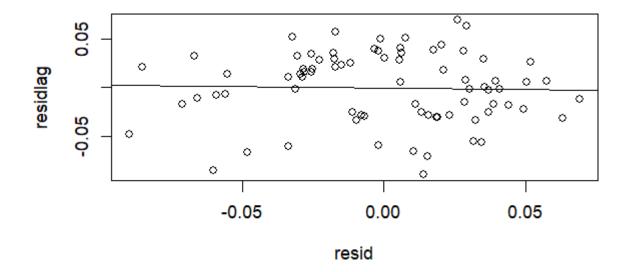
**Conclusión:** la mayoría de las pruebas que se aplicaron a los residuales del modelo propuesto indican que éstos cumplen con el supuesto de la distribución normal. Por ello, es probable que el modelo tenga capacidad de predicción y estimadores eficientes y exactos.

- b) Pruebas de Autocorrelación.
  - Gráfica de dispersión de residuales.



Puesto que el inicio y, en especial, el extremo derecho de la línea de dispersión los residuos estandarizados muestran una menor alineación, es posible que el modelo tenga algún grado de autocorrelación.

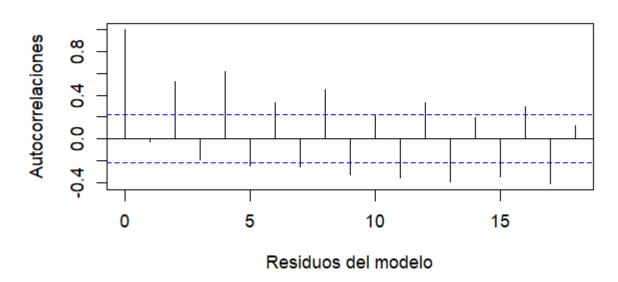
Gráfica de regresión de residuos contra residuos rezagados.



A partir de esta gráfica, es difícil observar una relación fuerte entre los residuos (quizá una relación negativa ligera). De esta prueba se puede inferir que es posible que los residuos del modelo no sufran de una autocorrelación notable.

Correlograma.

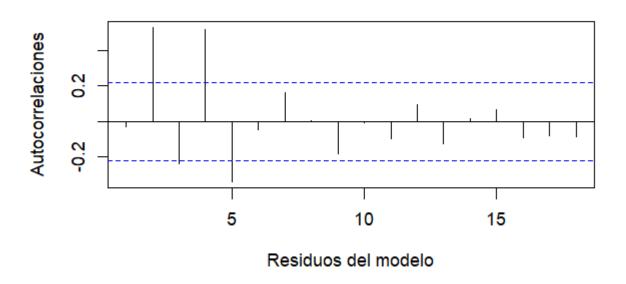
## Correlograma del modelo



Puesto que, muchos de los residuos del modelo, se salen del umbral anticipado del correlograma, es posible que éstos tengan autocorrelación.

Correlograma Parcial.

## Correlograma del modelo



Puesto que los residuos del modelo, en especial los que están ubicados en al inicio, se salen del umbral anticipado del correlograma, es posible que éstos sufran de autocorrelación.

Prueba Durbin-Watson.

```
data: modfin

DW = 2.0496, p-value = 0.521

alternative hypothesis: true autocorrelation is greater than 0
```

La cifra de 2.0496 cae en la zona deseable de la prueba. Por eso, no se puede rechazar la hipótesis nula, la cual establece que los residuos no tienen autocorrelación. Este resultado se podría deber a que el modelo posee variables rezagadas, las cuales son correctores ante la autocorrelación de los residuos.

Prueba Breusch-Godfrey.

```
data: modfin

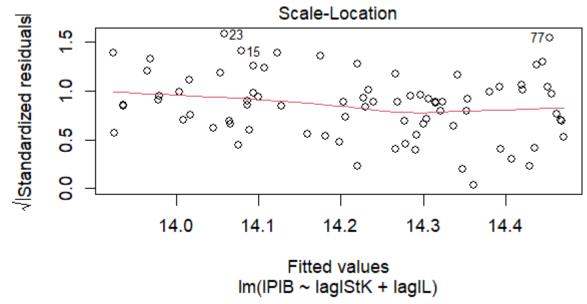
LM test = 0.072382, df = 1, p-value = 0.7879
```

Puesto que la cifra de la prueba no supera al umbral (3.841), pero no alcanza una significancia estadística de 5%, por lo que es posible rechazar la hipótesis nula (no hay autocorrelación) a favor de la hipótesis alternativa (hay autocorrelación) en 1 grado de libertad.

Conclusión: algunas de las pruebas muestran que los residuos del modelo tienen autocorrelación (correlogramas y Prueba Breusch-Godfrey), mientras que otras indican que ésta no es relevante (Prueba Durbin-Watson). Es probable que el modelo sufra de autocorrelación, ya que se comporta como una serie de tiempo (el valor de una observación depende del valor de otra observación). Sin embargo, el hecho de que este modelo contenga variables rezagadas mitiga los problemas que causan la autocorrelación en la estimación. Por ello, es probable que los estimadores del modelo se sigan teniendo algún grado de eficiencia.

