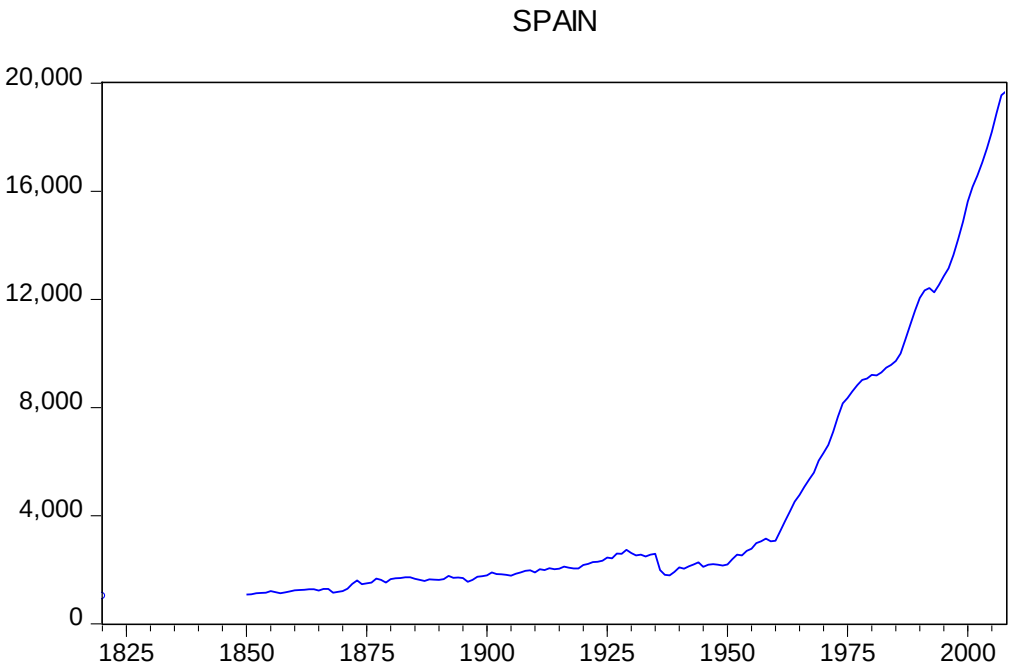
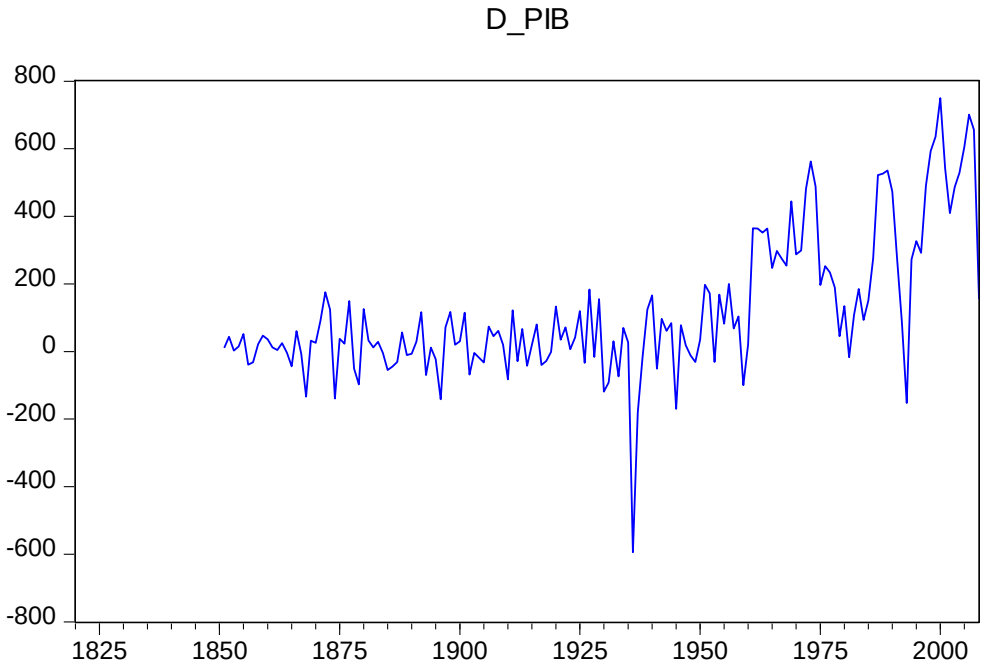
El objetivo de la práctica es discutir que modelo puede ser el más adecuado para predecir la evolución del PIB español.

