Military Expenditure, Economic Growth and Structural Stability:
A Case Study of South Africa

Goodness C. Aye*, Mehmet Balcilar**, John P. Dunne***, Rangan Gupta* and Renée van Eyden**

Department of Economics, University of Pretoria, Pretoria, 0002, South Africa.
** Department of Economics, Eastern Mediterranean University, Famagusta, Turkish Republic of Northern Cyprus, via Mersin 10, Turkey.
*** Corresponding author. School of Economics, University of Cape Town, Cape Town, South Africa. Email: jpdbris@gmail.com.
* Department of Economics, University of Pretoria, Pretoria, 0002, South Africa.
* Department of Economics, University of Pretoria, Pretoria, 0002, South Africa.

Abstract

This paper makes two contributions to the growing literature on the military expenditure-economic growth nexus. It provides a case study of a developing country, South Africa, and considers the possibilities of structural breaks in the relationship, applying newly developed econometric methods. Taking annual data from 1951 to 2010 and full sample bootstrap Granger non-causality tests, initially we find no causal link between military expenditure and GDP. Then, using parameter instability tests, the estimated VARs are found to be unstable. However, when a bootstrap rolling window estimation procedure is used to deal with time variation in the parameters, bidirectional Granger causality between the two series becomes evident in various subsamples. While military expenditure has positive predictive power for GDP at certain initial periods, it has negative predictive power at some later periods in the sample. Similar results were obtained for the causality running from GDP to military expenditure. These findings illustrate that conclusions based on the standard Granger causality tests, which neither account for structural breaks nor time variation in the relationship may be invalid.

Keywords: Military spending, Economic growth, Bootstrap, Time varying causality.

JEL Codes: C32, H56, O40

Introduction

There is a large empirical literature that considers the economic effects of military spending on growth and no consensus on the sign of the effects. In a recent survey, Dunne and Tian (2013) found that there does seem to be a more common finding that military spending has a negative effect on economic growth when post-Cold War data dominates the coverage of the study, but there still remains a range of estimates from different types of studies and the debate continues. The results can vary depending upon the theoretical underpinnings, model specifications, estimation methods, sample of countries covered and time period examined. Cross-country panel data studies have come to dominate the literature and provide valuable evidence in the attempt to determine the impact of military spending on growth at a general level, but it is still important to use case studies to try to understand the dynamic nature of the relationship and to
investigate specific issues at country level. In addition, the literature has identified possible issues of structural stability of the military spending-growth relationship, implying that it may not be sensible to make generalisations over long period of time, given the changes that took place, such as the strategic factors resulting from the end of the Cold War.

This paper makes two contributions to the literature; it provides a case study of the dynamics of the military spending growth relationship within a developing country, South Africa, and considers the possibilities of structural breaks in the relationship, using newly developed econometric methods. This is a particularly valuable case study, both because of the importance of South Africa within sub-Saharan Africa, but also because the history of apartheid military build-up and the post-apartheid military decline present an unusual degree of change in the military burden over time, thus providing more information on how the economy adjusts to changes in military spending, but also more potential for structural breaks in the relationship.

This study contributes to the debate by examining the military spending-growth nexus in South Africa. It uses Granger non-causality tests for this purpose for two reasons. First, because they allow the complexity of the underlying theoretical arguments to be avoided, by simply considering bivariate relations between military spending and economic growth (Dunne and Smith, 2010). Secondly, because they allow the use of recently developed tools to investigate structural stability. Granger non-causality testing is usually done in the context of a vector autoregression (VAR), and have been criticised because the test results are sensitive to the variables and deterministic terms included in the VAR, the lag order, sample or observation window used, treatment of integration and cointegration of the variables and level of significance. In addition, since the parameters are not structural, the test results may not be stable over different time periods (Dunne and Smith, 2010). These issues can be dealt with to some extent by the approach taken in this paper, using bootstrap tests and a rolling window estimation approach, with subsample rolling bootstrap tests to account for subsample variability (time variation) of the Granger causality tests.

The remainder of the paper is structured as follows: section two considers the theoretical channels and empirical studies of military spending and economic growth. Section three describes the data and explains the empirical model used. Results are discussed in section four while section five concludes.

