Time-Varying Causality between Research Output and Economic Growth in US

Roula Inglesi-Lotz* (roula.inglesi-lotz@up.ac.za)**, Mehmet Balcilar** (mehmet@mbalcilar.net) **and Rangan Gupta** (rangan.gupta@up.ac.za)

University of Pretoria, Department of Economics

Abstract:

This main purpose of this paper is to investigate the causal relationship between knowledge (research output) and economic growth in US over 1981 to 2011. To overcome the issues of ignoring possible instability and hence, falsely assuming a constant relationship through the years, we use bootstrapped Granger non-causality tests with fixed-size rolling-window to analyze time-varying causal links between two series. Instead of just performing causality tests on the full sample which assumes a single causality relationship, we also perform Granger causality tests on the rolling sub-samples with a fixed-window size. Unlike the full-sample Granger causality test, this method allows us to capture any structural shifts in the model, as well as, the evolution of causal relationships between sub-periods, with the bootstrapping approach controlling for small-sample bias. Full-sample bootstrap causality tests reveal no causal relationship between research and growth in the US. Further, parameter stability tests indicate that there were structural shifts in the relationship, and hence, we cannot entirely rely on full-sample results. The bootstrap rolling-window causality tests show that during the sub-periods of 2003-2005 and 2009, GDP Granger caused research output; while in 2010, the causality ran in the opposite direction. Using a two-state regime switching vector smooth autoregressive model, we find unidirectional Granger causality from research output to GDP in the full sample.

Keywords: Research Output; Scientometrics; Economic Growth; Causality

* Corresponding author

1. Introduction

Studying of the possible effects of improved human capital on economic growth is not novel or recent. For example, Romer (1986) argued that a firm's productivity level is higher, the higher the average knowledge stock acquired by its labour is. Past theoretical studies (Romer, 1986; Lucas, 1988 Tamura, 1991; Schumpeter, 2000) and applied studies (Price, 1978; Kealey, 1996; De Moya-Anego and Herrero Solana, 1999; King, 2004; Fedderke, 2005; Fedderke and Schirmer, 2006; Vinkler, 2008; Lee et al., 2011; Shelton and Leydersdorff, 2011; Inglesi-Lotz and Pouris, 2013; Inglesi-Lotz et al. 2013) have shown that there is some evidence on this relationship stressing out that improved human capital can be demonstrated through the accumulation of knowledge. From a microeconomic side, knowledge externalities are considered positive for the economic productive capacity of a company but also macroeconomically, higher degrees of knowledge and hence, better quality of labor provides a country with numerous advantages with regards to its innovation, development and economic growth.

The question that arises first is how quality of human capital can be measured and how it can be further improved. A number of activities can assist towards that direction such as higher education, training and life education. Research activities, such as reading local and international literature, learning new methods and producing academic papers (research output) can improve the academic human capital that is mostly responsible for the level of human capital of a country (Inglesi-Lotz and Pouris, 2013).

Although, economic growth can easily be measured by the country's Gross Domestic Product (GDP) or GDP per capita, measuring a country's level of knowledge is challenging. In the literature, a number of different indicators were used such as expenditures on Research and

Development (R&D) (Fedderke and Schirmer, 2006) or various scientometric indicators (De Moya-Anegon & Herrero Solana, 1999; King, 2004; Vinkler, 2008; Lee et al. 2011; Inglesi-Lotz and Pouris, 2013; Inglesi-Lotz et al. 2013). These indicators may refer to the quantity of research output (number of published academic papers), specific quantity or share (number of published academic papers per capita or share of a country's published academic papers to the world) or impact to the literature (number of citations or average number of citations per published academic paper). Pouris and Pouris (2009) have analysed the superiority of scientometric indicators in such studies by mentioning that scientometric analysis is one of the most objective and straightforward ways of measuring the research performance of a country.