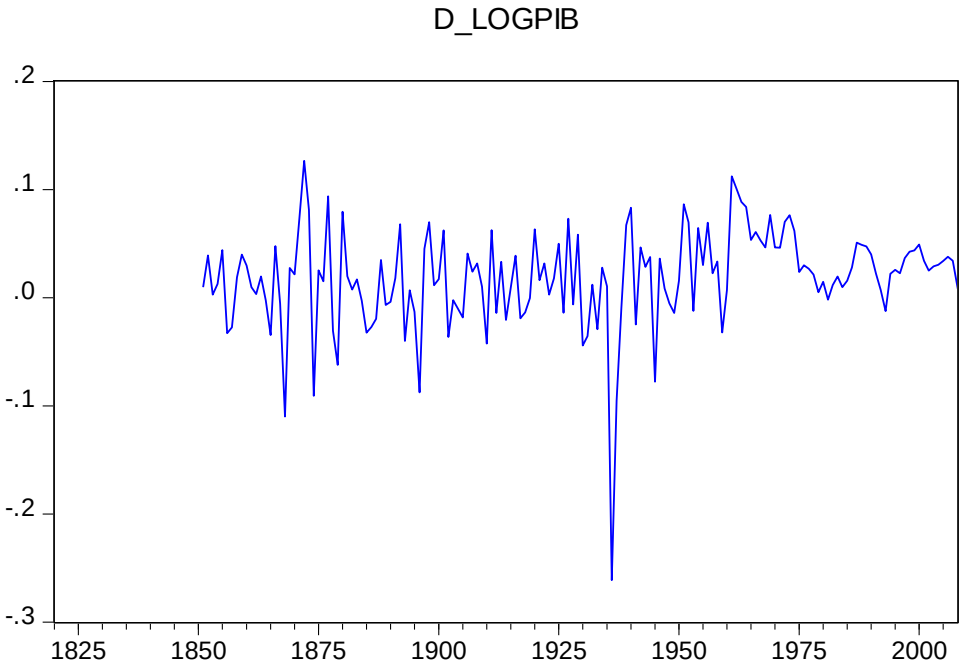
La serie del PIB utilizada sigue la siguiente evolución:



El primer paso para estimar un modelo adecuado es extraer la tendencia de la serie, para eso aplicamos una primera diferencia.

