Time Series Fall 2022

Final Paper

By Simone Carugno

Introduction

Time series analysis aims at investigating and creating models to analyse the evolution, over time, of one specific variable or a set of variables in a systematic and statistical manner.[1] This paper has a three-fold objective.

First, investigating the volatility of the risk premium between South Africa 3-Year bond yields versus 10-year bonds, as well as the volatility of the risk premium between South Africa 10-year bond yields and the US Treasury bond yields with a 10-year maturity. To do so, the analysis will revolve around the class of volatility models, originally theorized by Engle and Bollerslev, called autoregressive conditional heteroskedasticity models (ARCH models) and their generalization, the generalized autoregressive conditional heteroskedasticity models (GARCH models).[2]

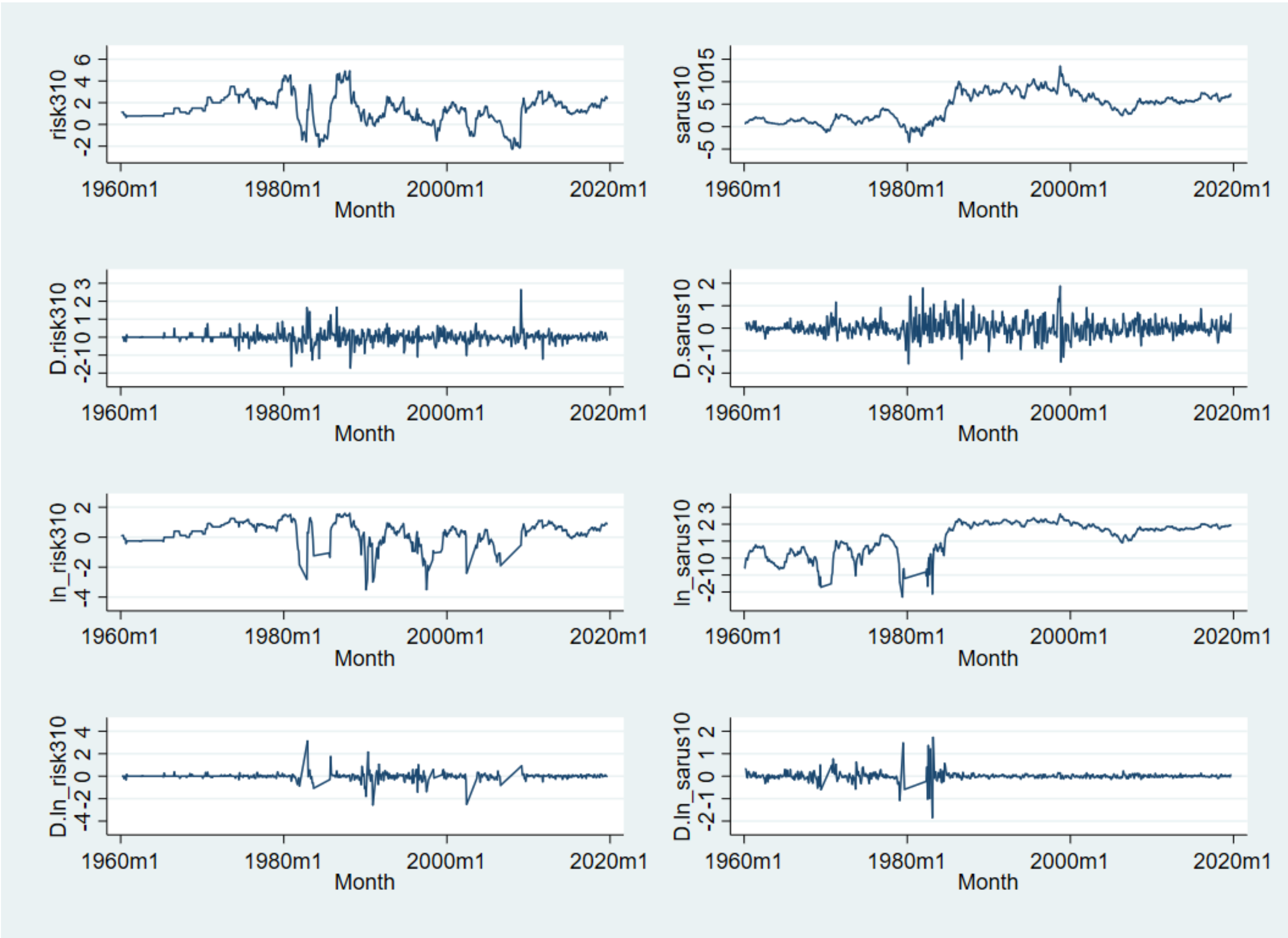
Second, the paper aims at testing the economic theory that suggests that capital intensity of production is a core determinant of the steady state real per capita GDP of a country, and, additionally, it influences the growth rate in real per capita GDP by testing for the presence of structural breaks.

Finally, the paper will estimate a structural model to specify the relationship between growth and inequality in South Africa by using Johansen's vector error correction model, which allows for the possibility of more than one independent cointegrating relation between variables that may not be stationary and treats all variables symmetrically.

[1] K. Neusser (2016), Time Series Econometrics, Springer Texts in Business and Economics, p.3.
[2] Ibid., p.167.

1.1 Volatility of Bond Yields' Risk Premia

Before conducting any formal ARCH and GARCH estimations, it is necessary to determine the scale and lag structure of the risk premia. As the graphs in the next page show, the two time series seem to be not stationary on the level, while their first difference appears to be stationary. In order to choose the appropriate scale, the time series were also log-transformed. The plotted log-transformed series appear to behave oddly, and the same can be said for the differenced log-transformed series. However, formal testing is necessary to determine the final scale, and the test used is Ermini-Hendry. Before conducting the Ermini-Hendry test, the Kwiatkowski–Phillips–Schmidt–Shin (KPSS) test and the Elliott, Rothenberg, and Stock test are performed to determine whether the series is stationary on the level and after it is first differenced and to determine the lag structure in order to test for scale.



1.2 KPSS and ERS tests

The KPSS test's null hypothesis is that the data is stationary, and the alternative hypothesis is that the data is not stationary. According to the KPSS test on the level, the risk premium of South African 10- versus 3-year bond yields is stationary because we fail to reject the null hypothesis, while its first difference is not because we do reject the null hypothesis. The KPSS test identified 18 and 14 lags, respectively, for the level and first difference series. At the same time, the risk premium of South African 10-year bond yields against US 10-year T-bond yields was flagged as stationary on the level, with a lag order of 18 lags, and not stationary in first difference, with a lag order of 2.

```
KPSS test for risk310

Automatic bandwidth selection (maxlag) = 18
Autocovariances weighted by Bartlett kernel

Critical values for H0: risk310 is trend stationary

10%: 0.119  5% : 0.146  2.5%: 0.176  1% : 0.216

Lag order    Test statistic
   18            .144
```

Figure 1 - KPSS Test SA 3-10 level

```
KPSS test for D.risk310

Automatic bandwidth selection (maxlag) = 14
Autocovariances weighted by Bartlett kernel

Critical values for H0: D.risk310 is trend stationary

10%: 0.119  5% : 0.146  2.5%: 0.176  1% : 0.216

Lag order    Test statistic
   14           .0197
```

Figure 2 - KPSS Test SA 3-10 1st Difference

```
KPSS test for D.sarus10

Automatic bandwidth selection (maxlag) = 2
Autocovariances weighted by Bartlett kernel

Critical values for H0: D.sarus10 is trend stationary

10%: 0.119  5% : 0.146  2.5%: 0.176  1% : 0.216

Lag order    Test statistic
    2            .041
```

Figure 4 - KPSS Test SA vs US 10y level

```
KPSS test for sarus10

Automatic bandwidth selection (maxlag) = 18
Autocovariances weighted by Bartlett kernel

Critical values for H0: sarus10 is trend stationary

10%: 0.119  5% : 0.146  2.5%: 0.176  1% : 0.216

Lag order    Test statistic
   18            .425
```

Figure 3 - KPSS Test SA vs US 10y Differenced

In the case of the ERS test, the null hypothesis is that the series presents a unit root (i.e., it is not stationary), while the alternative hypothesis is that there is no unit root in the series (i.e., it is stationary). For the South African 3- versus 10-year bond yields risk premium, the ERS test has similar results as the KPSS; the series is stationary on the level, but the ERS highlights the series as stationary also on the first difference, with 17 and 16 lags, respectively. On the other hand, the ERS test suggest that the risk premium of South African 10-year bond yields against US 10-year T-bond yields is not stationary on the level, but it is in first difference, with 9 and 18 lags, respectively, selected using the Ng-Perron sequence.

```
DF-GLS test for unit root                    Number of obs = 696
Variable: risk310
Lag selection: Schwert criterion             Maximum lag   =  19
```

[lags]	DF-GLS tau	Critical value 1%	Critical value 5%	Critical value 10%
19	-3.842	-3.480	-2.825	-2.541
18	-3.822	-3.480	-2.828	-2.543
17	-3.995	-3.480	-2.830	-2.545
16	-4.365	-3.480	-2.833	-2.548
15	-4.494	-3.480	-2.835	-2.550
14	-4.111	-3.480	-2.837	-2.552
13	-4.200	-3.480	-2.840	-2.554
12	-4.478	-3.480	-2.842	-2.556
11	-4.305	-3.480	-2.844	-2.558
10	-4.119	-3.480	-2.847	-2.560
9	-4.197	-3.480	-2.849	-2.562
8	-4.118	-3.480	-2.851	-2.564
7	-3.954	-3.480	-2.853	-2.566
6	-4.300	-3.480	-2.855	-2.568
5	-4.001	-3.480	-2.857	-2.570
4	-4.156	-3.480	-2.859	-2.572
3	-4.031	-3.480	-2.861	-2.574
2	-3.691	-3.480	-2.863	-2.575
1	-3.987	-3.480	-2.865	-2.577

```
Opt lag (Ng-Perron seq t) = 17 with RMSE = .3086331
Min SIC  = -2.276397 at lag  1 with RMSE = .3173968
```

Figure 5 - ERS Test SA 3vs10 Level

```
DF-GLS test for unit root                          Number of obs = 695
Variable: D.risk310
Lag selection: Schwert criterion                   Maximum lag   =  19
```

[lags]	DF-GLS tau	Critical value 1%	Critical value 5%	Critical value 10%
19	-6.550	-3.480	-2.825	-2.541
18	-6.817	-3.480	-2.828	-2.543
17	-7.067	-3.480	-2.830	-2.545
16	-6.967	-3.480	-2.833	-2.548
15	-6.533	-3.480	-2.835	-2.550
14	-6.479	-3.480	-2.837	-2.552
13	-7.285	-3.480	-2.840	-2.554
12	-7.356	-3.480	-2.842	-2.556
11	-7.092	-3.480	-2.844	-2.558
10	-7.609	-3.480	-2.847	-2.560
9	-8.275	-3.480	-2.849	-2.562
8	-8.473	-3.480	-2.851	-2.564
7	-9.064	-3.480	-2.853	-2.566
6	-10.025	-3.480	-2.855	-2.568
5	-9.779	-3.480	-2.857	-2.570
4	-11.344	-3.480	-2.859	-2.572
3	-11.909	-3.480	-2.861	-2.574
2	-13.743	-3.480	-2.863	-2.575
1	-18.054	-3.480	-2.865	-2.577

```
Opt lag (Ng-Perron seq t) = 16 with RMSE = .3131313
Min SIC  = -2.255857 at lag  1 with RMSE = .3206695
```

Figure 6 - ERS Test SA 3vs10 Differenced

```
DF-GLS test for unit root                    Number of obs = 696
Variable: sarus10
Lag selection: Schwert criterion             Maximum lag   =  19
```

		Critical value		
[lags]	DF-GLS tau	1%	5%	10%
19	-2.147	-3.480	-2.825	-2.541
18	-2.055	-3.480	-2.828	-2.543
17	-2.049	-3.480	-2.830	-2.545
16	-2.108	-3.480	-2.833	-2.548
15	-2.170	-3.480	-2.835	-2.550
14	-2.165	-3.480	-2.837	-2.552
13	-2.226	-3.480	-2.840	-2.554
12	-2.131	-3.480	-2.842	-2.556
11	-2.050	-3.480	-2.844	-2.558
10	-2.168	-3.480	-2.847	-2.560
9	-2.050	-3.480	-2.849	-2.562
8	-2.206	-3.480	-2.851	-2.564
7	-2.164	-3.480	-2.853	-2.566
6	-2.201	-3.480	-2.855	-2.568
5	-2.274	-3.480	-2.857	-2.570
4	-2.385	-3.480	-2.859	-2.572
3	-2.548	-3.480	-2.861	-2.574
2	-2.341	-3.480	-2.863	-2.575
1	-2.724	-3.480	-2.865	-2.577

```
Opt lag (Ng-Perron seq t) = 9 with RMSE = .3693119
Min SIC  =  -1.94401 at lag 2 with RMSE = .3730245
```

Figure 7 - ERS Test SAvsUS 10y Level

```
DF-GLS test for unit root                    Number of obs = 695
Variable: D.sarus10
Lag selection: Schwert criterion             Maximum lag   =  19
```

[lags]	DF-GLS tau	Critical value 1%	5%	10%
19	-3.393	-3.480	-2.825	-2.541
18	-3.596	-3.480	-2.828	-2.543
17	-3.931	-3.480	-2.830	-2.545
16	-4.157	-3.480	-2.833	-2.548
15	-4.275	-3.480	-2.835	-2.550
14	-4.393	-3.480	-2.837	-2.552
13	-4.659	-3.480	-2.840	-2.554
12	-4.805	-3.480	-2.842	-2.556
11	-5.349	-3.480	-2.844	-2.558
10	-5.998	-3.480	-2.847	-2.560
9	-6.147	-3.480	-2.849	-2.562
8	-7.124	-3.480	-2.851	-2.564
7	-7.292	-3.480	-2.853	-2.566
6	-8.266	-3.480	-2.855	-2.568
5	-9.159	-3.480	-2.857	-2.570
4	-10.117	-3.480	-2.859	-2.572
3	-11.144	-3.480	-2.861	-2.574
2	-12.168	-3.480	-2.863	-2.575
1	-16.729	-3.480	-2.865	-2.577

```
Opt lag (Ng-Perron seq t) = 18 with RMSE = .3750947
Min SIC  = -1.884046 at lag  2 with RMSE = .3843711
```

Figure 8 - ERS Test SAvsUS 10y Differenced

1.3 Ermini-Hendry Test

Ermini and Hendry studied how to determine whether models should be estimated using the linear values of variables or their logarithms.[3] Hence, to understand whether a log-transformation is justified we need to estimate the gamma and delta, after estimating lambda, in:

$$\Delta Y_t = \gamma + \sum \beta_i \Delta Y_{t-i} + \delta \exp(\lambda t)$$

If gamma is equal to zero, but the delta is not, then the log-transformation would be the cause behind the drift in the model with a normal scale.