**Military Expenditure and Economic Growth**

The standard economic account of the determination of military expenditures by a nation emphasises perceptions of: the threats to its security; its ability to pay, usually measured by GDP; and the opportunity costs of military expenditures. These perceptions are mediated by domestic political and bureaucratic institutions, including, perhaps, a military industrial complex (Dunne and Smith, 2010). In general, there appears to be no theoretical consensus on the nature and extent of economic effects of military or military expenditures, though a number of channels have been identified through which military spending can impact on the economy. In the short run, it can be through potential substitution effects with other government components, and in the long run through labour, capital, technology, external relations, socio-political effects, debt, conflicts, etc. (Dunne and Tian, 2013; D’Agostino et al., 2013). Keynesian demand side explanations might suggest a positive effect of military expenditure on output, while supply side displacement of factors of production might suggest a negative effect. Ability to pay arguments might suggest a positive effect of output on military expenditure, while military Keynesian effects
to stabilise output might suggest a negative effect, if output falls and military expenditure is increased to compensate (D’Agostino et al., 2010).

In the development literature, Benoit (1973, 1978) found a positive correlation between military spending and economic growth and started the empirical debate. Much of this used the developing Neoclassical (focusing on the supply-side, i.e., modernization positive externalities from infrastructure, technology, etc.) or Keynesian theoretical frameworks (focusing on the demand-side, i.e., crowding-out of investment, exports, education, health, etc.). These were operationalised as single growth equations, within a simultaneous equation framework (with a Keynesian aggregate demand and supply-side function) and as growth equations derived from aggregate production functions and as endogenous growth equations1.

Granger causality methods allow for the complexity of the underlying theoretical arguments to be ignored, by simply considering bivariate relations between military spending and growth, or with some ad hoc theoretical specification. Granger non-causality testing is usually done in the context of a VAR, with recent examples including Karagianni and Pempetzoglu (2009), Ozsoy (2008), and Kollias et al. (2007). Dunne and Smith (2010) provide a critical appraisal of this method because the test results are sensitive to the variables and deterministic terms included in the VAR, lag length, sample or observation window used, treatment of integration and cointegration of variables and the level of significance. In addition, since the parameters are not structural, the test results may not be stable over different time periods (Dunne and Smith, 2010). These issues can be dealt with to some extent by the approach taken in this paper, using bootstrap parameter instability tests and a rolling window estimation approach, with subsample rolling Granger non-causality bootstrap tests used to account for subsample variability (time variation) in the relationship between military spending and growth.

In general, empirical findings are mixed and inconclusive with results depending on the theoretical underpinnings, models and specifications, estimation methods, country or panel of countries and time periods examined as indicated by Dunne and Uye (2009) and Dunne and Nikolaidou (2012), but with no convincing evidence of positive effects of military spending on growth. More recently Dunne and Tian (2013) have identified a more robust result of military spending having a negative effect on growth when post-Cold War data dominates. These reviews highlight the possibilities of structural breaks and the need for case studies to understand the dynamics of the process.

There have been some studies of South Africa, although limited in number. A study focussing on apartheid South Africa was undertaken by Roux (1996, 2000), using a Keynesian aggregate production function model. Other studies include McMillan (1992), Dunne et al. (2000) and

---

Batchelor et al. (2000b), McMillan (1992) examines the link using a Feder Ram model based on a neoclassical production function for the period 1950-1985. The results indicate a negative size effect but a positive externality effect of military spending on economic growth. Furthermore, the findings show that decreasing foreign investment and increasing domestic unrest have negative effects on GDP. Using the same theoretical model, Batchelor et al. (2000a) estimate a neoclassical (supply-side) model for the manufacturing sector, as well as the aggregate macroeconomic level. They attempt to improve upon the model by allowing the data to determine the dynamic structure of the model through an ARDL procedure. Overall, military spending is found to have no significant impact in aggregate, but a significant negative impact for the manufacturing sector. Dunne et al. (2000) use a Keynesian (demand and supply-side) simultaneous equation model estimated for the period 1961 to 1997 and find a negative effect of military spending on economic growth. Birdi and Dunne (2002), criticise the Feder Ram model and use an aggregate production function model to underpin a VAR analysis of the impact of the growth of military spending on GDP growth, which is found to be negative and insignificant. When testing the effect of military spending on manufacturing output, a positive long-run effect, but negative short-run effect is observed.

Data and Empirical Model

As in most studies military expenditure is used as share of gross domestic product, defined as military burden (MB), and output is real gross domestic product (GDP). Data on military expenditure is sourced from the Stockholm International Peace Research Institute (SIPRI) for 1951-2010, while GDP data is from the South African Reserve Bank (SARB) Quarterly Bulletin. Note that the starting point of the sample is driven by the common date of data availability for the two series, while the end-point (2010) is also based on availability of data at the time of writing this paper. All variables are used in log levels and are plotted in Figure 1.

Figure 1: log of military burden (left) and log of real GDP (right)

Every country tries to ensure internal and external security for its inhabitants and consequently makes decisions on the magnitude of military expenditure. It is a major part of government expenditure and can significantly exceed expenditures on health, education and other socio-economic activities in developing countries (Hou and Chen, 2012).