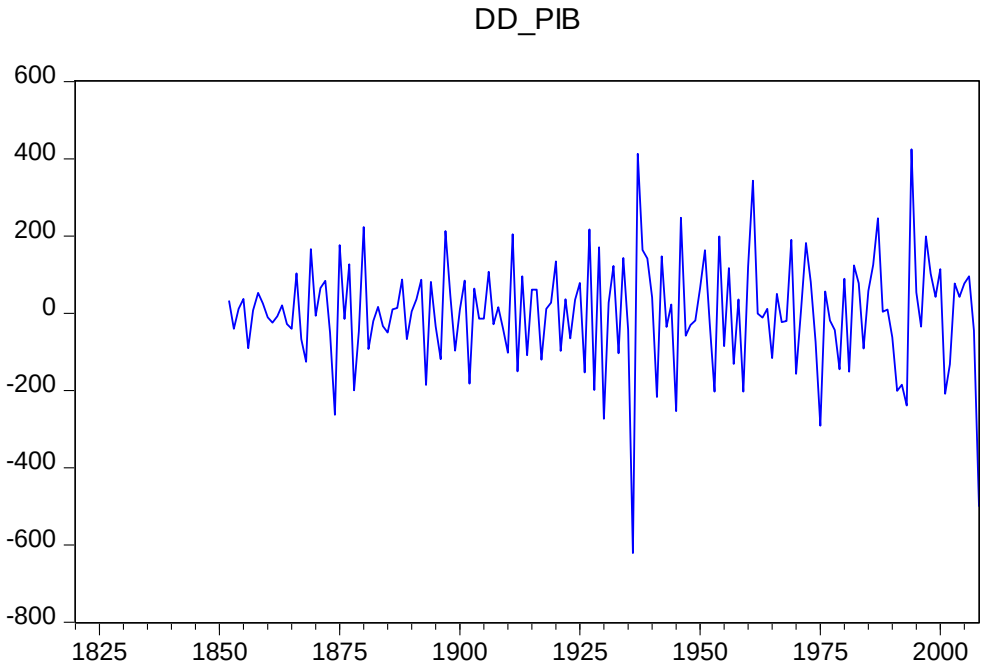


Puede parecer que la serie obtenida sufre de heterocedasticidad. También propondremos modelos sobre el logaritmo del PIB, para eliminar la heterocedasticidad si existiera.



A simple vista, aplicando el logaritmo la serie obtenida se parece mas a una serie estacionaria. Sin embargo, utilizaremos los dos ejemplos para buscar modelos adecuados y luego los compararemos.

Otra alternativa a la primera diferencia sobre el logaritmo es utilizar dos diferencias.



También utilizaremos este ejemplo para estimar modelos.

Una vez obtenido una serie estacionaria adecuada, sin tendencia y heterocedasticidad, buscamos el modelo más adecuado. Utilizamos tres sistemas diferentes:

- Una diferencia
- Diferencia del logaritmo
- Dos diferencias

Empezando por el primer ejemplo, observamos su correlograma para ayudarnos a proponer un modelo adecuado para predecir el PIB.

PRIMERA DIFERENCIA DEL PIB:

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.757	0.757	92.205	0.000
		2	0.633	0.142	157.22	0.000
		3	0.551	0.078	206.79	0.000
		4	0.488	0.047	245.93	0.000
		5	0.479	0.132	283.77	0.000
		6	0.438	0.003	315.67	0.000
		7	0.417	0.051	344.73	0.000
		8	0.413	0.072	373.42	0.000
		9	0.409	0.061	401.84	0.000

Este correlograma sigue la estructura normal de un AR(1), en la columna FAS todos los valores son significativos y decrecen poco a poco, mientras que en la columna de la FAC solo el primer valor es significativo.

Estimamos un AR(1):

Dependent Variable: D_PIB
Method: ARMA Maximum Likelihood (OPG - BHHH)
Date: 03/20/20 Time: 19:19
Sample: 1851 2008
Included observations: 158
Convergence achieved after 11 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	116.5902	44.63348	2.612170	0.0099
AR(1)	0.753519	0.042423	17.76195	0.0000
SIGMASQ	17566.43	1314.948	13.35903	0.0000

R-squared	0.573753	Mean dependent var	117.8960
Adjusted R-squared	0.568253	S.D. dependent var	203.6525
S.E. of regression	133.8149	Akaike info criterion	12.65491

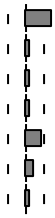
Los componentes son significativos, luego el siguiente paso es comprobar que los residuos son ruido blanco. Para obtener los residuos utilizo “correlogram squared residual”.