### c) Homocedasticidad.

Método Gráfico.



Puesto que la gráfica muestra que los residuos estandarizados no tienen una distribución definida, es posible que el modelo cumpla con el supuesto de homocedasticidad.

Prueba de Breusch-Pagan-Godfrey.

Studentized Breusch-Pagan test

data: mod1BP = 3.1167, df = 2, p-value = 0.2105

La prueba tiene una significancia estadística baja (un p-valor mucho menor a 0.05), lo que indica que no se puede rechazar la hipótesis nula de que el modelo posee homocedasticidad.

Prueba de White.

Estadístico White = 9.7551, df = 4, p-value = 0.04476

Puesto que la prueba tiene un valor ligeramente mayor a 9.44 (valor de la chi-cuadrada para 4 grados de libertad) y tiene un alto nivel de significancia estadística (un p-valor mayor a 0.05), indica que el modelo no cumple con el supuesto de homocedasticidad (tiene heterocedasticidad).

Conclusión: algunas pruebas indican que el modelo tiene homocedasticidad (Método Gráfico y la Prueba Breusch-Pagan-Godfrey) y otras indican que tiene heterocedasticidad, en el margen (Prueba de White). Por ello, es posible que el modelo presente una cantidad ligera de heterocedasticidad, lo cual podría indicar que los estimadores son menos eficientes y que su varianza podría ser menor que óptima.

- d) Multicolinealidad.
  - Correlación entre variables.

	PIB	StK	L
PIB	1.0000000	0.9285587	0.9752384
StK	0.9285587	1.0000000	0.8956729
L	0.9752384	0.8956729	1.0000000

Puesto que las correlaciones mostraron que las variables tienen, entre sí, una relación mayor a 0.8, a primera vista, parece que el modelo podría sufrir de una multicolinealidad relevante.

## Regla de Klein

```
Call: lm(formula = laglStK ~ laglL)
```

```
Residuals:
```

```
Min 1Q Median 3Q Max -0.30485 -0.05727 0.01337 0.06942 0.25940
```

### Coefficients:

```
(Intercept) -2.5301 1.0768 -2.35 0.0214 * laglL 1.6439 0.1024 16.05 <2e-16 ***
```

Signif. codes: 0 \\*\*\*' 0.001 \\*\*' 0.01 \\*' 0.05 \.' 0.1 \' 1

Residual standard error: 0.1105 on 77 degrees of freedom

Estimate Std. Error t value Pr(>|t|)

(1 observation deleted due to missingness)
Multiple R-squared: 0.7699, Adjusted R-squared: 0.7669

F-statistic: 257.6 on 1 and 77 DF, p-value: < 2.2e-16

```
Call: lm(formula = laglL ~ laglStK)
Residuals:
    Min 10 Median 30
                                    Max
-0.14457 -0.03829 0.00603 0.03415 0.12331
Coefficients:
           Estimate Std. Error t value Pr(>|t|)
(Intercept) 3.60412 0.43049 8.372 1.95e-12 ***
laglStK 0.46832 0.02918 16.050 < 2e-16 ***
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual standard error: 0.059 on 77 degrees of freedom
  (1 observation deleted due to missingness)
Multiple R-squared: 0.7699, Adjusted R-squared: 0.7669
```

Puesto que las  $R^2$  de 2 de las correlaciones auxiliares (0.7699 en ambos casos) son menores a la  $R^2$  original del modelo (0.9531), es posible que el modelo no tenga multicolinealidad relevante.

F-statistic: 257.6 on 1 and 77 DF, p-value: < 2.2e-16

• Regla de Theli.

$$R^2 \ original \ (y=PIB) = 0.9531$$
 
$$R^2 \ modelo \ auxiliar \ 1(y=Logaritmo \ del \ trabajo \ rezagado) = 0.7699$$
 
$$R^2 \ modelo \ auxiliar \ 2 \ (y=Logaritmo \ del \ capital \ rezagado) = 0.7699$$

0.9531 - (0.9531 - 0.7699) - (0.9531 - 0.7699) = 0.5867

Puesto que la cifra es ligeramente menor a la regla de decisión de la Regla de Theli (0.6), es posible que el modelo no sufra de multicolinealidad relevante.

• Factor de Inflación de la Varianza.

Puesto que las cifras son menores a la regla de decisión de la prueba de Inflación de la Varianza (arriba de 10.00), la prueba indica que no hay multicolinealidad relevante en el modelo.

**Conclusión:** casi la totalidad de las pruebas (a excepción de las correlaciones entre variables) mostraron que no existe multicolinealidad relevante en el modelo.

