Multivariate cointegration analysis of South African employment, inflation and output data for short- and long-run linkages

Sagaren Pillay*
Statistics South Africa, Pretoria, South Africa – Sagarenp@statssa.gov.za

Nazeem Mustapha
Human Sciences Research Council, Cape Town, South Africa – NMustapha@hsrc.ac.za

Abstract

This study empirically examines the short- run and long- run linkages between South Africa's output, inflation and employment using quarterly data from 1994 to 2013. The theoretical framework used for the study is based on Johansen's Maximum Likelihood cointegration technique which tests for both the existence and number of cointegration vectors that exists. The study finds that the three series are integrated of order one and are cointegrated. A statistically significant cointegrating relationship is found to exist between employment, inflation and output. The findings of the study confirm the existence of long run and short run linkages between the variables. This unique linear and lagged relationship is modelled using a Vector Error Correction Model (VECM). The VECM (2) model is found to be a good fit for the data with output and employment a weak exogeneity of inflation.

Key words: exogeneity lagged, linear, vector error correction model.

1. Introduction

The cointegration relationship between inflation and output as well as the relationship between employment and output has been researched extensively over the past decade. Research exploring the short-run and long run linkages and dynamic interaction between the three variables, employment, output and inflation considered jointly is very limited.

An early discussion of the relationship between fluctuations in employment and economic cycles, employment and inflation, and employment and output discussed these interactions and argued the need for knowledge on the quantitative relationships between these and other macro-economic variables Phillips (1958) and Phillips (1962). This discussion hinged on the premise that these three variables interact in an intrinsically non-linear fashion with bicausality or cyclic causality in what has come to be known as Phillips' Golden Triangle. Without quantitative knowledge of these relationships, it is not possible to choose optimally the particular inflation rate, level of economic activity or "natural rate of unemployment" Phillips (1962) that should be targeted by policy makers. Further, a proper understanding of the employment/inflation/output relationship might also be instrumental to avoiding or at least alleviating business cycles.

On the one hand, studies Okun (1981) that focused on the relationship between output and employment have found a positive relationship between output and employment. An approach that uses a model of the Phillips curve, Ball & Moffit (2001) argued that growth has only temporary, short-run effects on employment. On the other hand, higher rates of inflation were associated with slow real growth Motley (1998). A comprehensive quantitative comparative analysis Bruno & Easterly (1998) examined the relationship between inflation and economic growth for 127 countries between 1960 and 1992. The key conclusion was the finding of a negative relationship between inflation and growth when the inflation rate is very high (annual rate of 40 % and higher), but a trade-off relationship when inflation rates were moderate were not established.

There exists a body of literature which argues that inflation drags down growth over the medium to long term Fischer (1993) while at the same time allowing for the possibility of high inflation and

An analysis of the link between unemployment and inflation in South Africa for the period September 2000 to June 2011 Pillay (2013) brought out the existence of a unique linear and lagged relationship between unemployment and inflation, which indicated the presence of a cointegrating vector between the two variables. Empirical results also strongly support the existence of a single cointegrating relationship between employment growth, inflation and output growth. The findings Caporale & Škare (2011) from a panel analysis of 119 countries between 1970 and 2010 were that in the short run inflation was associated with positive employment and output growth; however in the long run higher inflation was linked with lower output and employment.

This study examines empirically the short-run and long-run linkages between South Africa's output, inflation and employment economic actors using quarterly data from 1994 to 2013. The theoretical framework used for the study was based on Johansen's Maximum Likelihood cointegration technique which tests for both the existence and number of cointegration vectors (cf. Johansen (1988)).

2. Data

This study obtained quarterly output (GDP), formal employment (emp) and inflation (CPI) data from 1994 to 2013 from Statistics South Africa. All three variables were expressed in natural logarithms to reduce the influence of possible heteroskedasticity and because multiplicative effects between output, inflation and employment is assumed. The official employment data consisted of many breaks, thus a linked series was used to model the data as in Pillay & Mosalo.

3. Methodology and Results

For a VAR set of variables, the usual analysis procedure is to difference the equation containing these non-stationary variables to produce difference stationary variables, and proceed with standard time series methods to solve the difference equations and integrate back to the original variables. This results in loss of information if the difference of residuals is not explicitly taken into account. However, under some circumstances, the original equation can be considered as a VAR process in the original variables plus a term that includes a matrix Π. If Π has zero rank, this identifies the process as a standard VAR process; whereas if it has full rank, the matrix is invertible and thus the variables are stationary. If the rank of Π is in between these two extremes, this implies that there is at least one vector that makes the original vector equation integrable and thus also encodes the long-run or equilibrium behaviour of the system Johansen & Juselius (1990).