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.075	0.075	0.9146	0.339
		2	-0.009	-0.014	0.9267	0.629
		3	0.028	0.030	1.0543	0.788
		4	-0.023	-0.028	1.1439	0.887
		5	0.021	0.026	1.2172	0.943
		6	0.131	0.127	4.0891	0.665
		7	0.010	-0.008	4.1054	0.768
		8	0.054	0.057	4.6002	0.799
		9	0.110	0.098	6.6428	0.674
		10	0.047	0.042	7.0273	0.723
		11	-0.003	-0.014	7.0291	0.797
		12	-0.028	-0.045	7.1672	0.846
		13	-0.035	-0.029	7.3802	0.882
		14	0.006	-0.005	7.3874	0.919

De esta forma compruebo que el modelo es adecuado para continuar con la predicción. La predicción la realizare junto con los otros modelos propuestos con la intención de compararlos y saber que modelo se ajusta mejor a la realidad.

Continuando con la serie PRIMERA DIFERENCIA DEL LOGARITMO DEL PIB:

El correlograma es el siguiente:

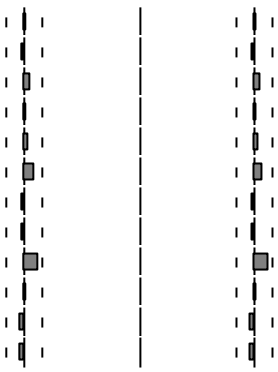
Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.223	0.223	7.9953	0.005
		2	0.039	-0.011	8.2392	0.016
		3	0.040	0.036	8.4993	0.037
		4	0.033	0.018	8.6781	0.070
		5	0.140	0.135	11.911	0.036
		6	0.079	0.019	12.953	0.044
		7	0.039	0.016	13.212	0.067

Su estructura no es tan clara como en el caso anterior por lo que probaremos varias estimaciones. Probamos varios modelos AR(p), MA(q) y ARMA(p,q) de los cuales sólo obtengo parámetros significativos en AR(1) Y MA(1).

Dependent Variable: D_LOGPIB
Method: ARMA Maximum Likelihood (OPG - BHHH)
Date: 03/20/20 Time: 19:55
Sample: 1851 2008
Included observations: 158
Convergence achieved after 5 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.018353	0.005372	3.416274	0.0008
AR(1)	0.221558	0.073704	3.006045	0.0031
SIGMASQ	0.001966	0.000113	17.46847	0.0000

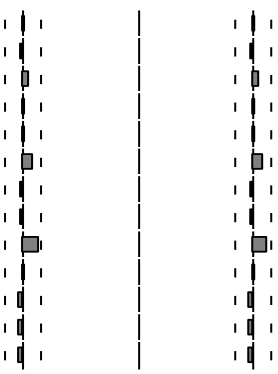
R-squared	0.049681	Mean dependent var	0.018387
Adjusted R-squared	0.037419	S.D. dependent var	0.045626

Autocorrelation	Partial Correlation
	

Dependent Variable: D_LOGPIB
Method: ARMA Maximum Likelihood (OPG - BHHH)
Date: 03/20/20 Time: 19:56
Sample: 1851 2008
Included observations: 158
Convergence achieved after 6 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.018350	0.005083	3.609870	0.0004
MA(1)	0.221411	0.075691	2.925220	0.0040
SIGMASQ	0.001967	0.000115	17.05816	0.0000

R-squared	0.049154	Mean dependent var	0.018387
Adjusted R-squared	0.036885	S.D. dependent var	0.045626

Autocorrelation	Partial Correlation
	

DOS DIFERENCIAS:

Al igual que en el caso anterior la estructura del correlograma no es clara y propongo varios modelos.

Para AR(1):

Dependent Variable: D(D(PIB))
Method: ARMA Maximum Likelihood (OPG - BHHH)
Date: 02/13/20 Time: 18:27
Sample: 1852 2008
Included observations: 157
Convergence achieved after 9 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1.563253	9.243833	0.169113	0.8659
AR(1)	-0.272707	0.067532	-4.038209	0.0001
SIGMASQ	18702.85	1395.276	13.40441	0.0000

R-squared	0.069332	Mean dependent var	0.920371
Adjusted R-squared	0.057246	S.D. dependent var	142.2146
S.E. of regression	138.0840	Akaike info criterion	12.71302
Sum squared resid	2936348.	Schwarz criterion	12.77142
Log likelihood	-994.9719	Hannan-Quinn criter.	12.73674
F-statistic	5.736312	Durbin-Watson stat	2.005616
Prob(F-statistic)	0.003955		