---

2 However, Batchelor et al. (2000b) and Birdi et al. (2000) considered the impact of military spending on corporate performance and industrial growth, respectively.
3 This is an extended data series from the one published and was made available by the SIPRI military spending project. We are grateful to Sam Perlo-Freeman and Mehmet Uye for providing the series.
South Africa’s military spending has fluctuated considerably over the past century. South African society was heavily militarised during the apartheid era. This was due to the use of the military to defend white minority rule against internal and external threats. The militarisation process was reflected in increasing levels of military spending, the introduction of compulsory conscription for all white males, and the development of an local arms production capability (Batchelor et al., 2000). Aside apartheid, the different war periods (First World War, Second World War, Cold War, Korean and Namibian wars) have served as the most strategic factor in shifting military expenditure. According to Global Security (2013), South Africa’s military burden (military spending as a share of GDP) reached a high of 6 per cent and 17 per cent respectively, during the First World War (1914-1918) and Second World War (1939-1945). Thereafter, it declined sharply to around 1 per cent. During the Namibian war (1974-1976), it rose to 5 per cent and from the 1990s, military spending averaged between 2 and 3 per cent of GDP.

In nominal terms, the military budget grew almost tenfold between 1975 and 1989, from R1 billion to R9.4 billion. In real terms, however, the increase was modest – from US$3 billion per year in the early 1980s to US$3.43 billion per year in the second half of that decade, based on 1988 prices (South African Military Review, 2012). The end of apartheid and the beginning of a democratic regime saw significant reductions in military expenditure. Between 1995 and the approval of the Military Review in 1998, the military budget was reduced from R10.9 to R9.5 billion, that is an 11.1 per cent (R1.4 billion) cut in nominal terms. The military budget further decreased from 1.54 per cent of GDP in 2004/05 and has levelled out in recent years at around 1.2 per cent to 1.1 per cent of GDP. While this may seem reasonably low, it is worth noting that South African military spending equals approximately 60 per cent of the total for sub-Saharan African.

**Methodology**

In setting up the empirical analysis the null hypothesis is Granger non-causality, defined as a situation when the information set on the first variable (e.g., military burden) does not improve the prediction of the second variable (e.g., GDP) over and above its own information. In the VAR framework, this implies testing whether the lagged values corresponding to the first variable are jointly significant or not, using Wald, Likelihood ratio (LR) and Lagrange multiplier (LM) tests. These tests assume that the underlying data is stationary and when this assumption does not hold, they may not have standard asymptotic distributions. The difficulties that arise when estimating these VAR models with non-stationary data have been shown by Park and Phillips (1989) and Toda and Phillips (1993, 1994), among others. One response includes the Toda and Yamamoto (1995) and Dolado and Lütkepohl (1996) modifications, which entail estimating a VAR($p+1$) and performing the Granger non-causality test on the first $p$ lags. This means one coefficient matrix, which relates to the $(p+1)^{th}$ lag, remains unrestricted under the null, giving the test a standard asymptotic distribution. In addition, residual-based bootstrap (RB) methods have been found to perform (in terms of power and size) considerably better than standard asymptotic tests, in a number of Monte Carlo simulation studies, regardless of the existence of cointegration or not (Horowitz, 1994; Shukur and Mantalos, 1997a, 1997b; Mantalos and Shukur, 1998; Shukur and Mantalos, 2000; Mantalos, 2000; Hacker and Hatemi-J, 2006). This led Balcilar and Ozdemir (2013) and Balcilar et al. (2013), to propose the use of RB modified-LR statistics to examine Granger causality and successfully apply them in considering the relation between growth and housing and growth and exports.
To illustrate the bootstrap modified-LR Granger causality, consider the following bivariate VAR($p$) process:

\[ z_t = \Phi_0 + \Phi_1 z_{t-1} + \cdots + \Phi_p z_{t-p} + \epsilon_t, \quad t = 1, 2, \ldots, T, \]  

where $\epsilon_t = (\epsilon_{1t}, \epsilon_{2t})'$ is a white noise process with zero mean and covariance matrix $\Sigma$ and $p$ is the lag order of the process. In the empirical section, the Akaike Information Criterion ($AIC$) is used to select the optimal lag order $p$. To simplify the representation, $z_t$ is partitioned into two sub-vectors, military expenditure ($m_t$) and GDP ($y_t$), so equation (1) becomes:

\[
\begin{bmatrix}
    m_t \\
    y_t
\end{bmatrix} = \begin{bmatrix}
    \phi_{m0} & \phi_{m(L)} & \phi_{m(L)} & \epsilon_{m_t} \\
    \phi_{y0} & \phi_{y(L)} & \phi_{y(L)} & \epsilon_{y_t}
\end{bmatrix},
\]

where $\phi_i(L) = \sum_{k=1}^{p} \phi_{ik} L^k$, $i, j = m, y$ and $L$ is the lag operator such that $L^i z_{it} = z_{i,t-k}$, $i = m, y$.