The long-run co-evolution of output, inflation and employment is examined here using the two-step test for cointegration proposed in Engle and Granger (1987). The first step tests whether each of the variables of interest has a stochastic trend. That is determined by performing the unit-root tests on the variables. The second step tests whether stochastic trends in these variables are related. That is investigated by estimating a fitted vector model under the hypothesis of cointegration identified with reduced rank of the long-run impact matrix Johansen (1988), followed by testing the suitability of the fitted vector model.

3.1 Unit Root Test

According to Hendry & Juselius (2001) if the time series are I(1), they have to be characterized by the presence of a unit root and their first difference by the absence of unit roots. The Augmented Dickey Fuller (ADF) unit root test was used to determine whether the series were stationary or non-stationary. The ADF test constructs a model with higher order lag terms and tests the significance of the parameter estimates using a non-standard *t*-test. The model used for this test was

$$\Delta x_t = \alpha_1 x_{t-1} + \beta_1 \Delta x_{t-1} + \beta_2 \Delta x_{t-2} + \ldots + \beta_{p-1} \Delta x_{t-p+1} + \varepsilon_t$$

where the *t*-test checks significance of the α_1 term. If $\alpha_1 = 0$ the series has a unit root.

The Dickey & Fuller (1979) test results for nonstationarity of each of the three series is shown in Table 1. The null hypothesis is to test a unit root. In the Dickey-Fuller test results in Table 1, the second column specifies three types of models, which are zero mean, single mean, or trend. The third column (Rho) and the fifth column (Tau) are the test statistics for unit root testing. Other columns are the *p*-values.

Table 1: Results of the Dickey-Fuller Unit Root Test

Variable	Type	Rho	Pr < Rho	Tau	Pr < Tau
y1	Zero Mean	0.16	0.7163	7.85	0.9999
	Single Mean	-0.34	0.936	-1.07	0.7254
	Trend	-21.22	0.0385	-3.01	0.1359
y2	Zero Mean	0.28	0.7475	5.59	0.9999
	Single Mean	-0.45	0.9285	-1.24	0.6538
	Trend	-13.27	0.2235	-2.98	0.1458
y3	Zero Mean	0.03	0.6864	3.27	0.9997
	Single Mean	-1.08	0.8756	-1.49	0.5313
	Trend	-1.73	0.9747	-0.63	0.9745

Consequently, all three series have a unit root and their first differences do not have any. Thus, the variables y_1= log (GDP), y_2 = log (CPI) and y_3 = log (emp) are first order difference stationary and are integrated, I (1).

3.2 Cointegration Test

The SAS procedure PROC VARMAX was used to model the data. The Minimum Information Criterion (not displayed) was used to inform the selection of an autoregressive order of *p*=2. The results of the cointegration tests are shown in Table 2.

Table 2: Results of the Cointegration Rank Test for I(2) Using Trace

r\k-r-s	**3**	**2**	**1**	**Trace**	**5% CV of I(1)**
0	175.7439	84.96163	36.68715	32.1758	29.38
1		81.46671	34.73748	3.2083	15.34
2			22.53594	1.1213	3.84
5% CV I(2)	29.38	15.34	3.84		

The Johansen rank test (Table 2) shows whether the series is integrated of order either 1 or 2. The last two columns in Table 2 explain the cointegration rank test with integrated order 1. The results indicate that there is the cointegrated relationship with the cointegration rank equal to 1 with respect to 0.05 significance level, because the test statistic of 3.2083 is smaller than the critical value of 15.34. The row associated with r=1 is used to determine the test statistic and critical value pairs such as (81.4667, 15.34) and (34.7374, 3.84). There is no evidence to indicate that the series are integrated of order 2 at the 0.05 significance level.

The fitted VECM (2) can be written in the following general form $\Delta y_t = \pi + \Pi\, y_{t-1} + \Phi_1 \Delta y_{t-1} + \varepsilon_t$ where $\Pi = \alpha\beta^T$ and $y_t \sim CI(1)$ is a k x 1 random vector; y_t is a VAR(p) process, α and β are k x 1 matrices, Φ_1 is the autoregressive (AR) coefficients of the differenced lag, Π is the vector of constant estimates, ε_t is a k x 1 white noise process.

The Johansen and Julius test -- with cointegration statistic $\lambda_{\text{trace}} = -T\sum_{\lambda=r+1}^{k}\log(1-\lambda_i)$ where T is the available number of observations and λ_i the eigenvalues -- determines the number of cointegrated vectors. The null hypothesis is that there are at most r cointegrated vectors versus the alternative hypothesis of there being more than *r* cointegrated vectors. The critical values at 0.05 significance level were used for testing. Firstly, the impact matrix Π is computed from the long-run parameter β and adjustment coefficient α.