Inverted AR Roots -.27

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.048	-0.048	0.3711	
		2	-0.187	-0.190	6.0028	0.014
		3	-0.096	-0.120	7.4915	0.024
		4	-0.114	-0.173	9.6111	0.022
		5	0.030	-0.040	9.7610	0.045
		6	-0.036	-0.118	9.9713	0.076
		7	-0.062	-0.124	10.618	0.101
		8	-0.018	-0.105	10.671	0.154
		9	0.021	-0.064	10.746	0.217

CORRELOGRAMA RESIDUOS

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.077	0.077	0.9522	0.329
		2	0.091	0.085	2.2840	0.319
		3	0.012	-0.001	2.3068	0.511
		4	-0.006	-0.015	2.3127	0.678
		5	-0.006	-0.006	2.3188	0.804
		6	0.074	0.078	3.2358	0.779
		7	0.014	0.005	3.2701	0.859
		8	0.040	0.025	3.5378	0.896

Para AR(2):

Dependent Variable: D(D(PIB))
Method: ARMA Maximum Likelihood (OPG - BHHH)
Date: 02/13/20 Time: 18:31
Sample: 1852 2008
Included observations: 157
Convergence achieved after 11 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	2.006651	7.842539	0.255868	0.7984
AR(1)	-0.325422	0.072566	-4.484489	0.0000
AR(2)	-0.185238	0.073577	-2.517602	0.0128
SIGMASQ	18103.99	1512.409	11.97030	0.0000

R-squared	0.099132	Mean dependent var	0.920371
Adjusted R-squared	0.081468	S.D. dependent var	142.2146
S.E. of regression	136.2986	Akaike info criterion	12.69366
Sum squared resid	2842327.	Schwarz criterion	12.77153
Log likelihood	-992.4526	Hannan-Quinn criter.	12.72529
F-statistic	5.612084	Durbin-Watson stat	1.960906
Prob(F-statistic)	0.001121		

Inverted AR Roots	-.16+.40i	-.16-.40i

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.025	-0.025	0.0980	
		2	-0.068	-0.069	0.8484	
		3	-0.164	-0.168	5.1946	0.023
		4	-0.143	-0.163	8.5135	0.014
		5	0.005	-0.037	8.5174	0.036
		6	-0.060	-0.122	9.1219	0.058
		7	-0.058	-0.136	9.6915	0.084
		8	-0.038	-0.111	9.9379	0.127
		9	0.017	-0.063	9.9872	0.189

CORRELOGRAMA RESIDUOS

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.069	0.069	0.7699	0.380
		2	0.069	0.065	1.5417	0.463
		3	0.064	0.056	2.2132	0.529
		4	-0.009	-0.022	2.2271	0.694
		5	-0.006	-0.012	2.2322	0.816
		6	0.131	0.132	5.0675	0.535
		7	0.014	0.001	5.1022	0.647
		8	0.032	0.015	5.2751	0.728
		9	0.078	0.060	6.3038	0.709

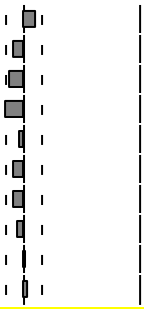
Para MA(1):

Dependent Variable: D(D(PIB))
Method: ARMA Maximum Likelihood (OPG - BHHH)
Date: 02/13/20 Time: 18:36
Sample: 1852 2008
Included observations: 157
Convergence achieved after 51 iterations
Coefficient covariance computed using outer product of gradients

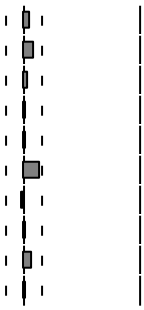
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	2.487815	5.760438	0.431880	0.6664
MA(1)	-0.511181	0.058820	-8.690655	0.0000
SIGMASQ	17779.21	1351.681	13.15341	0.0000