Source	SS	df	MS			
				Number of obs	=	613
				F(0, 612)	=	0.00
Model	0	0	.	Prob > F	=	.
Residual	65.0420186	612	.106277808	R-squared	=	0.0000
				Adj R-squared	=	0.0000
Total	65.0420186	612	.106277808	Root MSE	=	.326

D.ln_risk310	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
_cons	-.0102219	.0131671	-0.78	0.438	-.0360801	.0156364

Figure 9 - Ermini-Hendry Part A

Variable	Obs	Mean	Std. dev.	Min	Max
r_ln_risk310	613	1.80e-10	.3260028	-2.569995	3.152936

Figure 10 - Ermini-Hendry Part B

[3] Ermini L., Hendry D.F. (2008), “Log Income vs. Linear Income: An Application of the Encompassing Principle”, Oxford Bulletin of Economics and Statistics, Vol. 70, Supplement.

Model	5.55669876	17	.326864633	F(17, 516)	=	3.13
Residual	53.937905	516	.104530824	Prob > F	=	0.0000
				R-squared	=	0.0934
				Adj R-squared	=	0.0635
Total	59.4946037	533	.111622146	Root MSE	=	.32331

D.risk310	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
risk310						
LD.	.2295229	.0439943	5.22	0.000	.1430929	.3159528
L2D.	-.1017633	.0449424	-2.26	0.024	-.190056	-.0134707
L3D.	.0678093	.045117	1.50	0.133	-.0208263	.1564449
L4D.	.0568898	.0450822	1.26	0.208	-.0316775	.145457
L5D.	-.0830265	.0451368	-1.84	0.066	-.171701	.0056479
L6D.	.0971675	.0452856	2.15	0.032	.0082007	.1861342
L7D.	-.0703599	.0454107	-1.55	0.122	-.1595726	.0188528
L8D.	.0239232	.045514	0.53	0.599	-.0654923	.1133387
L9D.	.022977	.0455133	0.50	0.614	-.0664371	.1123911
L10D.	-.0598437	.0454021	-1.32	0.188	-.1490394	.0293519
L11D.	.0019076	.0452772	0.04	0.966	-.0870427	.090858
L12D.	.0259552	.0452677	0.57	0.567	-.0629764	.1148869
L13D.	-.0788915	.0451921	-1.75	0.081	-.1676747	.0098917
L14D.	-.0473745	.0452527	-1.05	0.296	-.1362767	.0415277
L15D.	.0962549	.0450702	2.14	0.033	.0077113	.1847985
L16D.	-.0360568	.0441689	-0.82	0.415	-.1228298	.0507163
expLamT	-6.03e-40	1.04e-39	-0.58	0.563	-2.65e-39	1.45e-39
_cons	.0005238	.0141597	0.04	0.971	-.0272939	.0283415

Figure 11 - Ermini-Hendry Est of gamma and delta

The output shows that, for the South African 3- versus 10-year bond yields series, the transformation is not justified as both the gamma and the delta are equal to zero. This would also be in line with the behaviour of the data once plotted, and with the outcome of the KPSS and ERS tests. The Ermini-Hendry test is also conducted on the risk premium of South African 10-year bond yields against US 10-year T-bond yields. The output yields the same results, so the log-transformation is not justified.

Source	SS	df	MS			
				Number of obs	=	665
				F(0, 664)	=	0.00
Model	0	0	.	Prob > F	=	.
Residual	26.7985773	664	.040359303	R-squared	=	0.0000
				Adj R-squared	=	0.0000
Total	26.7985773	664	.040359303	Root MSE	=	.2009

D.ln_sarus10	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
_cons	.0044053	.0077904	0.57	0.572	-.0108915	.0197021

Figure 12 - Ermini-Hendry SAvsUS 10y

Variable	Obs	Mean	Std. dev.	Min	Max
r_ln_sarus10	665	2.60e-10	.2008962	-1.863304	1.730196

Figure 13 - Ermini-Hendry SAvsUS 10y

Residual	94.3389173	677	.139348475	R-squared	=	0.1169
				Adj R-squared	=	0.0921
Total	106.822685	696	.15348087	Root MSE	=	.37329

D.sarus10	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
sarus10						
LD.	.3139483	.0385376	8.15	0.000	.2382807	.3896158
L2D.	-.1979591	.0404011	-4.90	0.000	-.2772856	-.1186325
L3D.	.0970459	.0410919	2.36	0.018	.0163629	.1777288
L4D.	-.0554973	.041259	-1.35	0.179	-.1365082	.0255137
L5D.	-.0532572	.0412956	-1.29	0.198	-.1343401	.0278256
L6D.	-.0149384	.041282	-0.36	0.718	-.0959945	.0661177
L7D.	-.0469226	.041281	-1.14	0.256	-.1279767	.0341315
L8D.	.0570572	.0412511	1.38	0.167	-.0239383	.1380528
L9D.	-.104812	.0412374	-2.54	0.011	-.1857806	-.0238433
L10D.	.0711211	.0412342	1.72	0.085	-.0098411	.1520833
L11D.	-.0611726	.0412686	-1.48	0.139	-.1422023	.0198572
L12D.	.013986	.041303	0.34	0.735	-.0671113	.0950833
L13D.	.0512403	.0413058	1.24	0.215	-.0298626	.1323432
L14D.	-.0336783	.0413305	-0.81	0.415	-.1148296	.047473
L15D.	.0035123	.0413872	0.08	0.932	-.0777504	.084775
L16D.	-.0204538	.0412237	-0.50	0.620	-.1013954	.0604879
L17D.	-.0303438	.0404992	-0.75	0.454	-.109863	.0491753
L18D.	.0009769	.0386174	0.03	0.980	-.0748475	.0768012
explamT	1.18e-20	1.82e-20	0.65	0.519	-2.40e-20	4.76e-20
_cons	.0058666	.0145359	0.40	0.687	-.0226744	.0344075

Figure 14 - Ermini-Hendry SAvsUS 10y

1.4 ARCH and GARCH Estimations

An asset is risky if its return is volatile, or if it changes over time. We use the variance to measure volatility, hence the risk. ARCH models, as proposed by Engle, assume that the variance of tomorrow's return is an equally weighted average of the squared residuals from the last X time periods. Equal weights seem counterintuitive, but ARCH models treat weights as parameters to estimate and determines the best weights to use in forecasting the variance. The original model proposed by Engle in 1982 models the variance of a regression model's disturbances as a function of lagged values of the squared regression disturbances. The model can be written as:

$$\begin{aligned} y_t &= \mathbf{x}_t\beta + \epsilon_t && \text{(conditional mean)} \\ \sigma_t^2 &= \gamma_0 + \gamma_1\epsilon_{t-1}^2 + \gamma_2\epsilon_{t-2}^2 + \cdots + \gamma_m\epsilon_{t-m}^2 && \text{(conditional variance)} \end{aligned}$$

ϵ_t^2 is the squared residuals (or innovations)

γ_i are the ARCH parameters

Bollerslev in 1986 generalized the ARCH model to include lagged values of the conditional variance (i.e., GARCH model). As Neusser points out, the GARCH model allows for "parsimonious specification of the volatility process".[4] However, the condition to satisfy is that all the parameters should be positive to make sure that the variance is always positive.

$$\begin{aligned} y_t &= \mathbf{x}_t\beta + \epsilon_t \\ \sigma_t^2 &= \gamma_0 + \gamma_1\epsilon_{t-1}^2 + \gamma_2\epsilon_{t-2}^2 + \cdots + \gamma_m\epsilon_{t-m}^2 + \delta_1\sigma_{t-1}^2 + \delta_2\sigma_{t-2}^2 + \cdots + \delta_k\sigma_{t-k}^2 \end{aligned}$$

γ_i are the ARCH parameters

δ_i are the GARCH parameters

Given that, in this case, the subject of analysis are risk premia, it is worth considering also ARCH-in-mean models, which allow the conditional variance of the time series to influence the conditional mean. It is of relevance because when it comes to modelling risk-return relationships, the riskier an investment is, the lower its expected return. The ARCH-in-mean model can be written as:

$$y_t = \mathbf{x_t}\beta + \psi\sigma_t^2 + \epsilon_t$$

Given that the South African 3- versus 10-year bond yields risk premia happens to be stationary on the level, I will conduct the ARCH estimations on the level. The series presents ARCH disturbance, as well as serial autocorrelation.

[4] K. Neusser (2016), Time Series Econometrics, Springer Texts in Business and Economics, p. 174.

```
. estat archlm, l(16)
LM test for autoregressive conditional heteroskedasticity (ARCH)
```

lags(p)	chi2	df	Prob > chi2
16	583.304	16	0.0000

H0: no ARCH effects vs. H1: ARCH(p) disturbance

Breusch-Godfrey LM test for autocorrelation

lags(p)	chi2	df	Prob > chi2
16	652.266	16	0.0000

H0: no serial correlation

The first-order generalized ARCH model (i.e., a GARCH 1,1) is the most widely used model for the conditional variance in empirical work. In this situation, the GARCH(1,1) model estimated on the South African 3- versus 10-year bond yields risk premium meets all of Bollerslev conditions (i.e., it has all coefficients greater than zero), and the ARCH(1) and GARCH(1) coefficients are significant collectively:

$$[1]\; y_t = 0.0941 + \varepsilon_t$$
$$[2]\; \sigma_t^2 = 0.0023 + 0.4086\varepsilon_{t-1}^2 + 0.6835\sigma_{t-1}^2$$

_cons	.0941086	.0080025	11.76	0.000	.078424	.1097933
ARCH						
arch						
L1.	.4086002	.056114	7.28	0.000	.2986188	.5185815
garch						
L1.	.(				6174372	.7497218
_cons	.(				0016454	.003111

```
( 1)  [ARCH]L.arch + [ARCH]L.garch = 1

          chi2(  1) =     8.62
        Prob > chi2 =    0.0033
```

Given the nature of the series, an ARCH-in-mean model was also fitted. If we specify ψ as:

$$\psi(\sigma_t^2) = \delta_0 + \delta_1\sigma_t^2$$

Then, given that higher volatility requires higher return, we expect δ_1 to be positive.

_cons	.0933064	.0101505	9.19	0.000	.0734118	.1132011
ARCHM						
sigma2	-.1036433	.0341447	-3.04	0.002	-.1705657	-.0367209
ARCH						
arch						
L1.	1.410059	.1427404	9.88	0.000	1.130293	1.689825
_cons	.0272901	.0022185	12.30	0.000	.022942	.0316383

The ARCHM model can be expressed as:

$$y_t = 0.0933 - 0.1036\sigma_t^2 + \varepsilon_t$$

However, given that the coefficient on sigma squared is negative, that would imply that the reward decreases with the risk, which is unrealistic. Therefore between the GARCH(1,1) and the ARCHM models, it is seems preferrable to use the GARCH(1,1) model. Then we conduct the same estimations on the second series. Given the mixed results yield by the unit root tests performed, the estimations will be conducted on both the level and the differenced series. Given that a first simple test with a GARCH(1,1) model did not meet the Bollerslev conditions, I plotted the ACF and PACF for both the series on the level and its first difference:

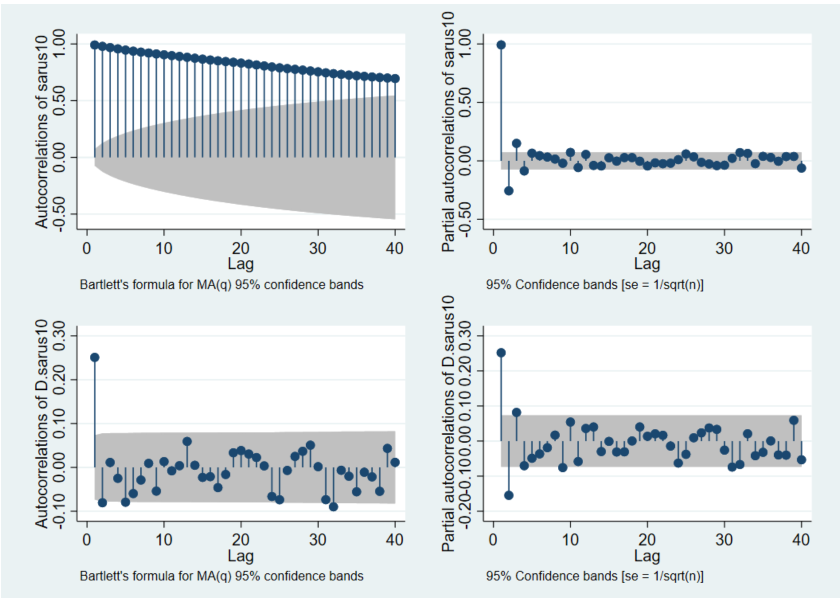


Figure 15 - ACF & PACF plots

As the plots show, the series seems to be an autoregressive process of order 3, which fits well a simple ARCH model. This is in line with the estimations conducted. The optimal model that meets the Engle conditions was an ARCH(3) on the level with 5 lags. No GARCH model was able to meet the Bollerslev conditions. The optimal model can be expressed as:

$$y_t = 0.0393 + \varepsilon_t$$

$$\sigma_t^2 = 0.0544 + 0.3230\varepsilon_{t-1}^2 + 0.2129\varepsilon_{t-2}^2 + 0.1216\varepsilon_{t-3}^2$$

```
ARCH family regression

Sample: 1960m7 thru 2019m10                     Number of obs     =        712
                                                Wald chi2(4)      =   88866.46
Log likelihood = -233.8863                      Prob > chi2       =     0.0000
```