The estimated cointegrating vector is β^T = [1, -1.47773, -0.79855], and α^T = [-0.08018, 0.13288, 0.02688]. The first element of β^T is 1 since y_1 = log (GDP) is specified as the normalised variable. The impact matrix is $\Pi = \alpha\beta^T$, which may be written as

$$\Pi = \begin{bmatrix} -0.08018 \\ 0.13288 \\ 0.02688 \end{bmatrix} [1, -1.47773, -0.79855] = \begin{bmatrix} -0.08018 & 0.11848 & 0.06402 \\ 0.13288 & -0.19636 & -0.10611 \\ 0.\ 02688 & -0.03972 & -0.02147 \end{bmatrix}$$

The long-run relationship of the series is $\beta^T y = [1, -1.47773, -0.79855]\ [y_1, y_2, y_3]^T$ or $y_{1t} = 1.47773\, y_{2t} + 0.79855\, y_{3t}$. By inserting the parameter estimates, the fitted VECM (2) model, written in the general form $\Delta y_t = \pi + \Pi\, y_{t-1} + \Phi_1 \Delta y_{t-1} + \varepsilon_t$ where $\Pi = \alpha\beta^T$, now becomes

$$\Delta y_t = \begin{bmatrix} -0.45052 \\ 0.80243 \\ 0.16256 \end{bmatrix} + \begin{bmatrix} -0.08018 & 0.11848 & 0.06402 \\ 0.13288 & -0.19636 & -0.10611 \\ 0.02688 & -0.03972 & -0.02147 \end{bmatrix} y_{t-1} + \begin{bmatrix} -0.27381 & 0.46253 & 0.11559 \\ -0.06933 & 0.15638 & 0.06068 \\ 0.02105 & 0.03035 & 0.27131 \end{bmatrix} \Delta y_{t-1} + \varepsilon_t$$

The long-run relationship may be interpreted as an error correction mechanism giving the form and size of output as a function of inflation and employment. Increases in output would result in increases in both inflation and formal employment. These changes occur elastically in the case of inflation with an elasticity of 147.7% and inelastically for formal employment with an elasticity of 79.9%. The employment elasticity of output is fixed for any given inflation level in the long-run. The system is constrained in such a way that in the long-run, inflation changes have asymmetric effects on output and formal employment. The coefficients of α may be interpreted as the average speed of adjustment towards the estimated equilibrium state, such that a low coefficient indicates slow adjustment and a high coefficient indicates rapid adjustment. The first equation, which measures the change in output, has an average speed of adjustment to the equilibrium path of approximately 0.08018.

4. Model diagnostics

Hendry and Juselius (2001) states that checking the assumptions of the model is not only crucial for meaningful statistical inference, but also for the economic interpretation of the model as a description of the behaviour of rational agents. The Portmanteau test Hosking (1980) results into significant (not shown) indicating that a VECM (2) model is a good fit.

Table 3: Univariate Model ANOVA and White Noise Diagnostics

			y1	y2	y3
ANOVA	**R-Square**		0.1416	0.2149	0.1566
	Standard Deviation		0.02689	0.0101	0.00863
	F Value		1.92	3.19	2.17
	Pr > F		0.0888	0.0079	0.0566
White Noise	**Durbin-Watson**		2.13656	2.06387	1.83296
	Normality	**Chi-Square**	49.12	17.76	8.67
		Pr > ChiSq	<.0001	0.0001	0.0131
	ARCH	**F Value**	12.12	0.05	4.75
		Pr > F	0.0008	0.8167	0.0325

The univariate model F statistics and R-square statistics indicate a good model fit. Goodness-of-fit for y_2 is indicated at the 0.05 significance level, and goodness-of-fit for y_1 and y_3 is indicated at the 0.1 significance level (Table 3). The residuals are investigated for normality, autoregressive conditional heteroskedasticity (ARCH) and autoregressive (AR) effects. The Durbin & Watson (1950, 1951) test statistic is used to test autocorrelation of the residuals. This statistic is always between nought and four. A value of two indicates that there is no autocorrelation, whereas values close to nought indicate positive autocorrelation and values toward the other extreme indicate negative autocorrelation. Table 3 shows that the residuals deviate from normality. There are ARCH effects on the residuals for y_1 and y_3, but not for y_2.

Table 4: Test Results for Autoregressive Effects on the Residuals

Variable	AR1		AR2		AR3		AR4	
	F Value	Pr > F	F Value	Pr > F	F Value	Pr > F	F Value	Pr > F
y1	0.37	0.5465	2.61	0.0806	1.76	0.1623	2.25	0.0728
y2	0.1	0.7482	1.18	0.313	1.25	0.2975	1.66	0.1694
y3	0.44	0.507	0.29	0.7455	0.35	0.7902	0.75	0.5594

There are no autoregressive effects on the residuals - for all residual series the autoregressive model fit to the residuals up to four lags show no significance indicating that the residuals are uncorrelated (Table 4).