R-squared	0.115293	Mean dependent var	0.920371
Adjusted R-squared	0.103804	S.D. dependent var	142.2146
S.E. of regression	134.6312	Akaike info criterion	12.66381
Sum squared resid	2791336.	Schwarz criterion	12.72221
Log likelihood	-991.1089	Hannan-Quinn criter.	12.68753
F-statistic	10.03451	Durbin-Watson stat	1.701350
Prob(F-statistic)	0.000080		

Inverted MA Roots .51

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.106	0.106	1.7944	
		2	-0.092	-0.104	3.1457	0.076
		3	-0.130	-0.111	5.8894	0.053
		4	-0.162	-0.150	10.152	0.017
		5	-0.043	-0.037	10.457	0.033
		6	-0.086	-0.130	11.692	0.039
		7	-0.097	-0.133	13.266	0.039
		8	-0.051	-0.102	13.703	0.057
		9	0.016	-0.048	13.745	0.089
		10	0.034	-0.058	13.938	0.125

CORRELOGRAMA RESIDUOS

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.046	0.046	0.3392	0.560
		2	0.084	0.082	1.4768	0.478
		3	0.041	0.034	1.7451	0.627
		4	0.021	0.012	1.8202	0.769
		5	0.024	0.017	1.9136	0.861
		6	0.145	0.141	5.3784	0.496
		7	-0.015	-0.032	5.4181	0.609
		8	0.024	0.002	5.5172	0.701
		9	0.071	0.066	6.3713	0.702
		10	0.024	0.015	6.4691	0.774

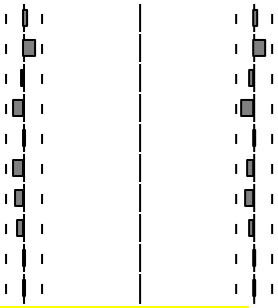
Para MA(2):

Dependent Variable: D(D(PIB))
Method: ARMA Maximum Likelihood (OPG - BHHH)
Date: 02/13/20 Time: 18:39
Sample: 1852 2008
Included observations: 157
Convergence achieved after 20 iterations
Coefficient covariance computed using outer product of gradients

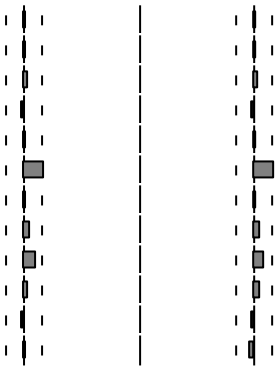
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	3.044567	2.279059	1.335888	0.1836
MA(1)	-0.494817	0.087473	-5.656815	0.0000
MA(2)	-0.308347	0.091033	-3.387210	0.0009
SIGMASQ	16404.88	1365.007	12.01817	0.0000

R-squared	0.183681	Mean dependent var	0.920371
Adjusted R-squared	0.167675	S.D. dependent var	142.2146
S.E. of regression	129.7450	Akaike info criterion	12.60001
Sum squared resid	2575567.	Schwarz criterion	12.67787
Log likelihood	-985.1005	Hannan-Quinn criter.	12.63163
F-statistic	11.47559	Durbin-Watson stat	1.857033
Prob(F-statistic)	0.000001		

Inverted MA Roots .86 -.36

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.037	0.037	0.2161	
		2	0.112	0.111	2.2386	
		3	-0.019	-0.027	2.2975	0.130
		4	-0.091	-0.103	3.6520	0.161
		5	-0.006	0.006	3.6574	0.301
		6	-0.084	-0.064	4.8342	0.305
		7	-0.075	-0.076	5.7683	0.329
		8	-0.055	-0.044	6.2829	0.392
		9	-0.003	0.015	6.2845	0.507
		10	0.016	0.010	6.3263	0.611