The exact statistical existence of the relationship between accumulated knowledge and economic growth as well as the direction of this relationship is still under debate. While the studies mentioned above agree in the existence of this link; however, causality can run from any of the two variables to the other. On the one side, countries with higher economic growth can promote better knowledge opportunities and hence better quality of human capital; while on the other side, it is this improved human capital that can enhance further the levels of economic growth and development. Lee et al. (2011) argue that the direction depends highly on the developmental stage of a country: weaker to no relationship for the developed economies of their study while stronger for the developing ones. The unambiguity in the direction of causality found in the international literature can be attributed to the country's level of economic growth and development or different periods examined or dissimilar academic and research systems.

While most studies assume that the existence and direction of the causality remain the same over long time periods, changes in the developmental level as well as policies in research and higher education might be responsible for altering the relationship between economic growth and research output from year to year. In the literature mostly full sample Granger causality tests were employed to establish the existence and the direction of this causality either in a single-country analysis (Inglesi-Lotz and Pouris, 2013) or multi-country analysis (Vinkler, 2008; Lee et al. 2011; Inglesi-Lotz et al. 2013). The Granger causality tests ignore structural shifts or instability in an economy; fact that may result in misleading results. Hence testing for instabilities is of paramount importance in an ever-changing world.

Against this backdrop, the objective of this paper is to investigate the relationship between knowledge (research output) and economic growth in US from 1981 to 2011. To overcome the issues of instability and the – possibly – wrong assumption of a constant relationship through the years, we use the approach developed by Balcilar et al. (2010) which involves using bootstrap Granger non-causality tests with fixed-size rolling sub-samples to analyze the time-varying causal links between the two series. This method allows us to capture any structural shifts in the model, as well as, the evolution of causal relationships between sub-periods, with the bootstrapping approach controlling for small-sample bias. Based on the time varying Granger causality relationship indicated by the bootstrap rolling causality tests, we build a vector smooth transition autoregressive (VSTAR) model. The nonlinear VSTAR model with two states, which fits well to the known recession and recovery periods, allows us to test for state dependent Granger causality. Using bootstrap approach to obtain the p-values, a VSTAR model indicates Granger causality from research output to GDP, but not from GDP to research output. Thus, we also obtain full sample evidence that research output Granger causes GDP.

The paper is structured as follows. The next section discusses the econometric method and the data. Section 3 presents the empirical results of the exercise while finally Section 4 discusses the policy implications and meanings of the results and concludes.

2. Methodology and Data

The main purpose of this paper is to specify the existence and direction of the causality between a country's research output, proxied with the share of the country's number of publications to the world, and its real GDP. The null hypothesis of the test states Granger non-causality from the one variable to the other. If the information on the one variable (i.e. research output) does not provide any improvement to the prediction of the second one (GDP) over and above its own information, then we can conclude that the first variable does not Granger cause the second one.

However, the standard Granger causality tests do not take into account possible non-stationarity in the time series. In that case, standard asymptotic distribution theory does not hold. To overcome this predicament, Toda and Yamamoto (1995) and Dolado and Lutkepohl (1996) propose the estimation of a VAR ($p+1$) in levels, where $p+1$ is the lag order. To use their method, the series should be confirmed to be I(1). Thus, the Granger causality tests remain valid without being dependent on the order of integration/cointegration of the variables (Hacker and Hatemi-J, 2006). The intuition behind this estimation is that the coefficient matrix now relates to the ($p+1$)st lag and is unrestricted under the null. This allows the test for a standard asymptotic distribution.

A number of studies (Shukur and Mantalos, 1997a, 1997b, 1998; Mantalos, 2000; Hacker and Hatemi-J, 2006) compared different Granger non-causality tests to conclude that the residual bootstrap (RB) based modified-*LR* statistics are superior to other tests in many aspects such as the power and size properties of the tests. Also, numerous studies have recognized the robustness of the bootstrap approach to test for Granger causality (Efron, 1979; Horowitz, 1994; Mantalos and Shukur, 1998; Mantalos, 2000). For a more detailed discussion on this, see Balcilar et al.