Table 5: Test for Weak Exogeneity of Each Variable

Variable	DF	Chi-Square	Pr > ChiSq
y1	1	0.7	0.4021
y2	1	12.8	0.0003
y3	1	0.76	0.3829

The weak exogeneity test results in Table 5 indicate that output (y_1) is the weak exogeneity of inflation (y_2) and formal employment (y_3), inflation is neither the weak exogeneity of output nor that of formal employment, and formal employment is the weak exogeneity of output and inflation.

5. Conclusion

The main objective of the study was to determine whether cointegration exists for South Africa's output, inflation and employment viewed as a coupled system. To this end quarterly data from 1994 to 2013 was modelled and estimated using the theoretical framework based on Johansen's Maximum Likelihood cointegration technique, which tests for both the existence and number of cointegration vectors. The study finds that formal employment, inflation and output are integrated of order one and are cointegrated. This unique linear and lagged relationship is modelled using a Vector Error Correction Model (VECM). The VECM (2) model is found to be a good fit for the data with output and employment a weak exogeneity of inflation. This finding sustained conditional inference from the model, as in Johansen & Juselius (1990). The cointegrated relationship between output (y_1), inflation (y_2) and formal employment (y_3), that is $y_{1t}= 1.47773y_{2t} + 0.79855y_{3t}$, must hold in the long-run. The model thus makes it clear that there is a long-term equilibrium solution and also gives a closed form expression for the equilibrium plane. Deviations from the equilibrium are stochastic, and thus in the short-run may be of any form and size for any of the three variables as long as they are attracted to the equilibrium plane in the long-run.

Future work could consider the effect of including informal employment in the analysis of the trivariate system, as well as analysing informal employment separately from the combination. This could bring out effects such as whether and to what extent the proliferation of small businesses may positively affect the employment elasticity of output.

References

Ball, L. M., & Moffit, R. A. (2001). Productivity Growth and the Phillips Curve. Cambridge: National Bureau of Economic Research.

Bruno, M., & Easterly, W. (1998). Inflation Crises and Long-Run Growth. Journal of Monetary Economics, 14(1), 3-26.

Caporale, G. M., & Škare, M. (2011). Employment Growth, Inflation and Output Growth: Was Phillips Right? Evidence from a Dynamic Panel. s.l.: Center for Economic Studies (CES), the Ifo Institute and the CESifo GmbH (Munich Society for the Promotion of Economic Research).

Dickey, D. A., & Fuller, W. A. (1979). Distribution of the Estimators for Autoregressive Time Series with a Unit Root. Journal of the American Statistical Association, 74(366), 427-431.

Durbin, J. & Watson, G. S. (1950). Testing for serial correlation in least squares regression. I. Biometrika, 37(3/4), 409-428.

Durbin, J., & Watson, G. S. (1951). Testing for serial correlation in least squares regression. II. Biometrika, 38(1/2), 159-177.

Engle, R. F., & Granger, C. W. J. (1987). Cointegration and Error Correction: Representation, Estimation and Testing. Econometrica, 55(2), 251–276.

Fischer, S. (1993). The Role of Macroeconomic Factors in Growth. Cambridge: National Bureau of Economic Research.

Hendry, D. F., & Juselius, K. (2001). Explaining cointegration analysis: Part II. The Energy Journal, 22(1), 75-120.

Hodge, D. (2006). Inflation and growth in South Africa. Cambridge Journal of Economics, 30, 163–180.

Hosking, J. R. (1980). The multivariate portmanteau statistic. Journal of the American Statistical Association, 75, 602-608.

Johansen, S. (1988). Statistical analysis of cointegration vectors. Journal of Economic Dynamics and Control, 12, 231-254.

Johansen, S., & Juselius, K. (1990). Maximum likelihood estimation and inference on cointegration—with applications to the demand for money. Oxford Bulletin of Economic Statistics, 52(2), 169-210.

Motley, B. (1998). Growth and Inflation: A Cross-Country Study. San Francisco: Federal Reserve Bank of San Francisco.

Okun, A. M. (1981). Prices and Quantities: A Macroeconomic Analysis. Washington (DC): The Brookings Institution.

Phillips, A. W. (1958). The Relation between Unemployment and the Rate of Change of Money Wage Rates in the United Kingdom. Economica, 25(17), 1861-1957.

Phillips, A. W. (1962). Employment, Inflation and Growth. Economica, 29(113), 1-16.

Pillay, S. (2013). The Link between Unemployment and Inflation Using Johansen's Co-Integration Approach and Vector Error Correction Modelling. International Journal of Social, Management, Economics and Business Engineering, 7(11), 1687-1691.

Pillay, S., & Mosalo, J., n.d. Linking and Adjusting the Levels of Survey of Total Employment and Earnings Series to the Quarterly Employment Statistics Series to Create Comparability. s.l.:s.n.