CORRELOGRAMA RESIDUOS

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.014	0.014	0.0294	0.864
		2	0.001	0.001	0.0296	0.985
		3	0.031	0.031	0.1856	0.980
		4	-0.016	-0.017	0.2295	0.994
		5	-0.002	-0.002	0.2302	0.999
		6	0.187	0.187	6.0372	0.419
		7	0.009	0.005	6.0497	0.534
		8	0.061	0.062	6.6811	0.571
		9	0.107	0.098	8.5991	0.475
		10	0.040	0.046	8.8715	0.544
		11	-0.013	-0.014	8.8994	0.631
		12	0.004	-0.033	8.9022	0.711

Para un ARMA(1,1):

Dependent Variable: D(D(PIB))
Method: ARMA Maximum Likelihood (OPG - BHHH)
Date: 02/13/20 Time: 18:41
Sample: 1852 2008
Included observations: 157
Convergence achieved after 11 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	2.896706	1.644609	1.761334	0.0802
AR(1)	0.467664	0.085336	5.480264	0.0000
MA(1)	-0.924097	0.036399	-25.38789	0.0000
SIGMASQ	15959.81	1176.941	13.56041	0.0000

R-squared	0.205828	Mean dependent var	0.920371
Adjusted R-squared	0.190256	S.D. dependent var	142.2146
S.E. of regression	127.9729	Akaike info criterion	12.57328
Sum squared resid	2505690.	Schwarz criterion	12.65114
Log likelihood	-983.0023	Hannan-Quinn criter.	12.60490
F-statistic	13.21785	Durbin-Watson stat	1.947001
Prob(F-statistic)	0.000000		

Inverted AR Roots	.47
Inverted MA Roots	.92

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.008	-0.008	0.0109	
		2	0.016	0.015	0.0496	
		3	0.005	0.005	0.0540	0.816
		4	-0.070	-0.070	0.8558	0.652
		5	0.045	0.043	1.1816	0.757
		6	-0.036	-0.033	1.3925	0.845
		7	-0.044	-0.046	1.7184	0.887
		8	-0.019	-0.023	1.7759	0.939
		9	0.017	0.025	1.8269	0.969
		10	0.012	0.006	1.8500	0.985

CORRELOGRAMA RESIDUOS

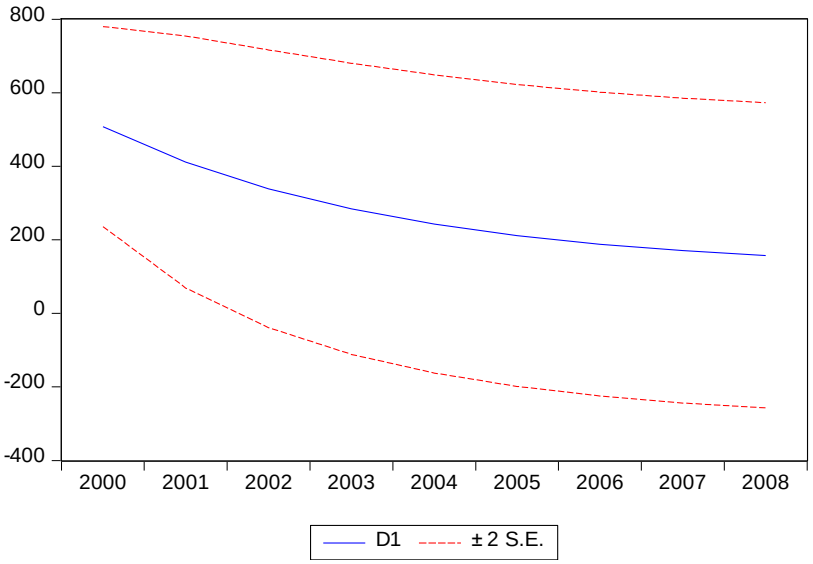
Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.006	0.006	0.0054	0.942
		2	-0.003	-0.003	0.0072	0.996
		3	0.017	0.017	0.0528	0.997
		4	-0.032	-0.032	0.2177	0.994
		5	0.011	0.011	0.2364	0.999
		6	0.140	0.139	3.4576	0.750
		7	-0.011	-0.012	3.4779	0.838
		8	0.046	0.046	3.8339	0.872
		9	0.095	0.093	5.3596	0.802

Una vez obtenidos los modelos se realizan las predicciones de cada uno de ellos.