In this setting, the null hypothesis that GDP does not Granger cause military spending can be tested by imposing zero restrictions $\phi_{ym,i} = 0$ for $i = 1, 2, \ldots, p$. In other words, GDP does not contain predictive content, or is not causal, for military expenditure if the joint zero restrictions under the null hypothesis:

\[ H_0^M : \phi_{ym,1} = \phi_{ym,2} = \cdots = \phi_{ym,p} = 0 \]  

are not rejected. Analogously, the null hypothesis that military spending does not Granger cause GDP implies that we can impose zero restrictions $\phi_{ym,i} = 0$ for $i = 1, 2, \ldots, p$. In this case, the military spending does not contain predictive content, or is not causal, for GDP if the joint zero restrictions under the null hypothesis:

\[ H_0^Y : \phi_{ym,1} = \phi_{ym,2} = \cdots = \phi_{ym,p} = 0 \]

are not rejected. Taking the bootstrap approach pioneered by Efron (1979), which uses critical or $p$ values generated from the empirical distribution derived for the particular test using the sample data to test the hypotheses, and combining this with Toda and Yamamoto’s (1995) modified causality tests provides a test that applies to both cointegrated and non-cointegrated I(1) variables (Hacker and Hatemi-J, 2006).^5^

A further issue is that Granger non-causality tests assume that parameters of the VAR model are constant over time and this is often violated because of structural change. Although the presence of structural changes can be detected beforehand and the estimations can be modified to address this issue, using dummy variables and sample splitting for example, such an approach can introduce pre-test bias. A more satisfactory alternative is to use rolling bootstrap estimation, which applies the bootstrap causality test to rolling window subsamples for $t = \tau + l + 1, \tau + l, \ldots, \tau = \tau + l + 1, \ldots, T$, where $l$ is the size of the rolling window.

---

^4 In the case that both hypotheses in Eq. (3) and Eq. (4) are rejected, then we have the case of bidirectional causality. Bidirectional causality between military spending and economic growth implies a feedback system where both variables react to each other. If the hypothesis in Eq. (3) is rejected, then military spending Granger causes economic growth. Analogously, if the hypothesis in Eq. (4) is rejected, economic growth Granger causes military spending. It is also possible to have a case of no Granger causality in either direction implying that neither of the two variables have predictive content for each other.

^5 See the Appendix of Balcilar and Ozdemir (2013) for technical details of the bootstrap procedure.

^6 Indeed, Granger (1996) argued that parameter non-constancy was one of the most challenging issues confronting empirical studies.

^7 Details of the rolling window technique are also explained in the Appendix of Balcilar and Ozdemir (2013).
In the empirical analysis, the first step is to test for stationary of the data series, using the \( Z_u \) unit root test of Phillips (1987) and Philips and Perron (1988) (PP), Augmented Dickey Fuller (ADF) test and the \( MZ_u \) test of Ng and Perron (2001), and for cointegration between the variables, using the Johansen’s (1991) maximum likelihood cointegration method. The results of these tests will of course be sensitive to sample period used and the order of the VAR model, if the parameters are temporally unstable (Balcilar and Ozdemir, 2013). The results based on the full sample will also be invalid in the presence of structural breaks, so in what follows, tests for parameter stability in the estimated VAR models are undertaken.

In practice, a number of tests exist for examining the temporal stability of VAR models (e.g. Hansen, 1992; Andrews, 1993; Andrews and Ploberger, 1994). These can be applied in a straightforward manner to stationary models, but there is the possibility that the variables in the VAR models may be nonstationary and/or cointegrated, so both the long-run cointegrating relation parameters and the short-run dynamic adjustment parameters need to be investigated for stability. Given the super consistency of the estimators of the cointegration parameters, testing for parameter stability can be done in two steps. First, the stability of the cointegration parameters are tested using the Nyblom-Hansen \( L_c \) test (Nyblom, 1989; Hansen, 1992). This is an \( LM \) test for parameter constancy against the alternative hypothesis that the parameters follow a random walk process and are thus time-varying, since the first two moments of a random walk are time dependent (Balcilar et al., 2013). If the series are I(1), the Hansen–Nyblom \( L_c \) test still serves as stability test and can be interpreted as a test of cointegration (Balcilar et al., 2010). The \( L_c \) test is calculated using the fully modified OLS (FM-OLS) estimator of Phillips and Hansen (1990).