(2013). Due to these reasons, in this paper, we are using the bootstrap approach with the Toda and Yamamoto (1995) modified causality tests. We generate the bootstrap samples using the parametric bootstrap approach. We resample from residuals with replacement and generated the data for variables using the restricted VAR model under the null hypothesis.

To illustrate the bootstrap modified-*LR* Granger causality test procedure, consider the following bivariate *VAR*(*p*) process[1]:

$$z_t = \Phi_0 + \Phi_1 z_{t-1} + \ \dots \ + \Phi_p z_{t-p} + \varepsilon_t , \quad t = 1,\ 2,\ \dots ,T , \tag{1}$$

where $\varepsilon_t = (\varepsilon_{1t}, \varepsilon_{2t})'$ is a white noise process with zero mean and covariance matrix Σ and p is the lag order of the process. In the empirical section, we use the Schwarz Information Criterion (*SIC*) to select the lag order *p*. To simplify, we partition z_t into two sub-vectors, the research output (z_{rt}) and real GDP (z_{yt}) and rewrite equation (1) as follows:

$$\begin{bmatrix} z_{rt} \\ z_{yt} \end{bmatrix} = \begin{bmatrix} \varphi_{r0} \\ \varphi_{y0} \end{bmatrix} + \begin{bmatrix} \varphi_{rr}(L) & \varphi_{ry}(L) \\ \varphi_{yr}(L) & \varphi_{yy}(L) \end{bmatrix} \begin{bmatrix} z_{rt} \\ z_{yt} \end{bmatrix} + \begin{bmatrix} \varepsilon_{rt} \\ \varepsilon_{yt} \end{bmatrix} \tag{2}$$

where $\phi_{i\,j}(L) = \sum_{k=1}^{p+1} \phi_{i\,j,k} L^k$, $i, j = r, y$ and L is the lag operator such that $L^k z_{it} = z_{it-k}$, $i = r, y$. In this setting, the null hypothesis that real GDP output does not Granger cause the research output implies that we can impose zero restrictions $\varphi_{ry,i}=0$ for $i = 1,\ 2,\ \dots ,\ p$. In other words, real GDP does not contain predictive content, or is not causal, for the research output when we cannot reject the joint zero restrictions under the null hypothesis:

$$H_0 \colon \varphi_{ry,1} = \varphi_{ry,2} = \cdots = \varphi_{ry,p} = 0 \tag{3}$$

[1] The details of the bootstrap *LR*-modified Granger causality test can be found in Balcilar et al. (2010).

Analogously, the null hypothesis that the research output does not Granger cause real GDP implies that we can impose zero restrictions $\varphi_{rh,i} = 0$ for $i = 1, 2, \ldots, p$. Now, the research output does not contain predictive content, or is not causal, for real GDP when we cannot reject the joint zero restrictions under the null hypothesis:

$$H_0: \varphi_{yr,1} = \varphi_{yr,2} = \cdots = \varphi_{yr,p} = 0 \tag{4}$$

All the standard Granger non-causality tests make a strong assumption that the VAR model's parameters remain constant over time. In an ever-changing socio-economic environment, this assumption hardly ever holds being a puzzling topic for economic empirical studies (Granger, 1996). Common practice would be to test for the presence of structural breaks in advance and modify the estimation in various ways, for example, with the use of dummy variables or sample splitting. All of these methods, however, introduce some pre-test bias so we employ a rolling bootstrap estimation to account for parameter instability. Structural changes may change the pattern of the causal relationship between the two variables over time. We will be testing for parameter inconsistency in the full sample. However, if the parameters prove to be unstable, the Granger causality tests and the cointegration tests to the full sample are proven invalid. All in all, parameter instability can occur in many ways. That is the reason why tests that leave the alternative specifications against the null hypothesis unspecified are preferred. Given the difficulty in test selection, we use several tests, namely, *Sup-F*, *Mean-F*, *Exp-F* (Andrews, 1993; Andrews and Ploberger, 1994) and L_c (Hansen, 1992) tests, based on their optimality properties. (2010).