L.sarus10	Coefficient	OPG std. err.	z	P>|z|	[95% conf.	interval]
sarus10						
sarus10						
L2.	1.274017	.0420601	30.29	0.000	1.19158	1.356453
L3.	-.4440399	.0678134	-6.55	0.000	-.5769517	-.311128
L4.	.2362399	.0682485	3.46	0.001	.1024752	.3700046
L5.	-.0762399	.036832	-2.07	0.038	-.1484292	-.0040506
_cons	.0393766	.0198834	1.98	0.048	.0004059	.0783473
ARCH						
arch						
L1.	.3230186	.0586593	5.51	0.000	.2080485	.4379886
L2.	.2129007	.0509882	4.18	0.000	.1129657	.3128357
L3.	.1216661	.0571135	2.13	0.033	.0097256	.2336066
_cons	.0544073	.0051019	10.66	0.000	.0444077	.0644069

Other estimations that do not meet either the Engle or the Bollerslev conditions are presented here:

_cons	-.0018864	.012436	-0.15	0.879	-.0262605	.0224877
ARCH						
arch						
L1.	.4473894	.0717699	6.23	0.000	.3067229	.5880558
_cons	.081617	.0050371	16.20	0.000	.0717445	.0914895

Figure 16 - ARCH(1) model

_cons	-.002751	.0116579	-0.24	0.813	-.0256	.0200981
ARCH						
arch						
L1.	.1424338	.0254622	5.59	0.000	.0925287	.1923389
garch						
L1.	.8481592	.0255551	33.19	0.000	.7980721	.8982462
_cons	.0025506	.0009303	2.74	0.006	.0007273	.004374

Figure 17 - GARCH(1,1) Model

_cons	.018988	.01783	1.06	0.287	-.0159582	.0539342
ARCH						
arch						
L2.	.1699975	.0240846	7.06	0.000	.1227925	.2172025
garch						
L2.	.8235397	.0220393	37.37	0.000	.7803435	.8667359
_cons	.0031696	.0010189	3.11	0.002	.0011726	.0051667

Figure 18 - GARCH(2,2) Model

2.1 Structural Breaks & GDP

Unit-root tests are heavily reliant on the appropriate specification of the deterministic component. To achieve a correct specification of the deterministic component, it is necessary to test for the presence of structural breaks. If structural breaks are ignored, it is possible to incur in the risk of favouring the null hypothesis (i.e., that the unit root is present, in the null hypothesis for most tests). Although it is often possible to be aware of specific events that may correspond to sudden movements in time series, it is important to generalize the test for structural breaks.

The first step is to determine the scale of the variables we will use when testing for structural breaks. To do so, information criteria were used. The outcome of testing whether a log-transformation is necessary is that the information criteria do not justify a log-transformation of any of the variables.

Akaike's information criterion and Bayesian information criterion

Model	N	ll(null)	ll(model)	df	AIC	BIC
.	64	-781.4429	-772.9827	11	1567.965	1591.713

Note: BIC uses N = number of observations. See [R] BIC note.

Figure 19 - IC D.Rgdp

Akaike's information criterion and Bayesian information criterion

Model	N	ll(null)	ll(model)	df	AIC	BIC
.	64	142.7792	155.5586	11	-289.1172	-265.3695

Note: BIC uses N = number of observations. See [R] BIC note.

Figure 120 - IC D.ln_Rgdp

Akaike's information criterion and Bayesian information criterion

Model	N	ll(null)	ll(model)	df	AIC	BIC
.	64	-652.2558	-642.5201	11	1307.04	1330.788

Note: BIC uses N = number of observations. See [R] BIC note.

Figure 21 - IC D.CapitGov

Akaike's information criterion and Bayesian information criterion

Model	N	ll(null)	ll(model)	df	AIC	BIC
.	64	55.58916	63.64692	11	-105.2938	-81.54612

Note: BIC uses N = number of observations. See [R] BIC note.

Figure 22 - IC D.ln_CapitGov

Akaike's information criterion and Bayesian information criterion

Model	N	ll(null)	ll(model)	df	AIC	BIC
.	64	-664.1553	-655.4346	11	1332.869	1356.617

Note: BIC uses N = number of observations. See [R] BIC note.

Figure 23 - IC D.CapitPubCorp

Akaike's information criterion and Bayesian information criterion

Model	N	ll(null)	ll(model)	df	AIC	BIC
.	64	23.59848	30.03299	11	-38.06598	-14.31827

Note: BIC uses N = number of observations. See [R] BIC note.

Figure 24 - IC D.ln_CapitPubCorp

Akaike's information criterion and Bayesian information criterion

Model	N	ll(null)	ll(model)	df	AIC	BIC
.	64	-718.5178	-711.8196	11	1445.639	1469.387

Note: BIC uses N = number of observations. See [R] BIC note.

Figure 25 - IC D.CapPrivCorp

Akaike's information criterion and Bayesian information criterion

Model	N	ll(null)	ll(model)	df	AIC	BIC
.	64	67.67715	80.00426	11	-138.0085	-114.2608

Note: BIC uses N = number of observations. See [R] BIC note.

Figure 13 - IC D.ln_CapPrivCorp

Once determined the scale of the variables, they are then fitted in a regression model.

. reg Rgdp Population capitgov cappubcorp capPrivCorp

Source	SS	df	MS			
				Number of obs	=	75
				F(4, 70)	=	7641.55
Model	5.5910e+13	4	1.3978e+13	Prob > F	=	0.0000
Residual	1.2804e+11	70	1.8291e+09	R-squared	=	0.9977
				Adj R-squared	=	0.9976
Total	5.6038e+13	74	7.5727e+11	Root MSE	=	42769

Rgdp	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
Population	.0372876	.001109	33.62	0.000	.0350758	.0394994
capitgov	3.75794	.3366646	11.16	0.000	3.086483	4.429396
cappubcorp	1.822279	.4252716	4.28	0.000	.9741014	2.670456
capPrivCorp	1.674663	.1836266	9.12	0.000	1.308432	2.040895
_cons	-271879.4	22767.55	-11.94	0.000	-317287.8	-226470.9

All variables seem to be statistically significant; however, to better isolate and understand the impact of the capital formation of the government, public corporations, and private corporations, the population is removed from the regression and the model is refitted. Before doing so, the optimal lag structure is identified through Ng-Perron and Schwert Criterion, and the variables are tested for the presence of a unit root by conducting a Phillips-Perron test, KPSS test, and ERS test. The variables pass the KPSS test, however, they do not pass the Phillips-Perron test [see Appendix A]. Once first differenced, they do succeed in passing the Phillips-Perron test and reject the null hypothesis of the presence of a unit root.

Source	SS	df	MS
Model	5.3842e+13	3	1.7947e+13
Residual	2.1959e+12	71	3.0929e+10
Total	5.6038e+13	74	7.5727e+11

Number of obs	=	75
F(3, 71)	=	580.29
Prob > F	=	0.0000
R-squared	=	0.9608
Adj R-squared	=	0.9592
Root MSE	=	1.8e+05

Rgdp	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
capitgov	2.279625	1.372515	1.66	0.101	-.4570924	5.016342
cappubcorp	-1.128975	1.711072	-0.66	0.512	-4.540757	2.282806
capPrivCorp	6.744994	.4308501	15.66	0.000	5.885903	7.604085
_cons	310277	60794.23	5.10	0.000	189056.7	431497.2

As the output shows, the government's capital formation and the public corporations' capital formation do not seem to be statistically significant anymore. However, refitting the regression after differencing the variables and setting the correct lag structure as established by Ng-Perron, the public corporations' capital formation appears to be statistically significant while the government's capital formation still is not.

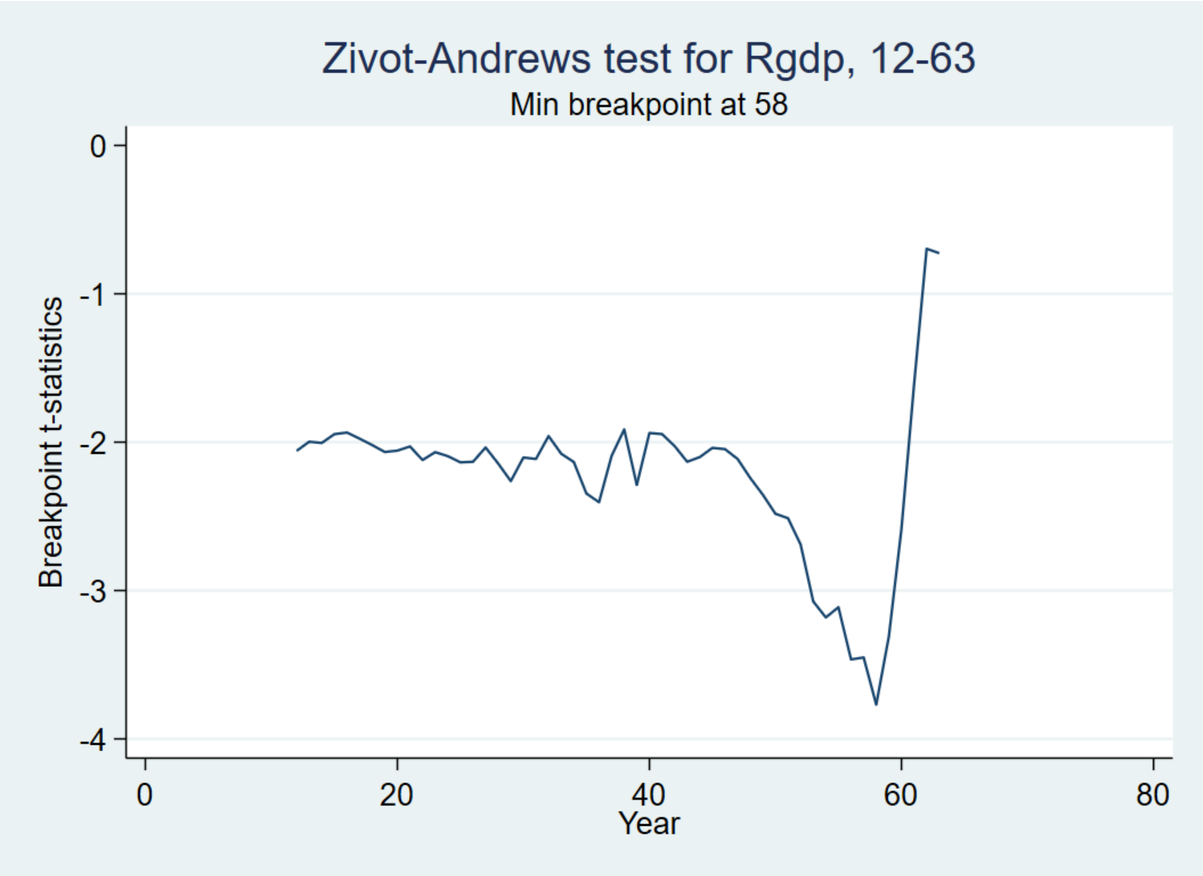
Source	SS	df	MS
Model	4.4724e+10	3	1.4908e+10
Residual	1.0612e+11	59	1.7987e+09
Total	1.5085e+11	62	2.4330e+09

Number of obs	=	63
F(3, 59)	=	8.29
Prob > F	=	0.0001
R-squared	=	0.2965
Adj R-squared	=	0.2607
Root MSE	=	42411

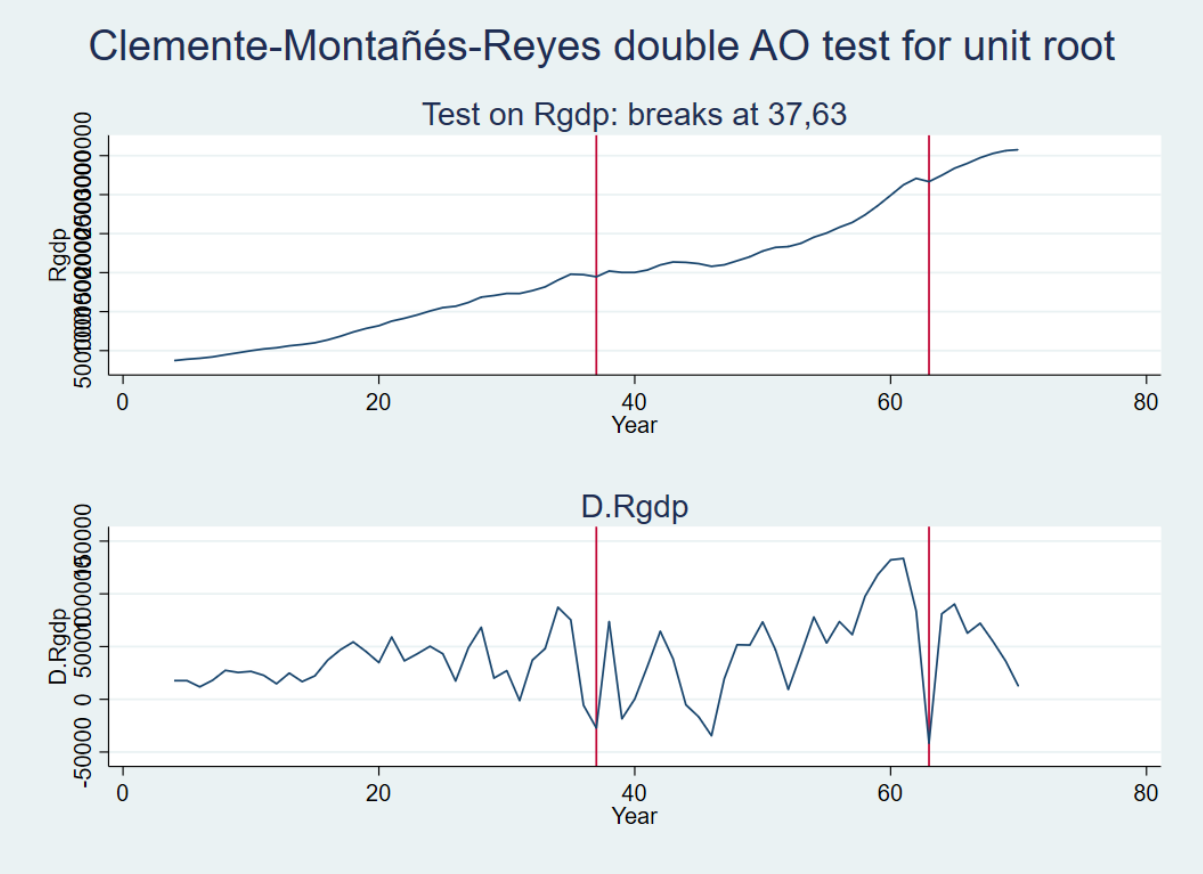
D.Rgdp	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
capitgov						
L8D.	.5385192	.884374	0.61	0.545	-1.231109	2.308147
cappubcorp						
L11D.	-3.304789	.8327743	-3.97	0.000	-4.971167	-1.638412
capPrivCorp						
L11D.	1.404392	.3961045	3.55	0.001	.611789	2.196995
_cons	36509.78	5952.769	6.13	0.000	24598.32	48421.25