If the long-run parameters are stable, the \( Sup-F \), \( Ave-F \) and \( Exp-F \) tests developed by Andrews (1993) and Andrews and Ploberger (1994) are used to investigate the stability of the short-run parameters. These tests are computed from the sequence of \( LR \) statistics that test constant parameters against the alternative of a one-time structural change at each possible point of time in the full sample and exhibit non-standard asymptotic properties\(^8\). To avoid the use of asymptotic distributions, the critical values and \( p \)-values are obtained using the parametric bootstrap procedure. Specifically, the \( p \)-values are obtained from a bootstrap approximation to the null distribution of the test statistics, constructed by means of Monte Carlo simulation using 2000 samples generated from a VAR model with constant parameters. The \( Sup-F \), \( Ave-F \) and \( Exp-F \) tests need to be trimmed at the ends of the sample and following Andrews (1993) 15 percent is trimmed from both ends.

**Empirical Analysis and Results**

In this section we apply the procedure described above to the GDP and military spending series. The results of the ADF, PP and NP unit root tests including an intercept, as well as an intercept and trend in the test regression are reported in Table 1. These test statistics have nonstandard distributions and the response surface critical values computed by Mackinnon (1996) are used. The null hypothesis of nonstationarity could not be rejected for GDP and military spending at the 5 per cent significance level, but could for first differences, meaning the series are integrated of order one, i.e. I(1). Testing for a common stochastic trend, which implies a cointegrating relationship between GDP and military spending, is done using the Johansen’s (1991) maximum likelihood cointegration method. An optimal lag order of one for the VAR is suggested by the

\(^8\) The critical values are reported in Andrews (1993) and Andrews and Ploberger (1994)
Akaike Information Criterion (AIC) and the cointegration results based on the Trace and Maximum Eigen-value statistics are reported in Table 2. The null hypothesis of no cointegration could not be rejected at 5 per cent significance level, suggesting no long-run relationship between GDP and military expenditure.

Given that no cointegration is found between military spending and economic growth, the next step is to use a VAR rather than a VECM. Table 3 shows the estimation results for an optimal lag order of two, as indicated by AIC, and the Wald-statistics testing for Granger causality in the VAR. This fails to reject the null hypothesis that military expenditure does not Granger cause GDP at any of the conventional significance levels and the null that GDP does not Granger cause military expenditure, implying no Granger causality either way. As a check of robustness of the result, bootstrap LR-tests using the p-values obtained with 2000 replications were undertaken and provided similar results. Thus for the full sample of 1951 to 2010, there is no evidence of long-run nor short-run Granger causality between military expenditure and economic growth.

Table 1: Unit Root Tests

<table>
<thead>
<tr>
<th>Series</th>
<th>ADF</th>
<th>PP</th>
<th>NP</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Constant and trend</td>
<td>Constant and trend</td>
<td>Constant and trend</td>
</tr>
<tr>
<td>GDP</td>
<td>-2.322</td>
<td>-2.149</td>
<td>-2.620*</td>
</tr>
<tr>
<td>Military spending</td>
<td>-1.544</td>
<td>-1.442</td>
<td>-1.348</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Series</th>
<th>ADF</th>
<th>PP</th>
<th>NP</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Constant and trend</td>
<td>Constant and trend</td>
<td>Constant and trend</td>
</tr>
<tr>
<td>GDP</td>
<td>-4.368***</td>
<td>-4.822***</td>
<td>-4.857***</td>
</tr>
<tr>
<td>Military spending</td>
<td>-5.065***</td>
<td>-5.123***</td>
<td>-5.000***</td>
</tr>
</tbody>
</table>

Note: *, **, and *** denote significance at 10%, 5%, and 1%, respectively.

Table 2: Multivariate Cointegration Test Results

<table>
<thead>
<tr>
<th>Series</th>
<th>$H_0$</th>
<th>$H_1$</th>
<th>Trace Statistic</th>
<th>Maximum Eigenvalue Statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Military Spending and GDP</td>
<td>$r = 0$</td>
<td>$r &gt; 0$</td>
<td>9.01</td>
<td>8.20</td>
</tr>
<tr>
<td></td>
<td>$r \leq 1$</td>
<td>$r &gt; 1$</td>
<td>0.80</td>
<td>0.80</td>
</tr>
</tbody>
</table>

Notes: * indicates significance at the 5 per cent level.