There are four common approaches, commonly employed in econometric applications, for estimation when structural breaks are exists: recursive estimation, rolling estimation, regime switching, and time varying parameters (TVP). Recursive and TVP estimation are analogue, as they keep the lower end of the estimation window and move towards and with a grooving window. As the window grows, it accumulates more information and when they reach the last observation, they will be equivalent to the full sample estimation. If the parameters are stable, then the recursive and TVP estimators will converge to the constant parameters as the sample size grows with increasing window size. This implies that successive prediction errors will diminish for the estimate of the parameters, as the information already incorporated in the estimation increases. A consequence of this is that all previous observations will have impact on the successive estimates. In the presence of multiple structural breaks such an approach is not optimal since the impact of previous breaks on the later ones will not be isolated.

In case of multiple breaks, it is preferable to give more weight to recent observations and discard the data that has reached certain age and passed the date of expiry. One way of better accommodating parameter variability is then to base the estimation only on the most recent portion of the data. This leads to rolling estimation, which is used in this study. Our preference for rolling estimation is based on its better capability to accommodate parameter variations, particularly multiple ones.

Stock and Watson (1996) use TVP and rolling estimation and find that they equivalently outperform the other approaches. In application to time varying betas, Groenewold and Fraser (1999) conclude that rolling estimation shows greatest variation in the sample and thus better captures the structural breaks. Barnett et al. (2012) also find that rolling estimation slightly outperforms other approaches, including TVP.

Regime switching models assume a certain mechanism for the parameter variation. Threshold, smooth transition, and Markov switching are among these and successfully used in applications.

When a system evolves switching among a finite number of states, then we can use regime-switching models to describe this dynamic evolution. The most well-known state switching in economic data relates to the business cycle. Regime switching models have been used successfully to model the business cycle.

Our study first focuses on showing the existence of instabilities and their influence on the causality tests. Second, it tests for Granger causality by taking into account the form of the parameter instabilities. We utilize rolling estimation for the first task. We utilize two approaches to isolate the impact of parameter instability on the causality tests. First, rolling bootstrap causality tests are used so that that sample period used is sufficiently homogenous and not affected by the structural breaks. Second, VSTAR model is used to model the state switching and causality is tested within this model, which uses full sample information.

To sum up the methodology that we use here: a) we specify the order of integration of the two series using the Phillips (1987) and Phillips Perron (1988) test. We test for all three types of specifications (constant/ constant with time trend/ none); b) we test the existence of cointegration by using the L_c test performed on the long-run relationship between our two variables of concern, with the long-run equation being estimated based on FM-OLS method; c) we perform Granger non-causality tests for the full-sample to identify if there is an overall causal relationship; d) we test for parameter stability of the short-run coefficient estimates based on the *Sup-F*, *Mean-F* and *Exp-F*; e) we estimate rolling VAR regressions and employ Granger non-causality tests with a fixed 15-year window, if structural breaks are detected; and f) we use VSTAR model to test for state dependent causality, which uses full sample information.

3. Data

In this paper, we test for a causal relationship between the research output of the US economy and real GDP, using annual data from 1981 to 2011. The GDP data come from World Development Indicators (WDI) at constant 2005 US dollars. For the second important variable of the analysis (research output), we follow the Inglesi-Lotz and Pouris (2013) approach of proxying research output by the share of number of publications of the country to the rest of the world. Inglesi-Lotz and Pouris (2013) argue that "the link between economic growth of and the growth in the number of publications in a country should be measured vis-à-vis the research performance of the rest of the world. It is research and innovation performance vis-à-vis the rest of the world that may lead to economic growth…Furthermore, such an approach neutralizes the fact that Thomson Reuters, in their indexing efforts, changes the set of journals indexed from time-to-time" (Inglesi-Lotz and Pouris, 2013: 132). This indicator is derived by the Institute for Scientific Information (ISI) Thomson Reuters family of databases is employed. In the National Science Indicators database, the ISI counts articles, notes, reviews and proceeding papers, but not other types of items and journal marginalia such as editorials, letters, corrections, and abstracts (Inglesi-Lotz and Pouris 2011).