2.2 Structural Breaks

After the presented evaluations and estimations, the series was tested for the presence of structural breaks. First, the test for one structural break was conducted using the Zivot-Andrews test. The test highlights a potential structural break in the year 2004:



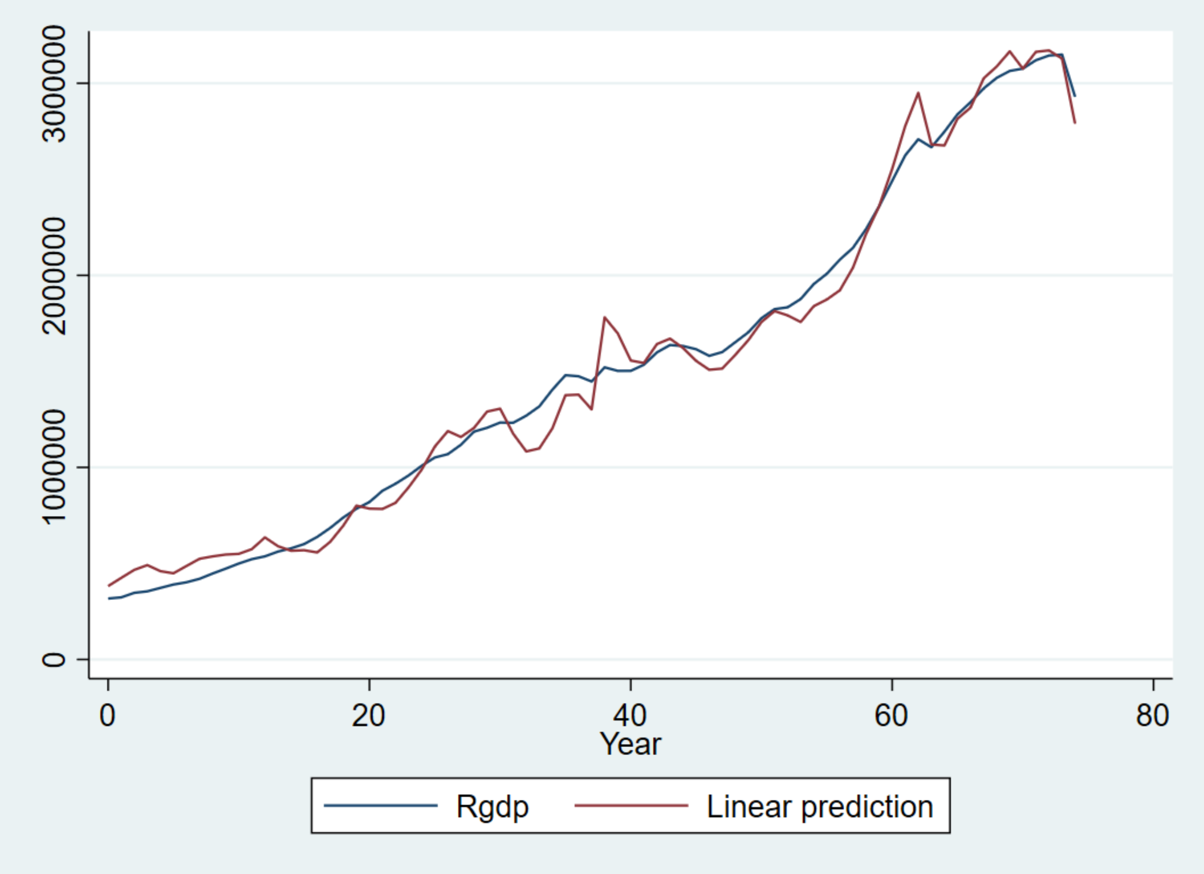
The test was also repeated using Clemente-Montañes-Reyes. The test was conducted to identify 2 potential structural breaks. The test reports two structural breaks at the year 1983 and the year 2009. Both years are likely to be significantly related with the behaviour of RGDP. The year 1983 was in the midst of the early 1980s recession, considered the most severe recession since World War II at the time, caused by the 1979 energy crisis. On the other hand, 2009 is characterized by the recession caused by the most impactful global financial crisis since the 1929 recession.

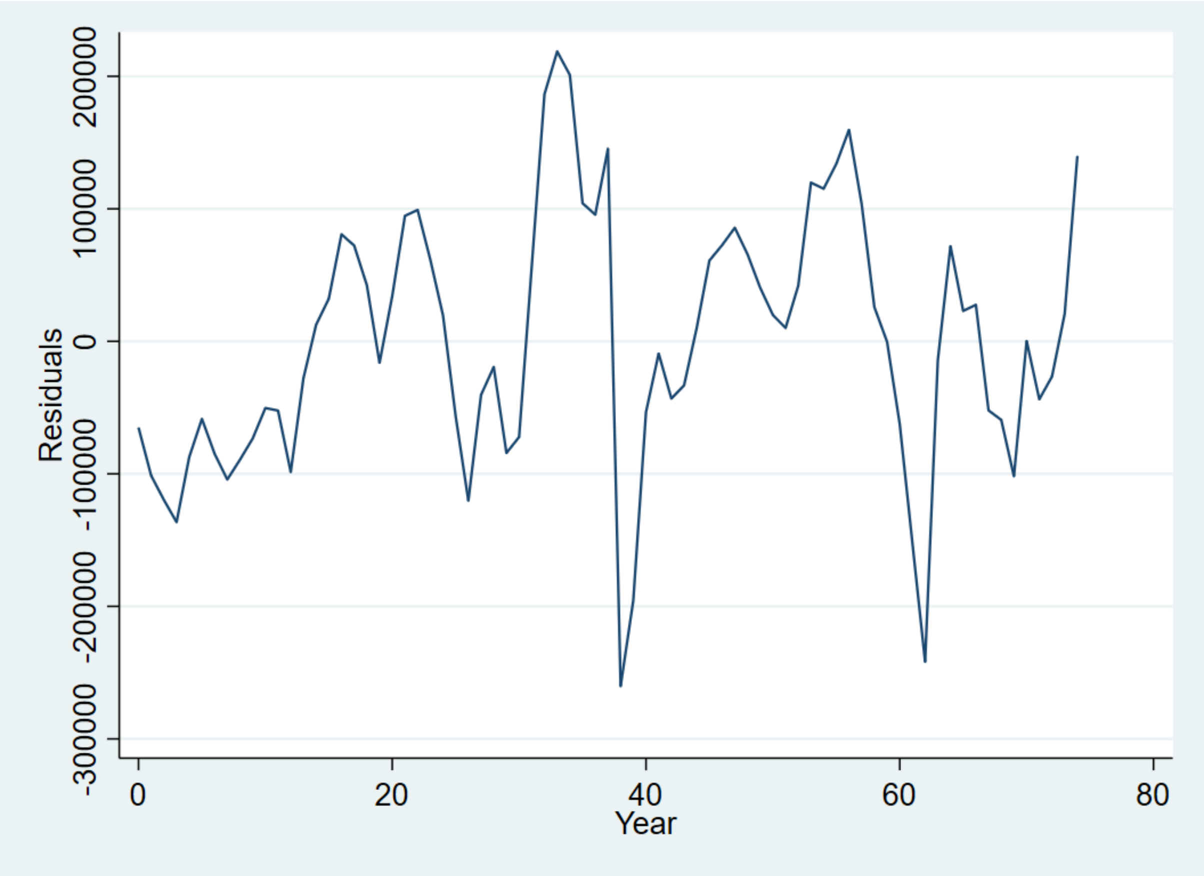


After identifying the two structural breaks, a regression model is refitted to take them into account:

Source	SS	df	MS			
				Number of obs	=	75
				F(5, 69)	=	1099.55
Model	5.5344e+13	5	1.1069e+13	Prob > F	=	0.0000
Residual	6.9460e+11	69	1.0067e+10	R-squared	=	0.9876
				Adj R-squared	=	0.9867
Total	5.6038e+13	74	7.5727e+11	Root MSE	=	1.0e+05

Rgdp	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
capitgov	7.013529	.8908546	7.87	0.000	5.236323	8.790735
cappubcorp	-.3551209	1.061637	-0.33	0.739	-2.473029	1.762787
capPrivCorp	4.118106	.3328776	12.37	0.000	3.454033	4.782178
D37	516009	43173.03	11.95	0.000	429881.1	602136.8
D63	81123.11	56814.55	1.43	0.158	-32218.82	194465
_cons	177832.5	36511.62	4.87	0.000	104993.8	250671.2


3000000
2000000
1000000
0
0
20
40
60
80
Year
Rgdp
Linear prediction


Residuals
200000
100000
0
-100000
-200000
-300000
0
20
40
60
80
Year

3.1 VECM, Growth and Inequality in South Africa

In this final section, the aim is to investigate the relationship between growth and inequality in South Africa. The dataset used contains the following variables: Gini coefficient, index of property rights in South Africa, total employment rate, private sector employment rate, public sector employment rate, openness of the economy, household credit as a proportion of GDP, corporate sector credit as a proportion of GDP, real GDP, gross fixed capital formation as a proportion of GDP, government expenditure as a proportion of GDP. Before creating a vector error correction model and estimating its coefficients, it is necessary to correctly identify the lag structure of the model. First, after specifying the scale of each variable, all variables were used to estimate the correct lag structure through the use of information criteria. However, since the software was unable to allocate a matrix using all the variables, the estimation was conducted again using the Gini coefficient, public sector employment rate, private sector employment rate, openness of the economy, RGDP, gross fixed capital formation as a proportion of GDP, and government expenditure as a proportion of the GDP. The maximum lag length was specified as 6 lags because when estimating for the number of cointegrating equations in the next step any lag structure greater than 6 would result in multicollinearity and failure to identify the number of equations.

Lag-order selection criteria

Sample: 6 thru 53 Number of obs = 48

Lag	LL	LR	df	p	FPE	AIC	HQIC	SBIC
0	-735.081				63256.6	30.9201	31.0232	31.1929
1	-340.323	789.52	49	0.000	.035804	16.5135	17.3385	18.6965*
2	-274.557	131.53	49	0.000	.020264	15.8149	17.3617	19.9081
3	-209.826	129.46	49	0.000	.015152	15.1594	17.4281	21.1628
4	-156.414	106.82	49	0.000	.028658	14.9756	17.9662	22.8892
5	-68.4304	175.97*	49	0.000	.033625	13.3513*	17.0637*	23.1751
6	.	.	49	.	-3.6e-32*	.	.	.

* optimal lag
Endogenous: gini ln_privsecemprate ln_pubsecemprate open rgdp ln_investgdp ln_govexpgdp
Exogenous: _cons

The selected number of lags is 1 in accordance with the Schwarz's Bayesian information criterion (SBIC). Next, the number of cointegrating equations was estimated under no trend and no constant, as well as under unrestricted constant but no trend to allow for the cointegrating or long-run relationship to be stationary around a constant mean without including a trend in the relationship.

Johansen tests for cointegration
Trend: <none> Number of obs = 53
Sample: 1 thru 53 Number of lags = 1

Maximum rank	Params	LL	Eigenvalue	Trace statistic	Critical value 5%
0	0	-495.50697	.	214.2565	109.99
1	13	-442.86998	0.86280	108.9826	82.49
2	24	-421.65712	0.55089	66.5568	59.46
3	33	-407.3133	0.41799	37.8692*	39.89
4	40	-399.40495	0.25802	22.0525	24.31
5	45	-392.46183	0.23049	8.1663	12.53
6	48	-388.38556	0.14258	0.0137	3.84
7	49	-388.37871	0.00026		

Maximum rank	Params	LL	Eigenvalue	Eigenvalue Maximum	Critical value 5%
0	0	-495.50697	.	105.2740	41.51
1	13	-442.86998	0.86280	42.4257	36.36
2	24	-421.65712	0.55089	28.6876	30.04
3	33	-407.3133	0.41799	15.8167	23.80
4	40	-399.40495	0.25802	13.8862	17.89
5	45	-392.46183	0.23049	8.1525	11.44
6	48	-388.38556	0.14258	0.0137	3.84
7	49	-388.37871	0.00026		

* selected rank

No trend and no constant

Johansen tests for cointegration
Trend: Constant Number of obs = 53
Sample: 1 thru 53 Number of lags = 1

Maximum rank	Params	LL	Eigenvalue	Trace statistic	Critical value 5%
0	7	-460.36957	.	169.9961	124.24
1	20	-430.06035	0.68138	109.3777	94.15
2	31	-408.96561	0.54888	67.1882*	68.52
3	40	-398.15799	0.33491	45.5730	47.21
4	47	-387.75348	0.32472	24.7640	29.68
5	52	-380.17956	0.24859	9.6161	15.41
6	55	-376.57425	0.12720	2.4055	3.76
7	56	-375.3715	0.04437		

Maximum rank	Params	LL	Eigenvalue	Eigenvalue Maximum	Critical value 5%
0	7	-460.36957	.	60.6184	45.28
1	20	-430.06035	0.68138	42.1895	39.37
2	31	-408.96561	0.54888	21.6152	33.46
3	40	-398.15799	0.33491	20.8090	27.07
4	47	-387.75348	0.32472	15.1478	20.97
5	52	-380.17956	0.24859	7.2106	14.07
6	55	-376.57425	0.12720	2.4055	3.76
7	56	-375.3715	0.04437		