* One-sided test of the null hypothesis ($H_0$) that the variables are not cointegrated against the alternative ($H_1$) of at least one cointegrating relationship. The critical values are taken from MacKinnon et al., (1992) with 5 per cent critical values equal to 15.49 for testing $r = 0$ and 3.84 for testing $r \leq 1$ for the Trace test. The corresponding values for the Maximum Eigenvalue tests are 14.26 and 3.84.
Table 3: Full Sample Granger Causality Tests

<table>
<thead>
<tr>
<th></th>
<th>$H_0$: Military spending does not Granger cause GDP</th>
<th>$H_1$: GDP does not Granger cause Military spending</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Statistics</td>
<td>$p$-value</td>
</tr>
<tr>
<td>$Wald$ test</td>
<td>1.641</td>
<td>0.461</td>
</tr>
<tr>
<td>$LR$ test</td>
<td>1.613</td>
<td>0.461</td>
</tr>
</tbody>
</table>

To investigate whether this result is underpinned by parameter constancy, the $Sup-F$, $Ave-F$, $Exp-F$ and $L_c$ tests were computed and the results are presented in Table 4. The results for the $L_c$ test of stability of the cointegration parameters indicate that the military equation has stable long-run parameters at the one per cent level, but not for the GDP equation. There is however no evidence of parameter constancy for the unrestricted VAR(2) model.

Table 4 Parameter Stability Tests for VAR(2) Model

<table>
<thead>
<tr>
<th></th>
<th>Military Equation</th>
<th>GDP Equation</th>
<th>VAR(2) System</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Statistics</td>
<td>Bootstrap</td>
<td>Statistics</td>
</tr>
<tr>
<td>$Sup-F$</td>
<td>42.84***</td>
<td>&lt;0.01</td>
<td>124.32***</td>
</tr>
<tr>
<td>$Ave-F$</td>
<td>11.44***</td>
<td>&lt;0.01</td>
<td>35.74***</td>
</tr>
<tr>
<td>$Exp-F$</td>
<td>17.66***</td>
<td>&lt;0.01</td>
<td>58.40***</td>
</tr>
<tr>
<td>$L_c$</td>
<td>0.59</td>
<td>0.56</td>
<td>1.63**</td>
</tr>
</tbody>
</table>

Notes: *, **, and *** denote significance at 10, 5 and 1 per cent, respectively.

*p-values are calculated using 2000 bootstrap repetitions.

Moving to consider the short-run parameter stability in this study, the results for the sequential $Sup-F$, $Ave-F$, and $Exp-F$ tests are reported in Table 4. The $Sup-F$ statistic tests parameter constancy against a one-time sharp shift in parameters, while the $Ave-F$ and $Exp-F$, are appropriate if the regime shift is gradual and assume the parameters follow a martingale process.9 These findings indicate instability in the short-run parameters of the VAR model, with evidence of both a one-time shift and gradual shifting of the parameters, implying that Granger causality tests based on the full sample VAR model are not reliable.

As a further test, the cointegration relation:

\[ GDP_t = \alpha + \beta MB_t + \varepsilon_t \]  

is estimated using the FM-OLS estimator, giving the results in Table 5. In this case the Nyblom-Hansen $L_c$ test does not reject the null hypothesis of cointegration at any reasonable level and the the $Mean-F$ and $Exp-F$ tests do not reject the null hypothesis of unchanging parameters in the cointegration equation, and so do not find evidence of a gradual shifting of the parameters of the

9 Both the $Ave-F$ and the $Exp-F$ statistics test the overall constancy of the parameters and are optimal tests as shown by Andrews and Ploberger (1994).
cointegration equation. The Sup-F test does suggest a one-time shift in the cointegration relationship.

Table 5: Parameter Stability Tests in Long-Run Relationship

<table>
<thead>
<tr>
<th></th>
<th>Sup-F</th>
<th>Ave-F</th>
<th>Exp-F</th>
<th>Lc</th>
</tr>
</thead>
<tbody>
<tr>
<td>GDP = ( \alpha + \beta \times MB )</td>
<td>534.88</td>
<td>330.77</td>
<td>264.43</td>
<td>0.09</td>
</tr>
<tr>
<td>Bootstrap p-value</td>
<td>&lt;0.01</td>
<td>1.00</td>
<td>1.00</td>
<td>0.85</td>
</tr>
</tbody>
</table>

Notes: We calculate p-values using 2,000 bootstrap repetitions.

While the tests are not completely consistent, they do indicate the presence of structural change in the dynamic relationship between economic growth and military expenditure.

To investigate this further the VAR model can be estimated using the rolling window regression techniques discussed above. This method entails performing the causality test using the residual based bootstrap method on a changing subsample of fixed length that moves sequentially from the beginning to the end of the sample by adding one observation from the front and dropping one from the end. With window size \( l \) and full series length \( T \), this provides a \( T-l \) sequence of causality tests.