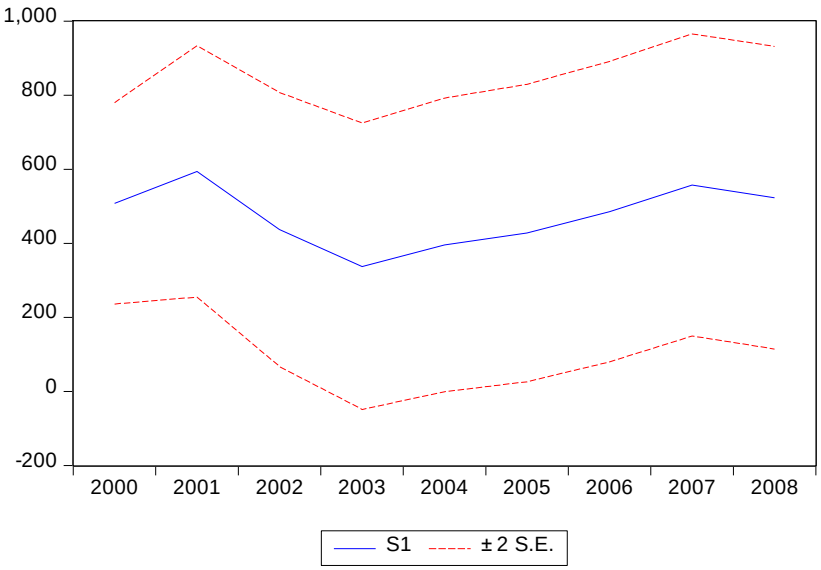
En total comparo 8 modelos diferentes.

PREDICCIÓN PRIMER MODELO:

Realizando el forecast dinámico y estático para el periodo 2000-2008 obtenemos las predicciones realizadas para esos años.



Forecast: D1	
Actual: D_PIB	
Forecast sample: 2000 2008	
Included observations: 9	
Root Mean Squared Error	308.7557
Mean Absolute Error	258.5701
Mean Abs. Percent Error	42.56200
Theil Inequality Coefficient	0.357386
Bias Proportion	0.698764
Variance Proportion	0.032975
Covariance Proportion	0.268262
Theil U2 Coefficient	1.714960
Symmetric MAPE	60.25399



Forecast: S1	
Actual: D_PIB	
Forecast sample: 2000 2008	
Included observations: 9	
Root Mean Squared Error	190.2827
Mean Absolute Error	162.8545
Mean Abs. Percent Error	46.33690
Theil Inequality Coefficient	0.182446
Bias Proportion	0.111143
Variance Proportion	0.231867
Covariance Proportion	0.656990
Theil U2 Coefficient	0.986538
Symmetric MAPE	35.00739

Representar estas predicciones ofrece poca información sobre qué modelo es mejor que otro, por eso calculare los errores mediante una tabla de Excel. El sumatorio de errores al cuadrado se obtiene de la diferencia de los datos del PIB conocidos respecto a las predicciones de cada uno de los modelos propuestos. Una vez calculados los sumatorios de los errores al cuadrado de cada una de las 8 predicciones propuestas, las comparo.

SUMATORIO CUADRADOS			Error d1	error d2	Error d3	Error d4	Error d5	Error d6	Error d7	Error d8
			857971.057	0.00212332	0.00221826	354700.902	362483.052	374713.582	394389.868	375586.622
			error s1	error s2	error s3	error s4	error s5	error s6	error s7	error s8
			190568.519	0.00165232	0.00176868	354700.902	360109.327	356336.834	304383.463	304814.385

El segundo modelo propuesto es el mejor se ajusta a la realidad