* selected rank

Unrestricted constant, but no trend

The selected rank for the estimation of the parameters and their short-term adjustment coefficients (alpha coefficients) and the long run coefficients (beta coefficients) is 2. Given the chosen rank (r), at least r^2 = 4 restrictions must be imposed. To identify the impact that employment rates have on inequality and the impact of capital investments and government expenditures have, respectively, the model was overidentified with 6 restrictions.

```
Sample: 1 thru 53                                Number of obs     =        53
                                                 AIC               =  16.64193
Log likelihood = -413.0113                       HQIC              =  17.04222
Det(Sigma_ml)  =  .0138453                       SBIC              =  17.68284
```

Equation	Parms	RMSE	R-sq	chi2	P>chi2
D_gini	3	1.14424	0.1765	10.50209	0.0147
D_ln_pubsecemp~e	3	.022756	0.3371	24.9157	0.0000
D_ln_privsecem~e	3	.029902	0.2338	14.95241	0.0019
D_open	3	4.28231	0.0400	2.04247	0.5636
D_rgdp	3	35198.6	0.6650	97.28686	0.0000
D_ln_investgdp	3	.06406	0.0600	3.127824	0.3723
D_ln_govexpgdp	3	.046328	0.1392	7.923437	0.0476

	Coefficient	Std. err.	z	P>|z|	[95% conf.	interval]
D_gini						
_ce1 L1.	-.0208608	.0072084	-2.89	0.004	-.0349889	-.0067326
_ce2 L1.	-.0048533	.0041098	-1.18	0.238	-.0129083	.0032017
_cons	.3245021	.4655655	0.70	0.486	-.5879895	1.236994

D_ln_pubsecemprate						
_ce1						
L1.	.0007075	.0001434	4.93	0.000	.0004265	.0009884
_ce2						
L1.	-.0000125	.0000817	-0.15	0.878	-.0001727	.0001477
_cons	-.0219777	.009259	-2.37	0.018	-.040125	-.0038304
D_ln_privsecemprate						
_ce1						
L1.	.0002845	.0001884	1.51	0.131	-.0000847	.0006537
_ce2						
L1.	.0003911	.0001074	3.64	0.000	.0001806	.0006016
_cons	.0274527	.0121666	2.26	0.024	.0036067	.0512987
D_open						
_ce1						
L1.	-.0317953	.0269773	-1.18	0.239	-.0846699	.0210793
_ce2						
L1.	-.013491	.0153809	-0.88	0.380	-.043637	.016655
_cons	-.208488	1.742384	-0.12	0.905	-3.623498	3.206522

D_rgdp						
_ce1						
L1.	49.81897	221.7411	0.22	0.822	-384.7856	484.4236
_ce2						
L1.	-437.9808	126.4237	-3.46	0.001	-685.7667	-190.1948
_cons	-.0000543	14321.59	-0.00	1.000	-28069.79	28069.79
D_ln_investgdp						
_ce1						
L1.	-.0000116	.0004036	-0.03	0.977	-.0008026	.0007794
_ce2						
L1.	-.000338	.0002301	-1.47	0.142	-.000789	.0001129
_cons	-.0245901	.0260646	-0.94	0.345	-.0756758	.0264957
D_ln_govexpgdp						
_ce1						
L1.	.0005992	.0002919	2.05	0.040	.0000271	.0011712
_ce2						
L1.	.0001399	.0001664	0.84	0.400	-.0001862	.000466
_cons	.007923	.0188499	0.42	0.674	-.0290221	.0448681

Cointegrating equations

Equation	Parms	chi2	P>chi2
_ce1	4	51.49352	0.0000
_ce2	4	111.1268	0.0000

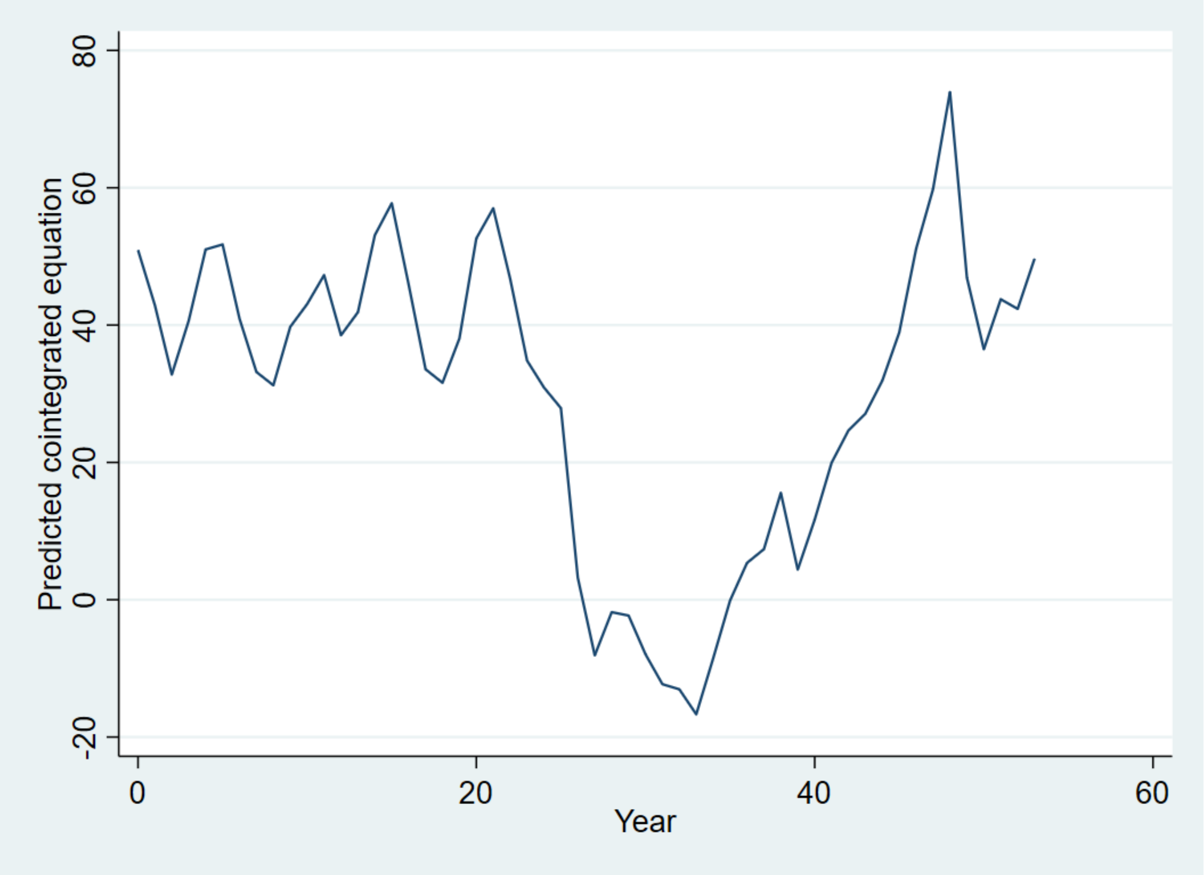
Identification: beta is overidentified

(1) [_ce1]gini = 1
(2) [_ce1]ln_pubsecemprate = 0
(3) [_ce1]ln_privsecemprate = 0
(4) [_ce2]gini = 1
(5) [_ce2]ln_investgdp = 0
(6) [_ce2]ln_govexpgdp = 0

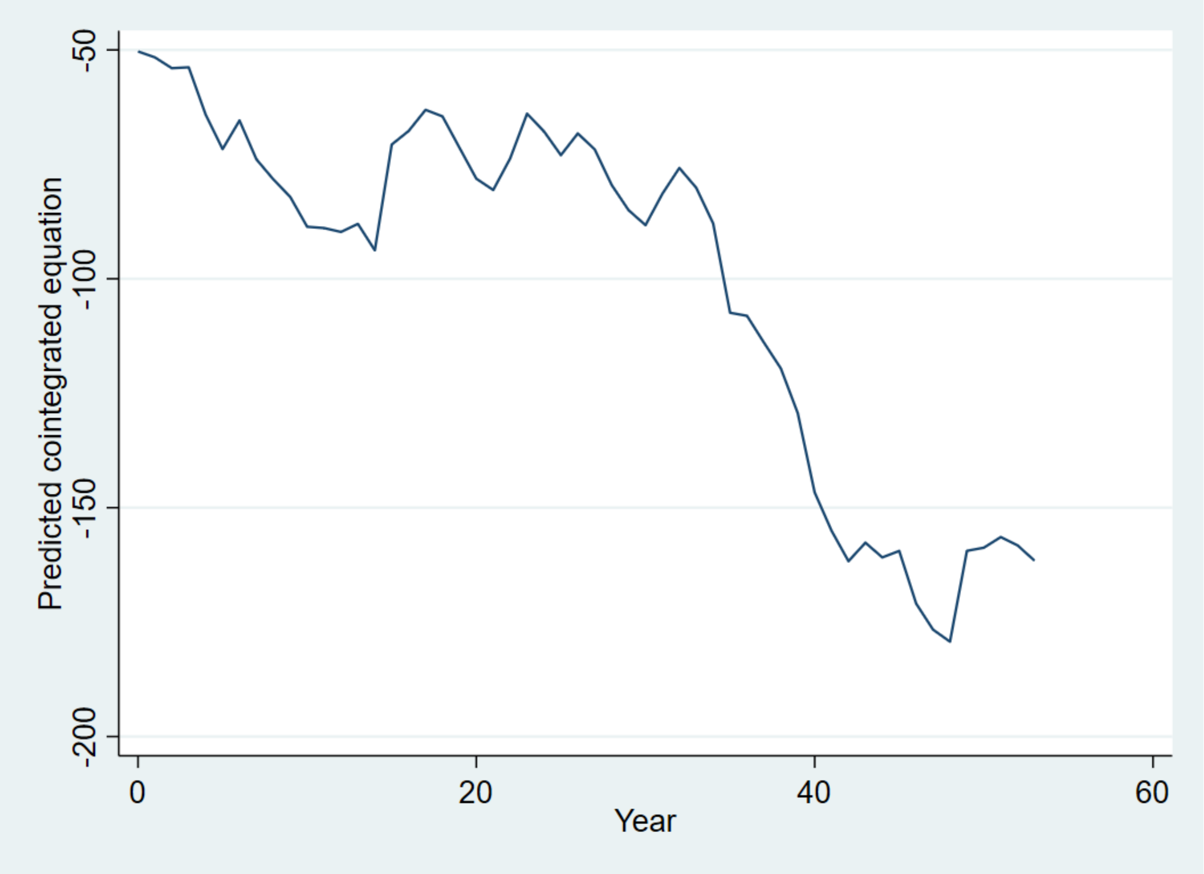
beta	Coefficient	Std. err.	z	P>|z|	[95% conf.	interval]
_ce1						
gini	1	.	.	.	.	.
ln_pubsecemprate	0	(omitted)				
ln_privsecemprate	0	(omitted)				
open	1.006204	.6706437	1.50	0.134	-.3082331	2.320642
rgdp	-7.08e-06	9.89e-06	-0.72	0.474	-.0000265	.0000123
ln_investgdp	91.71492	23.68039	3.87	0.000	45.30221	138.1276
ln_govexpgdp	-74.77259	34.67323	-2.16	0.031	-142.7309	-6.814312
_cons	335.7035	.	.	.	.	.
_ce2						
gini	1	.	.	.	.	.
ln_pubsecemprate	221.4777	52.42945	4.22	0.000	118.7178	324.2375
ln_privsecemprate	-63.75885	48.2756	-1.32	0.187	-158.3773	30.85959
open	-.5851244	.7071699	-0.83	0.408	-1.971152	.8009031
rgdp	-.0000537	8.03e-06	-6.69	0.000	-.0000694	-.0000379
ln_investgdp	0	(omitted)				
ln_govexpgdp	0	(omitted)				
_cons	540.3506	.	.	.	.	.

LR test of identifying restrictions: chi2(3) = 8.096 Prob > chi2 = 0.044

After imposing the restriction and estimating the parameters, the output shows that, in the long run, the gross capital formation as a proportion of GDP has a statistically significant and negative impact on the Gini coefficient, while government expenditure as a proportion of GDP has a statistically significant but positive impact on the Gini coefficient. On the other hand, the RGDP has a statistically significant and positive impact on Gini coefficient of South Africa, while the public sector employment rate has a statistically significant and negative impact on the Gini coefficient. Although the relationship between government expenditure, RGDP, and Gini is reasonable, the relationship between public sector employment, capital formation, and Gini is less so given that a higher employment rate and gross capital formation lead to a higher volume of money circulating into the economy and boost the income of the population.