An important choice parameter in rolling estimations is the window size \( l \) as the precision and representativeness of the subsample estimates are determined by the window size. Pesaran and Timmerman (2005), using root mean square error measures, show that the optimal window size depends on persistence and the size of the structural break. Their Monte Carlo simulations show that the bias in autoregressive (AR) parameters can be minimized by a window size as low as 20 when there are frequent breaks present. In determining the window size, there is a need to balance between two conflicting demands. First, the accuracy of parameter estimates, which depends on the degrees of freedom, is improved by a larger window size. Second, in the presence of multiple regime shifts a smaller window size reduces the probability of including some of these multiple shifts in the subsample (Balcilar and Ozdemir, 2013).

Based on the simulation results in Pesaran and Timmerman (2005) a window size of 15 is chosen\(^\text{10}\) and the VAR model in Eq. (1) is estimated for 15 years rolling through \( t = \tau - 14, \tau - 13, \ldots, \tau = 15, \ldots, T \). The bootstrap p-values of the null hypothesis that military expenditure does not Granger cause GDP and that GDP does not Granger cause military expenditure are then calculated using the residual based method. More precisely, the residual based p-values of the modified LR-statistics that test the absence of Granger causality from military to GDP or vice-versa are computed from the VAR(\( p+1 \)) defined in Eq. (2) fitted to a rolling window size of 15 observations.

The effect of military expenditure on GDP is then calculated as the mean of all bootstrap estimates, \( \sum_{k=1}^{N_b} \hat{\phi}_{m,y,k} \), where \( N_b \) equals the number of bootstrap repetitions and the effect of GDP on military spending is similarly calculated as \( \sum_{k=1}^{N_b} \hat{\phi}_{y,m,k} \). The estimates \( \hat{\phi}_{m,y,k} \) and \( \hat{\phi}_{y,m,k} \) are the bootstrap least squares estimates from the VAR in equation (2) estimated with the lag order of \( p \) determined by the AIC. The 90-percent confidence intervals are also calculated, where the lower and upper limits equal the 5\(^{\text{th}}\) and 95\(^{\text{th}}\) quantiles of each of \( \hat{\phi}_{m,y,k} \) and \( \hat{\phi}_{y,m,k} \), respectively.

\(^{10}\) This excludes the observations required for lags and hence is the actual number of observations in the VAR.
Bootstrap $p$-values of the rolling test statistics and the impact of each series on the other are shown graphically in Figures 1 to 4, with the horizontal axes showing the final observation in each of the 15-year rolling windows. Figure 1 presents the bootstrap $p$-values of the rolling test statistics, testing the null hypothesis that military expenditure ($MB$) does not Granger-cause GDP and shows that the null hypothesis is not rejected for most of the periods at the 10 per cent significance level. The only rejections are during the 1973, 1982 and 2000-2002 subperiods. Figure 2 shows the bootstrap estimates of the sum of the rolling coefficients for the impact of military expenditure on GDP. The results suggest that at a 10 per cent level of significance, military expenditure has positive predictive power for GDP during the 1973-1975 subperiod with a mean coefficient of about 0.05, but negative predictive power during the 1998-2005 subperiod with coefficients ranging between -0.04 to -0.07.

An interesting observation is the positive effect of military spending on growth from the beginning of the analysis in the mid-sixties up to the mid-seventies, even though this effect was only statistically significant for the period 1973 to 1975. Starting as early as 1966, South Africa was involved in a counter insurgency war against SWAPO (South West Africa People's Organisation). During 1972 conscription (national service) was increased from 9 months to one year for all white males 17 years and older, and by the middle of 1974, control of the northern part of South West Africa (present day Namibia) was handed over to the South African Defence Force (SADF) and the South African Police (SAP). The conflict deepened when the Popular Movement for the Liberation of Angola (MPLA), aided by Cuba, also got involved in the struggle after the independence of Angola in 1975. During this time South Africa sided with the Angolan rival UNITA party against the MPLA's armed force.

It was also during this period, in 1968, that Armscor, a South African government-supported weapon producing conglomerate was officially established, primarily in response to a tightening UN arms embargo. Armscor produced small arms ammunition as well as heavy armament, in addition to sophisticated military aircraft and vehicles. They also produced for the export market, including export destinations like Iraq. The positive effect on growth during the sixties and seventies on growth is therefore likely attributable to a Keynesian demand side channel.