## e) Especificación del Modelo.

#### Pruebas de indicio.

```
Call: lm(formula = lPIB ~ laglStK + laglL)
```

#### Residuals:

```
Min 1Q Median 3Q Max -0.089699 -0.025698 0.005854 0.028584 0.068884
```

### Coefficients:

```
Estimate Std. Error t value Pr(>|t|)

(Intercept) 0.60478 0.36849 1.641 0.10488

laglStK 0.12391 0.03767 3.289 0.00152 **

laglL 1.12164 0.07058 15.892 < 2e-16 ***
```

```
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '' 1
```

Residual standard error: 0.03654 on 76 degrees of freedom

(1 observation deleted due to missingness)

Multiple R-squared: 0.9531, Adjusted R-squared: 0.9518

F-statistic: 771.6 on 2 and 76 DF, p-value: < 2.2e-16

Puesto que ambas variables regresoras del modelo en cuestión son estadísticamente significativas a un intervalo de confianza mayor a 99%, su magnitud tiene un efecto notable sobre la dependiente y la bondad de ajuste conjunta es mayor a .9, es probable que el modelo no contenga una variable irrelevante.

Prueba RESET de Ramsay.

```
RESET test
```

```
data: modfin

RESET = 4.7835, df1 = 1, df2 = 75, p-value = 0.03185
```

Puesto que la cifra tiene una significancia estadística (p-value) mayor a 5%, la prueba Ramsay indica que el modelo captura que, en cierta medida, el modelo podría estar mejor especificado.

• Prueba Log likehood.

```
logLik(modfin, todas las variables)
'log Lik.' 150.8666 (df=4)

logLik(modfinalt1, sin trabajo)
'log Lik.' 93.03786 (df=3)

logLik(mod3, sin capital)
'log Lik.' 145.6089 (df=3)
```

De acuerdo con los resultados de la prueba, parece que el mejor modelo es el original.

Prueba Akaike.

```
AIC (modfin, todas las variables)
[1] -293.7331

AIC (modfinalt1, sin trabajo)
[1] -180.0757

AIC (modfinalt2, sin capital)
[1] -285.2178
```

De acuerdo con los resultados de la prueba, parece que el mejor modelo es el modelo original.

**Conclusión:** las pruebas mostraron resultados mixtos. Por un lado, la prueba de Reset Ramsay indica que la especificación del modelo podría mejorar. Por otro lado, las pruebas de indicio, Log-likehood y Akaike mostraron que el modelo original mejor a las alternativas que omiten alguna de las regresoras. Con esta información, es posible que la estimación del modelo tenga algún sesgo por omisión de una variable, por lo que se podría agregar otra variable para mejorar los cálculos.

### f) Cambio Estructural.

Prueba Chow.

Chow test

## En la observación 5:

F = 1.9931, p-value = 0.1225

### En la observación 15:

data: lPIB ~ laglStK + laglL

F = 5.623, p-value = 0.001594

## En la observación 25:

data: lPIB ~ laglStK + laglL

F = 6.1935, p-value = 0.0008286

## En la observación 35:

data: lPIB ~ laglStK + laglL

F = 6.8812, p-value = 0.0003812

```
En la observación 45:
```

data: lPIB ~ laglStK + laglL

F = 4.2008, p-value = 0.00847

## En la observación 55:

data: lPIB ~ laglStK + laglL

F = 3.1654, p-value = 0.02946

# En la observación 65:

data: lPIB ~ laglStK + laglL

F = 1.583, p-value = 0.2008

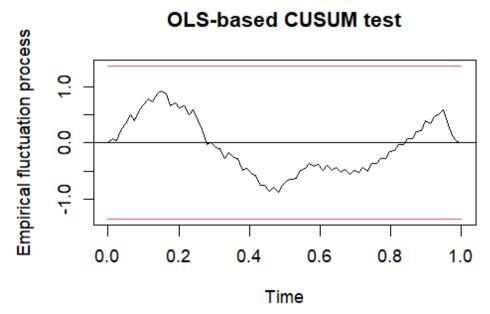
## En la observación 75:

data: lPIB ~ laglStK + laglL

F = 3.9781, p-value = 0.01105

A partir de una serie de pruebas en diferentes puntos, se observan posibles cambios estructurales en varias de las observaciones.

Prueba Cusum.



A partir de la gráfica, se puede observar que el modelo no cruza alguno de los umbrales de cambio estructural, por lo que es probable que la estimación del modelo no tiene cambios estructurales serios.

**Conclusión:** las pruebas mostraron resultados mixtos. La prueba Chow indica que hay varias observaciones que se comportan de manera muy diferente que el resto. Sin embargo, la prueba CUSUM parece indicar que no existe algún cambio que afecte notablemente los datos. Puesto que las observaciones del modelo conciernen un periodo de tiempo inestable en la economía mexicana (Crisis de 1994), es probable que haya cambios estructurales en el modelo. Sin embargo, la conversión del modelo a formato log y el uso de rezagos pudieron haber ayudado a mitigar estos efectos.