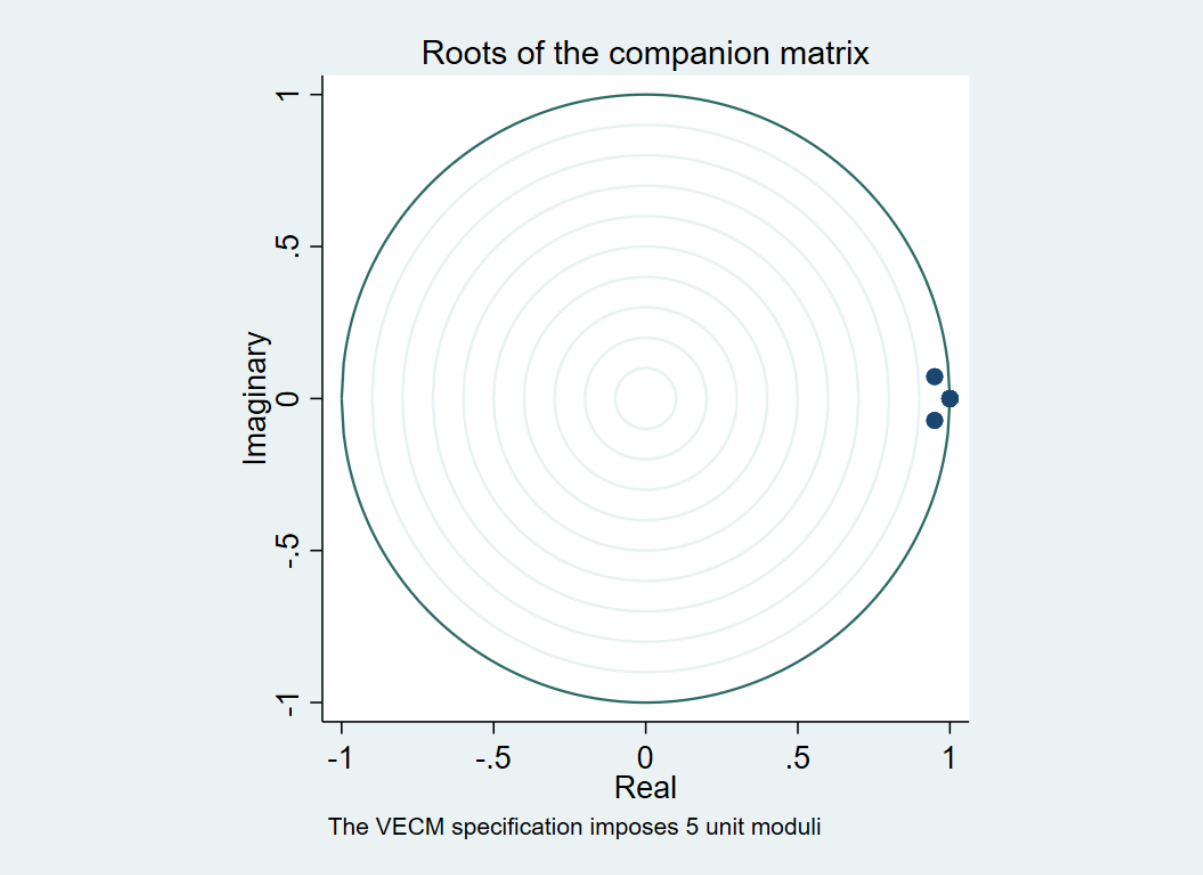


Cointegrating equation 1

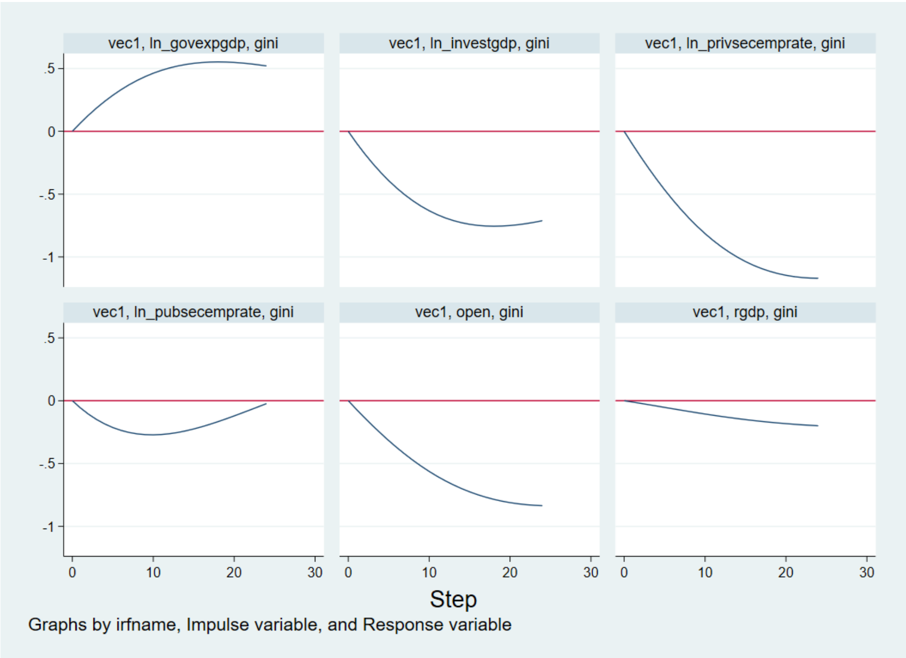


Cointegration equation 2

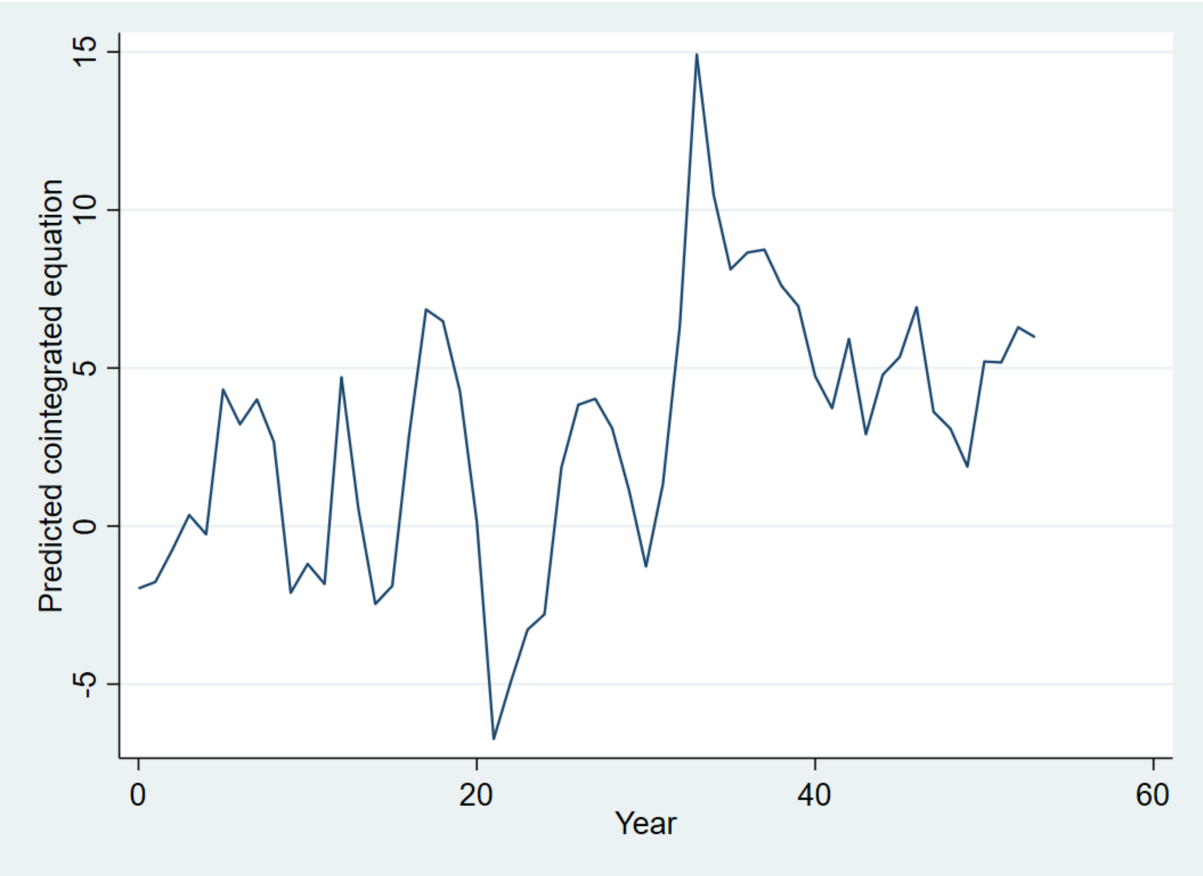
The model is also tested for stability. However, the test shows that the model is not stable.



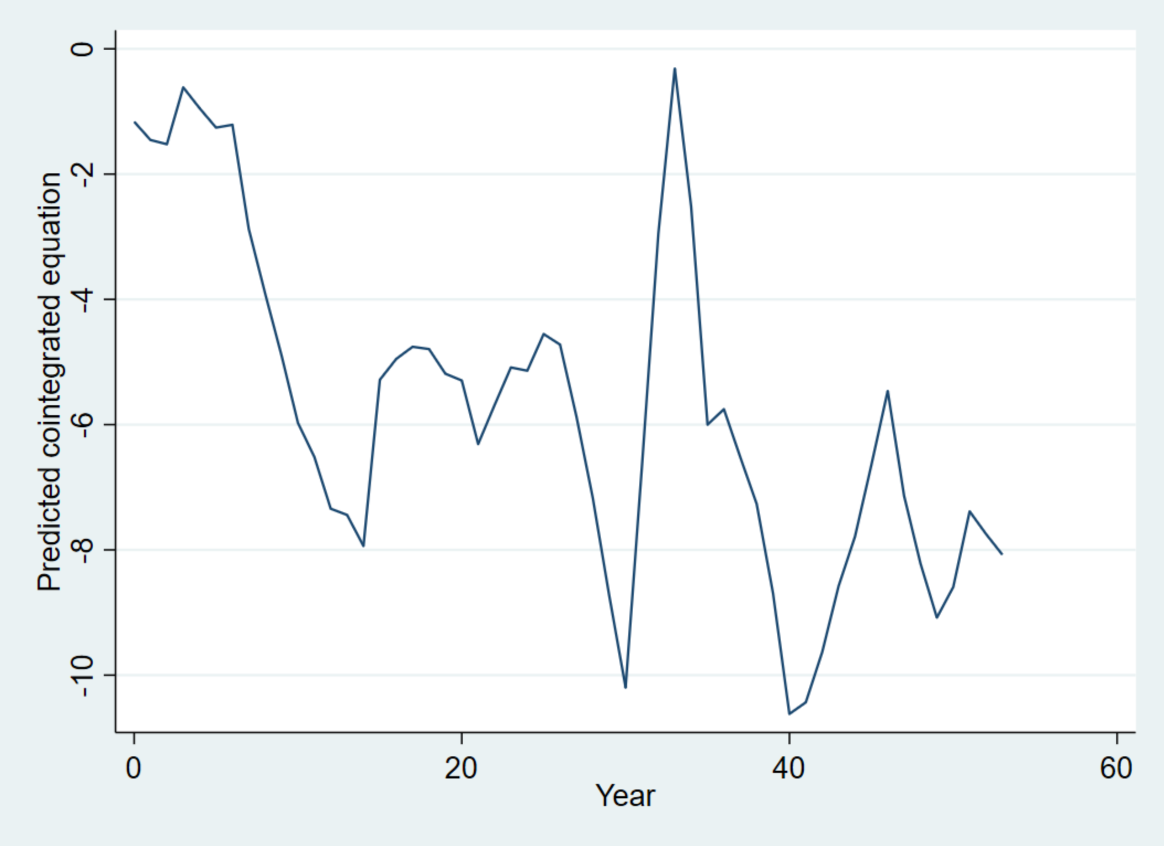
Despite the lack of stability, an orthogonal impulse response function was generated to explore the impact of a shock in each of the variables in the model on the Gini coefficient. As the plots show, there seems that all of the shocks are permanent because none of them subsides over time or plateaus.



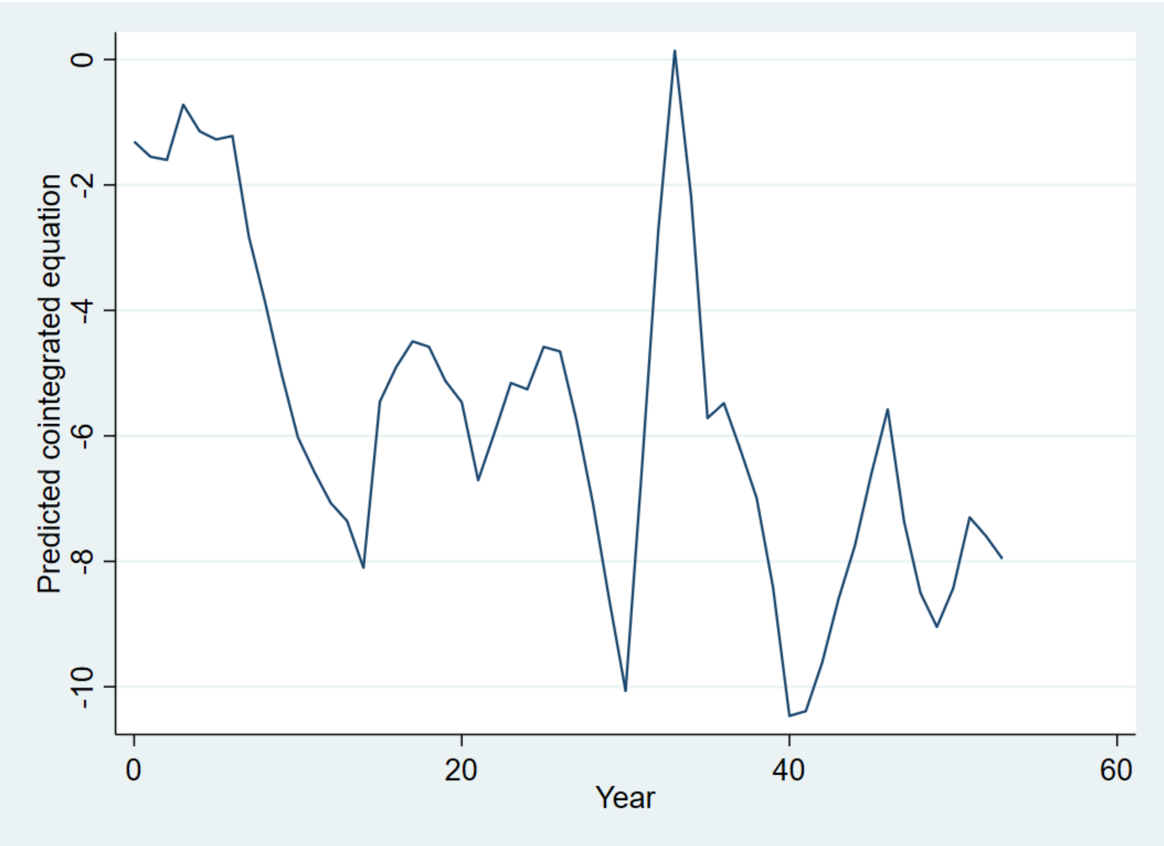
Given these results, another VECM was created with 2 lags, 3 cointegrating equations and 9 restrictions. The output [Appendix B] shows that, in the short term, government expenditure has a statistically significant impact on the public employment rate, the RGDP on the openness of the economy, which also has a significant impact on the RGDP, and the RGDP has a statistically significant impact on the investments as rate of GDP. In the long run, in the first cointegrating equation (ce1), after restricting Gini, and the two employment rates, RGDP and government expenditures have a statistically significant positive and negative impact on Gini, respectively. In the second cointegrating equation (ce2), after restricting Gini, the gross investments, and government expenditures, both employment rates have a negative impact on Gini while the RGDP has a positive impact on it. Finally, in the third cointegrating equation (ce3), after restricting Gini, the private sector employment rate, and the openness of the economy, the public sector employment rate has a negative impact on Gini, the RGDP a positive one, and the government expenditure a negative one.