The rise in military spending continued until 1977, but from the evidence here its initial demand stimulus effect had started to become a negative overall effect on growth by this time (Batchelor et al., 2002). After that the effects were mainly negative, possibly reflecting the effect of misallocation of investment from productive to less productive sectors, namely the defence and other strategic industries, as argued by Batchelor et al. (2000a). Conscription was also once again increased in 1977, this time from one year to two years and 30 days annually for 8 years, representing a further displacement of productive factors of production.

By this time, the country has also seen the Soweto uprising of 1976 and the death of the political activist, Steve Biko in September 1977, and by 1980 international opinion has turned decisively against the apartheid regime. During the 1980s the state was pre-occupied with security and much effort and resources went into nuclear and biological warfare research. By July 1985 a State of Emergency was declared and South Africa experienced a host of cultural, political, economic sanctions from the international community. Apartheid was however dismantled in a series of negotiations from 1990 to 1993, starting with the release of Nelson Mandela in February 1990 and finally culminating in democratic elections in 1994, and the establishment of a democracy. During all this time, from 1997 leading up to 1994, no significant impact of military spending on growth is evident from the analysis. Between 1994 and 2004 government embarked upon an
integration process of the SADF with forces from freedom movements as well as defence forces of the formerly independent homelands, bearing a cost no necessarily justified by economic productivity or efficiency, and thereby contributing to the significant negative impact on output and growth that we observe for the period between 1998 and 2005. Also, by 1998, the newly elected ANC government announced a military procurement package of weaponry that involved US$4.8 (R30 billion in 1999 rands). The deal has been subject to repeated allegations of corruption. Towards the end of this subsample, the significantly negative impact subsided, and disappeared after 2005.

Figure 1: Bootstrap $p$-values of $LR$ test statistic testing the null hypothesis that $MB$ does not Granger cause $GDP$

Figure 2: Bootstrap estimates of the sum of the rolling window coefficients for the impact of $MB$ on $GDP$

Figure 3 shows the bootstrap $p$-values of the rolling test statistics, testing the null hypothesis that $GDP$ does not Granger-cause military expenditure ($MB$). Again, the non-causality tests are
evaluated at 10 per cent significance level. The figure shows that the null hypothesis that GDP does not Granger-cause military expenditure is rejected at a 10 per cent significance level for only three periods namely 1981, 1995 and 1997. Figure 4 shows the bootstrap estimates of the sum of the rolling coefficients for the impact of GDP on military expenditure and these suggest that GDP has positive predictive power for military expenditure during the 1966-1972, 1975, 1977, 1979-1984 subperiods with coefficients ranging between 0.1 and 1.2, but negative predictive power during the 1995-2001 subperiod with coefficients ranging between -2.04 to -0.57. Again, this pattern makes sense, with the affordability of the military burden declining over time and the large declines in military spending from 1989 when the security situation allowed. These included, the end of the Cold War, reduction in neighbouring countries’ military spending, withdrawal from Namibia and the reforms within South Africa, leading to the unbanning of the opposition groups and the release of Nelson Mandela (Batchelor et al., 2002)

Figure 3: Bootstrap $p$-values of $LR$ test statistic testing the null hypothesis that GDP does not Granger cause MB

Comment [R1]: Bottom line of frame not there.
Overall, our bootstrap rolling window Granger causality results support the hypothesis of bidirectional causality between military expenditure and economic growth. This is in contrast to the results using the full sample which could not establish any predictive content from military expenditure to GDP and *vice versa*. They also seem to fit with strategic and economic developments within South Africa. These findings point to the fact that using the standard Granger causality tests which neither accounts for structural breaks nor time variation in the relationship between economic variables may be misleading.

**Conclusion**

This paper contributes to the growing literature on the military spending and economic growth nexus, by providing a case study of South Africa and considering the possibilities of structural breaks in the relationship, using techniques that allow inference whether or not the series are integrated-cointegrated. Full sample bootstrap Granger non-causality tests suggested no significant Granger causality in either direction between that military expenditure and economic growth, a result that could be anticipated. Parameter stability tests, however, find the estimated VARs to be unstable, suggesting the inference may be invalid. Allowing for structural change by using the bootstrap rolling window estimation, it is found that military expenditure has positive predictive power for GDP during the 1966-1972, 1973-1975, 1975, 1977, and 1979-1984 subperiods but negative predictive power during 1995-2001 and 1998-2005 subperiods. These results support bidirectional causality and suggest that military spending may have had a positive effect on growth in the earlier apartheid period, but not later. The value of the approach is further reinforced by the fact that the pattern of the results seems to align with and support strategic and economic developments within South Africa. This means that the causal relation between military expenditure and economic growth within a country is likely to be non-linear,
asymmetric and time varying and that future research should try to take account of these properties.

References