Cointegrating equation 1

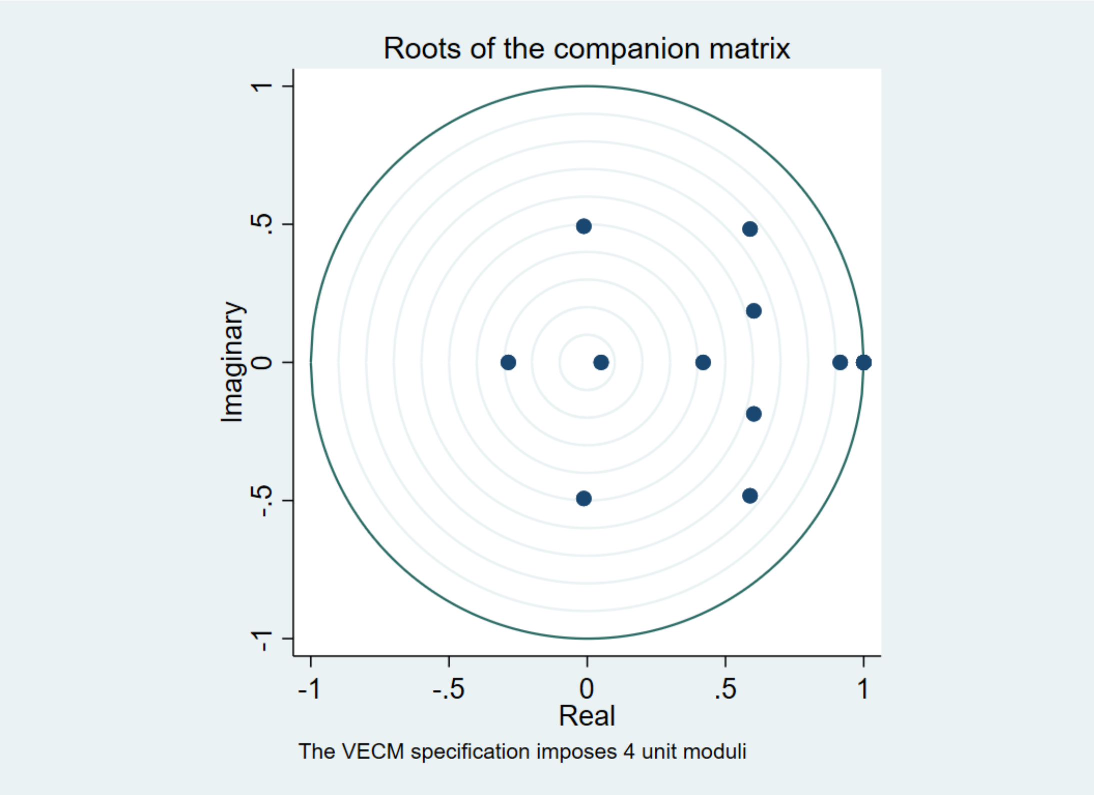


Cointegrating equation 2

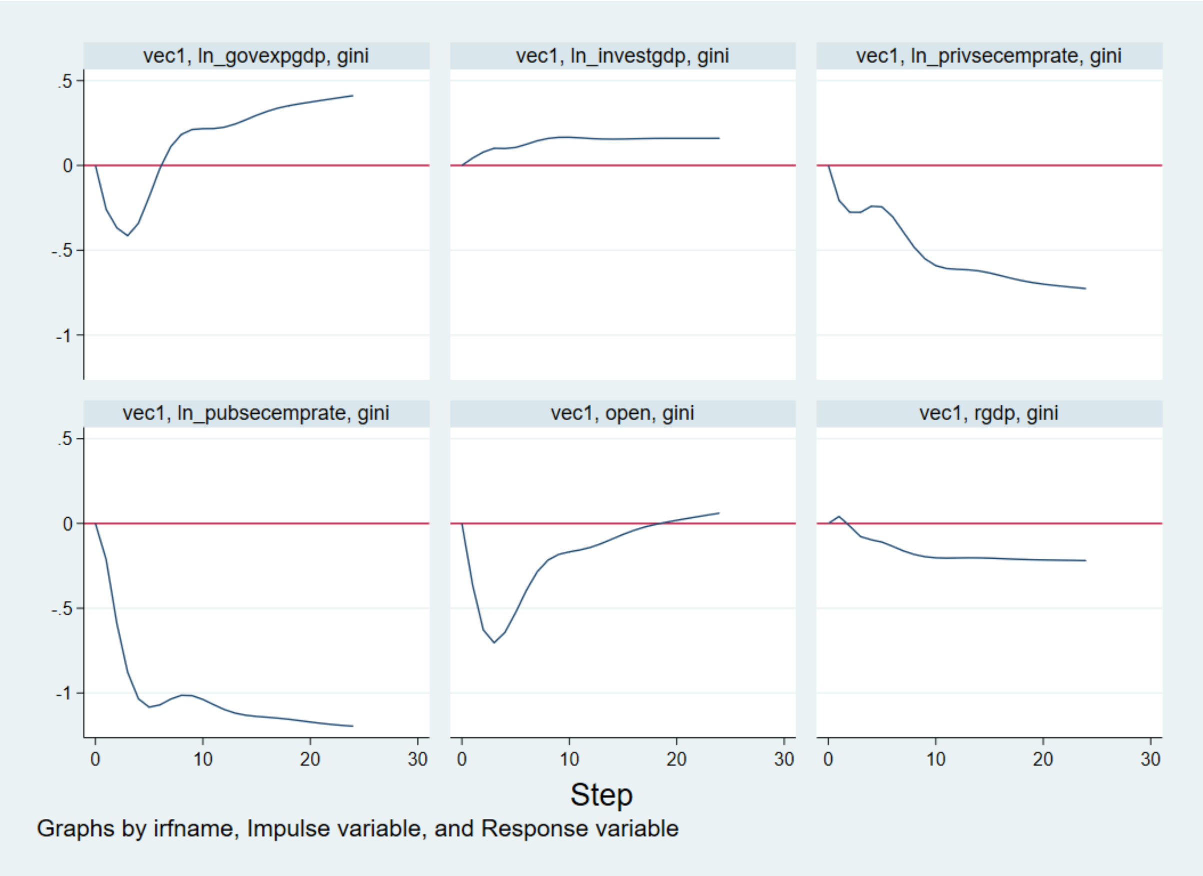


Cointegrating equation 3

The model is still unstable, after testing for instability:



Although unstable, an orthogonal impulse response function was plotted for each variable against the Gini coefficient to determine their impact.



A shock in the RGDP or the gross investments as percentage of GDP seems to stabilize and plateau after some time. On the other hand, a negative shock to either public or private employment rate appears to have a permanent negative impact on the inequality level of South Africa. Finally, a shock to the government expenditures has, according to this VECM, a permanent positive impact on Gini.

Total word count: 3050 words

APPENDIX A

Phillips-Perron test on variables RGDP, Capital Formation Government, Capital Formation Public Corporations, Capital Formation Private Corporations.

```
Phillips-Perron test for unit root          Number of obs    = 74
Variable: Rgdp                              Newey-West lags =  3

H0: Random walk without drift, a = 0, d = 0
```

	Test statistic	Dickey-Fuller critical value 1%	5%	10%
Z(rho)	1.364	-13.092	-7.796	-5.548
Z(t)	4.292	-2.610	-1.950	-1.610

Regression table

Rgdp	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
Rgdp L1.	1.018725	.0032556	312.92	0.000	1.012236	1.025213

Phillips-Perron CASE 1 no constant RGDP

```
Phillips-Perron test for unit root          Number of obs    = 74
Variable: Rgdp                              Newey-West lags =  3

H0: Random walk without drift, d = 0
```

	Test statistic	Dickey-Fuller critical value 1%	5%	10%
Z(rho)	0.217	-19.332	-13.492	-10.844
Z(t)	0.363	-3.546	-2.911	-2.590

MacKinnon approximate *p*-value for Z(t) = 0.9801.

Regression table

Rgdp	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
Rgdp L1.	1.003817	.0062968	159.42	0.000	.9912642	1.016369
_cons	29586.9	10853.84	2.73	0.008	7950.152	51223.64

Phillips-Perron CASE 2 constant RGDP

```
Phillips-Perron test for unit root          Number of obs    = 74
Variable: Rgdp                              Newey-West lags =  3

H0: Random walk with or without drift
```

	Test statistic	Dickey-Fuller critical value 1%	5%	10%
Z(rho)	-5.912	-26.516	-20.232	-17.136
Z(t)	-1.933	-4.097	-3.476	-3.166

MacKinnon approximate *p*-value for Z(t) = 0.6372.

Regression table

Rgdp	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
Rgdp L1.	.9436216	.0326529	28.90	0.000	.8785136	1.00873
_trend	2452.117	1306.043	1.88	0.065	-152.0581	5056.292
_cons	27749.08	10713.2	2.59	0.012	6387.57	49110.58

Phillips-Perron CASE 4 trend RGDP

```
Phillips-Perron test for unit root          Number of obs    = 74
Variable: capitgov                          Newey-West lags =  3

H0: Random walk without drift, a = 0, d = 0
```

	Test statistic	Dickey-Fuller critical value 1%	5%	10%
Z(rho)	0.504	-13.092	-7.796	-5.548
Z(t)	0.511	-2.610	-1.950	-1.610

Regression table

capitgov	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
capitgov L1.	1.008408	.0115539	87.28	0.000	.9853806	1.031434

Phillips-Perron CASE 1 no constant CapGov

```
Phillips-Perron test for unit root          Number of obs   = 74
Variable: capitgov                          Newey-West lags =  3

H0: Random walk without drift, d = 0

                                       Dickey-Fuller
                   Test        ──────── critical value ────────
                 statistic          1%           5%          10%
─────────────────────────────────────────────────────────────────
 Z(rho)            -3.632      -19.332      -13.492      -10.844
 Z(t)              -1.475       -3.546       -2.911       -2.590
─────────────────────────────────────────────────────────────────
MacKinnon approximate p-value for Z(t) = 0.5457.

Regression table
──────────────────────────────────────────────────────────────────────────────
    capitgov │ Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
─────────────┼────────────────────────────────────────────────────────────────
    capitgov │
         L1. │   .9612707   .0287339    33.45   0.000     .9039906    1.018551
             │
       _cons │   3189.299   1785.053     1.79   0.078    -369.1401    6747.737
──────────────────────────────────────────────────────────────────────────────
```

Phillips-Perron CASE 2 constant CapGov

```
Phillips-Perron test for unit root          Number of obs   = 74
Variable: capitgov                          Newey-West lags =  3

H0: Random walk with or without drift

                                       Dickey-Fuller
                   Test        ──────── critical value ────────
                 statistic          1%           5%          10%
─────────────────────────────────────────────────────────────────
 Z(rho)            -6.076      -26.516      -20.232      -17.136
 Z(t)              -1.770       -4.097       -3.476       -3.166
─────────────────────────────────────────────────────────────────
MacKinnon approximate p-value for Z(t) = 0.7188.

Regression table
──────────────────────────────────────────────────────────────────────────────
    capitgov │ Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
─────────────┼────────────────────────────────────────────────────────────────
    capitgov │
         L1. │   .9388805   .0397825    23.60   0.000     .8595563    1.018205
             │
      _trend │       37.4   45.84128     0.82   0.417    -54.00492    128.8049
       _cons │    3063.93   1795.799     1.71   0.092    -516.7922    6644.652
──────────────────────────────────────────────────────────────────────────────
```

Phillips-Perron CASE 4 trend CapGov

```
Phillips-Perron test for unit root          Number of obs    = 74
Variable: cappubcorp                        Newey-West lags =  3

H0: Random walk without drift, a = 0, d = 0
```

	Test statistic	Dickey-Fuller critical value 1%	5%	10%
Z(rho)	-0.509	-13.092	-7.796	-5.548
Z(t)	-0.312	-2.610	-1.950	-1.610

Regression table

cappubcorp	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
cappubcorp L1.	1.001033	.0166565	60.10	0.000	.9678366	1.034229

Phillips-Perron CASE 1 no constant CapPubCorp

```
Phillips-Perron test for unit root          Number of obs    = 74
Variable: cappubcorp                        Newey-West lags =  3

H0: Random walk without drift, d = 0
```

	Test statistic	Dickey-Fuller critical value 1%	5%	10%
Z(rho)	-2.979	-19.332	-13.492	-10.844
Z(t)	-1.265	-3.546	-2.911	-2.590

MacKinnon approximate *p*-value for Z(t) = 0.6451.

Regression table

cappubcorp	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
cappubcorp L1.	.9757968	.024051	40.57	0.000	.9278521	1.023741
_cons	1784.975	1235.418	1.44	0.153	-677.7861	4247.736

Phillips-Perron CASE 2 constant CapPubCorp

```
Phillips-Perron test for unit root          Number of obs    = 74
Variable: cappubcorp                        Newey-West lags =  3

H0: Random walk with or without drift
```

	Test statistic	Dickey-Fuller critical value 1%	5%	10%
Z(rho)	-8.863	-26.516	-20.232	-17.136
Z(t)	-2.049	-4.097	-3.476	-3.166

MacKinnon approximate *p*-value for Z(t) = **0.5744**.

Regression table

cappubcorp	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
cappubcorp L1.	.934337	.0442349	21.12	0.000	.8461351	1.022539
_trend	81.60861	73.12844	1.12	0.268	-64.20536	227.4226
_cons	271.2739	1833.28	0.15	0.883	-3384.183	3926.731

Phillips-Perron CASE 4 trend CapPubCorp

```
Phillips-Perron test for unit root          Number of obs    = 74
Variable: capPrivCorp                       Newey-West lags =  3

H0: Random walk without drift, a = 0, d = 0
```

	Test statistic	Dickey-Fuller critical value 1%	5%	10%
Z(rho)	1.068	-13.092	-7.796	-5.548
Z(t)	1.316	-2.610	-1.950	-1.610

Regression table

capPrivCorp	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
capPrivCorp L1.	1.015281	.0098469	103.11	0.000	.9956564	1.034906

Phillips-Perron CASE 1 no constant CapPrivCorp

```
Phillips-Perron test for unit root          Number of obs    = 74
Variable: capPrivCorp                       Newey-West lags =  3

H0: Random walk without drift, d = 0

                                     Dickey-Fuller
                  Test     ──────── critical value ─────────
              statistic         1%           5%          10%
─────────────────────────────────────────────────────────────
 Z(rho)          -0.487    -19.332      -13.492      -10.844
 Z(t)            -0.369     -3.546       -2.911       -2.590
─────────────────────────────────────────────────────────────
MacKinnon approximate p-value for Z(t) = 0.9151.

Regression table
─────────────────────────────────────────────────────────────────────────────
 capPrivCorp │ Coefficient  Std. err.       t    P>|t|     [95% conf. interval]
─────────────┼───────────────────────────────────────────────────────────────
 capPrivCorp │
         L1. │    .995579   .0161419    61.68   0.000     .9634008    1.027757
             │
       _cons │   5068.729   3308.183     1.53   0.130    -1526.013    11663.47
─────────────────────────────────────────────────────────────────────────────
```

Phillips-Perron CASE 2 constant CapPrivCorp

```
Phillips-Perron test for unit root          Number of obs    = 74
Variable: capPrivCorp                       Newey-West lags =  3

H0: Random walk with or without drift

                                     Dickey-Fuller
                  Test     ──────── critical value ─────────
              statistic         1%           5%          10%
─────────────────────────────────────────────────────────────
 Z(rho)          -6.968    -26.516      -20.232      -17.136
 Z(t)            -1.965     -4.097       -3.476       -3.166
─────────────────────────────────────────────────────────────
MacKinnon approximate p-value for Z(t) = 0.6206.

Regression table
─────────────────────────────────────────────────────────────────────────────
 capPrivCorp │ Coefficient  Std. err.       t    P>|t|     [95% conf. interval]
─────────────┼───────────────────────────────────────────────────────────────
 capPrivCorp │
         L1. │   .9218821   .0431453    21.37   0.000     .8358527    1.007911
             │
      _trend │   459.7035    250.233     1.84   0.070    -39.24701     958.654
       _cons │  -138.0898   4315.973    -0.03   0.975    -8743.896    8467.716
─────────────────────────────────────────────────────────────────────────────
```

Phillips-Perron CASE 4 trend CapPrivCorp

Appendix B

Sample: 2 thru 53		Number of obs =	52
		AIC =	16.14201
Log likelihood = -333.6923		HQIC =	17.37919
Det(Sigma_ml) = .0008842		SBIC =	19.36907

Equation	Parms	RMSE	R-sq	chi2	P>chi2
D_gini	11	.846698	0.6301	68.12776	0.0000
D_ln_pubsecemp~e	11	.02424	0.3815	24.67621	0.0102
D_ln_privsecem~e	11	.027587	0.4652	34.79754	0.0003
D_open	11	3.8496	0.3556	22.07286	0.0238
D_rgdp	11	31747.8	0.7760	138.5322	0.0000
D_ln_investgdp	11	.051314	0.5023	40.37275	0.0000
D_ln_govexpgdp	11	.042073	0.4166	28.56535	0.0027

	Coefficient	Std. err.	z	P>|z|	[95% conf.	interval]
D_gini						
_ce1 L1.	-.1690972	.0703732	-2.40	0.016	-.3070262	-.0311682
_ce2 L1.	-2.617855	1.552403	-1.69	0.092	-5.660509	.4247993
_ce3 L1.	2.531392	1.567589	1.61	0.106	-.5410268	5.603811
gini LD.	.5985203	.1071455	5.59	0.000	.388519	.8085217
ln_pubsecemprate LD.	-1.380564	5.306609	-0.26	0.795	-11.78133	9.020198
ln_privsecemprate LD.	-8.626488	5.78565	-1.49	0.136	-19.96615	2.713178
open LD.	-.0355	.0367134	-0.97	0.334	-.1074568	.0364569
rgdp LD.	-2.40e-07	5.21e-06	-0.05	0.963	-.0000105	9.98e-06
ln_investgdp LD.	1.144378	2.34844	0.49	0.626	-3.45848	5.747236
ln_govexpgdp LD.	-3.983366	2.89126	-1.38	0.168	-9.650131	1.6834
_cons	.0032227	.4169964	0.01	0.994	-.8140753	.8205207

D_ln_pubsecemprate						
_ce1 L1.	-.0013888	.0020147	-0.69	0.491	-.0053375	.0025599
_ce2 L1.	.0372378	.0444429	0.84	0.402	-.0498686	.1243442
_ce3 L1.	-.0395668	.0448776	-0.88	0.378	-.1275253	.0483918
gini LD.	.0001459	.0030674	0.05	0.962	-.0058661	.0061579
ln_pubsecemprate LD.	.1286989	.1519198	0.85	0.397	-.1690585	.4264564
ln_privsecemprate LD.	.1036229	.1656341	0.63	0.532	-.2210139	.4282597
open LD.	.0014439	.001051	1.37	0.170	-.0006161	.0035039
rgdp LD.	4.40e-08	1.49e-07	0.30	0.768	-2.48e-07	3.37e-07
ln_investgdp LD.	-.0158856	.0672321	-0.24	0.813	-.1476582	.115887
ln_govexpgdp LD.	.1901604	.0827722	2.30	0.022	.0279299	.352391
_cons	-.0113314	.011938	-0.95	0.343	-.0347293	.0120666
D_ln_privsecemprate						
_ce1 L1.	-.0053349	.0022929	-2.33	0.020	-.0098289	-.0008409
_ce2 L1.	-.1146752	.0505806	-2.27	0.023	-.2138113	-.0155392
_ce3 L1.	.1217324	.0510754	2.38	0.017	.0216266	.2218383
gini LD.	.0003291	.003491	0.09	0.925	-.0065132	.0071714
ln_pubsecemprate LD.	.0366835	.1729005	0.21	0.832	-.3021953	.3755622
ln_privsecemprate LD.	.2119881	.1885087	1.12	0.261	-.1574821	.5814583
open LD.	.0000891	.0011962	0.07	0.941	-.0022554	.0024336

rgdp						
LD.	3.10e-07	1.70e-07	1.82	0.068	-2.33e-08	6.42e-07
ln_investgdp						
LD.	.0028826	.0765171	0.04	0.970	-.1470883	.1528534
ln_govexpgdp						
LD.	-.0925395	.0942034	-0.98	0.326	-.2771746	.0920957
_cons	.0365126	.0135866	2.69	0.007	.0098833	.0631419
D_open						
_ce1						
L1.	-.5634904	.3199589	-1.76	0.078	-1.190598	.0636176
_ce2						
L1.	-22.79564	7.058156	-3.23	0.001	-36.62937	-8.96191
_ce3						
L1.	23.2016	7.127203	3.26	0.001	9.232538	37.17066
gini						
LD.	.3919441	.4871479	0.80	0.421	-.5628482	1.346736
ln_pubsecemprate						
LD.	28.41829	24.12703	1.18	0.239	-18.86982	75.70639

ln_privsecemprate						
LD.	-4.623646	26.30504	-0.18	0.860	-56.18057	46.93328
open						
LD.	-.1990941	.1669209	-1.19	0.233	-.5262532	.1280649
rgdp						
LD.	.0000789	.0000237	3.33	0.001	.0000324	.0001253
ln_investgdp						
LD.	-14.50225	10.67742	-1.36	0.174	-35.4296	6.425108
ln_govexpgdp						
LD.	-11.75429	13.1454	-0.89	0.371	-37.5188	14.01023
_cons	-.0002272	1.895916	-0.00	1.000	-3.716154	3.7157
D_rgdp						
_ce1						
L1.	4983.174	2638.712	1.89	0.059	-188.6073	10154.95
_ce2						
L1.	2590.094	58208.85	0.04	0.965	-111497.2	116677.3
_ce3						
L1.	-3358.535	58778.28	-0.06	0.954	-118561.8	111844.8
gini						
LD.	3564.391	4017.525	0.89	0.375	-4309.814	11438.6
ln_pubsecemprate						
LD.	167004.8	198976.4	0.84	0.401	-222981.8	556991.4
ln_privsecemprate						
LD.	194098.2	216938.5	0.89	0.371	-231093.5	619289.9
open						
LD.	-4831.333	1376.603	-3.51	0.000	-7529.425	-2133.242
rgdp						
LD.	.6113547	.1954618	3.13	0.002	.2282566	.9944529
ln_investgdp						
LD.	-27646.86	88057.03	-0.31	0.754	-200235.5	144941.8
ln_govexpgdp						
LD.	-36923.04	108410.6	-0.34	0.733	-249403.9	175557.8
_cons	.0286337	15635.68	0.00	1.000	-30645.35	30645.41
D_ln_investgdp						
_ce1						
L1.	-.0003968	.004265	-0.09	0.926	-.0087561	.0079624
_ce2						
L1.	-.1362891	.0940839	-1.45	0.147	-.3206902	.048112
_ce3						
L1.	.1387071	.0950043	1.46	0.144	-.047498	.3249121

gini						
LD.	.0045887	.0064936	0.71	0.480	-.0081385	.0173159
ln_pubsecemprate						
LD.	.1969327	.3216089	0.61	0.540	-.4334092	.8272746
ln_privsecemprate						
LD.	.5828449	.3506414	1.66	0.096	-.1043996	1.270089
open						
LD.	-.0037773	.002225	-1.70	0.090	-.0081383	.0005837
rgdp						
LD.	1.06e-06	3.16e-07	3.37	0.001	4.45e-07	1.68e-06
ln_investgdp						
LD.	.1154291	.1423281	0.81	0.417	-.1635288	.3943869
ln_govexpgdp						
LD.	-.1903557	.1752259	-1.09	0.277	-.5337921	.1530806
_cons	-.0260278	.0252722	-1.03	0.303	-.0755604	.0235048
D_ln_govexpgdp						
_ce1						
L1.	-.0019701	.0034969	-0.56	0.573	-.0088239	.0048837
_ce2						
L1.	.1131778	.0771406	1.47	0.142	-.038015	.2643706
_ce3						
L1.	-.1091214	.0778953	-1.40	0.161	-.2617933	.0435505
gini						
LD.	.0014366	.0053242	0.27	0.787	-.0089986	.0118718
ln_pubsecemprate						
LD.	.2671417	.2636912	1.01	0.311	-.2496836	.783967
ln_privsecemprate						
LD.	-.5510074	.2874953	-1.92	0.055	-1.114488	.0124731
open						
LD.	.0008848	.0018243	0.48	0.628	-.0026908	.0044604
rgdp						
LD.	-3.07e-07	2.59e-07	-1.19	0.235	-8.15e-07	2.00e-07
ln_investgdp						
LD.	.1552141	.1166966	1.33	0.183	-.073507	.3839352
ln_govexpgdp						
LD.	.0492295	.1436699	0.34	0.732	-.2323583	.3308174
_cons	.0556968	.020721	2.69	0.007	.0150844	.0963092

```
Cointegrating equations

Equation           Parms    chi2     P>chi2
---------------------------------------------
_ce1                   4   105.2474   0.0000
_ce2                   4   124.1155   0.0000
_ce3                   4   135.9756   0.0000
---------------------------------------------

Identification:  beta is overidentified

 ( 1)  [_ce1]gini = 1
 ( 2)  [_ce1]ln_pubsecemprate = 0
 ( 3)  [_ce1]ln_privsecemprate = 0
 ( 4)  [_ce2]gini = 1
 ( 5)  [_ce2]ln_investgdp = 0
 ( 6)  [_ce2]ln_govexpgdp = 0
 ( 7)  [_ce3]gini = 1
 ( 8)  [_ce3]open = 0
 ( 9)  [_ce3]ln_privsecemprate = 0
------------------------------------------------------------------------------------
               beta | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
--------------------+---------------------------------------------------------------
_ce1                |
               gini |          1          .        .       .            .           .
   ln_pubsecemprate |          0  (omitted)
  ln_privsecemprate |          0  (omitted)
               open |   .1600966   .0913668     1.75   0.080    -.0189791    .3391722
               rgdp |   -.000014   1.67e-06    -8.38   0.000    -.0000173   -.0000107
       ln_investgdp |  -2.211385   3.536934    -0.63   0.532    -9.143647    4.720878
       ln_govexpgdp |   49.18942    5.24408     9.38   0.000     38.91122    59.46763
              _cons |  -196.9702          .        .       .            .           .
--------------------+---------------------------------------------------------------
_ce2                |
               gini |          1          .        .       .            .           .
   ln_pubsecemprate |   29.57277    4.00853     7.38   0.000     21.71619    37.42934
  ln_privsecemprate |   .8685877   .1958259     4.44   0.000      .484776    1.252399
               open |   .0081987   .0047598     1.72   0.085    -.0011303    .0175276
               rgdp |  -7.22e-06   1.04e-06    -6.93   0.000    -9.26e-06   -5.18e-06
       ln_investgdp |          0  (omitted)
       ln_govexpgdp |          0  (omitted)
              _cons |    45.5035          .        .       .            .           .
--------------------+---------------------------------------------------------------
_ce3                |
               gini |          1          .        .       .            .           .
   ln_pubsecemprate |   29.11981   3.921756     7.43   0.000     21.43331    36.80631
  ln_privsecemprate |          0  (omitted)
               open |          0  (omitted)
               rgdp |  -7.72e-06   1.04e-06    -7.43   0.000    -9.76e-06   -5.68e-06
       ln_investgdp |   .1342435   .1876393     0.72   0.474    -.2335227    .5020097
       ln_govexpgdp |   2.160416   .3112061     6.94   0.000     1.550463    2.770368
              _cons |   37.10238          .        .       .            .           .
------------------------------------------------------------------------------------
LR test of identifying restrictions: chi2(3) = 6.018        Prob > chi2 = 0.111
```