4. Empirical results

As discussed in Section 3, we will follow a five step method to investigate the relationship between the research output (proxied by the share of number of publications to the rest of the world) and the GDP of the US economy.

a) Order of integration

To identify the order of integration and existence of non-stationarity of the two series, we use the Phillips (1987) and Phillips and Perron (1988) test. We included all three different specifications: constant, constant with time trend and none of the two. The MacKinnon (1996) one-sided p-values are used the test's critical values. According to the Table 1 that reports the results, the null hypothesis of non-stationarity cannot be rejected at levels. However it can be rejected for the first differences of the two series, implying that both our series are I(1).

Table 1: Phillips- Perron unit root test results

	Level			First differences		
Series	**Constant**	**Constant & trend**	**None**	**Constant**	**Constant & trend**	**None**
Real GDP	-1.719	-0.427	6.047	-4.181***	-5.128***	-1.827*
Research output	2.452	-1.388	-3.551***	-4.623***	-5.446***	-2.902***

Note: * (**) [***] denotes 10% (5%) [1%] level of significance.

b) Cointegration tests

Next we test for the existence of cointegration among the two series using the L_c test, by estimating the following cointegration equation between the two variables in question:

$$GDP_t = \alpha + \beta \, Research\ output_t + \varepsilon_t \qquad (5)$$

The parameters of (5) are estimated using the FM-OLS estimator. In Table 2, the results of the various parameter stability tests are presented. The L_c test rejects the null of hypothesis of parameter stability, implying lack of cointegration.[2] Structural breaks in the long-run relationship are also overwhelmingly supported by the *Sup-F*, *Mean-F* and *Exp-F* statistics.

[2] The lack of cointegration was also confirmed by the Trace and Maximum Eigen-Value statistics proposed by Johansen (1991). The details of these results are available upon request from the authors.

Table 2: Parameter stability tests in long-run relationship FM-OLS

	Mean-F	*Exp-F*	*Sup-F*	L_c
GDP=α+β* Research output	33.31	30.14	66.65	3.42
Bootstrap p-value	<0.01	0.01	<0.01	0.01

c) Full sample Granger non-causality tests

The lack of cointegration cannot influence the exercise because as Balcilar et al. (2010) state, the variables might exhibit Granger temporal causality. In Table 3, we present the results of from the bootstrap *LR* causality test performed on a VAR model of order 1, with the lag-length being chosen based on the *SIC*.. The test fails to reject the null hypotheses of Granger non-causality, thus implying that there are no causal links between research output and GDP for US.[3]

Table 2: Full-sample Granger non-causality tests

	H_0: Research output does not Granger cause Real GDP		**H_0: Real GDP does not Granger cause Research output**	
	Statistics	**p-value**	**Statistics**	**p-value**
Bootstrap *LR* Test	2.439	0.306	1.196	0.295

d) Parameter stability tests

Table 4 reports the results of the parameter constancy tests that investigate the temporal stability of the coefficients of the VAR model. The *p*-values come from a bootstrap approximation to the null distribution of the test statistics. Both *Mean-F* and *Exp-F* statistics test for the overall constancy of the parameters. The *Mean-F* statistic imply that there is evidence of parameter non-constancy for the GDP and research output equations but not for the case of the VAR(1) as a

[3] Granger causality results based on standard (non-bootstrapped) F-tests also yielded similar results, i.e., no causality could be detected between the two variables at standard levels of significance.

system system. The *Exp-F* test's results show some instability at the GDP equation but not for the Research output equation and the VAR(1) system. On the other hand, the *Sup-F* statistic, which tests for parameter constancy against the alternative of a one-time sharp shift in parameters, shows no evidence for parameter non-constancy for both the GDP and Research output equations, but one-time shift could not be rejected for the VAR(1) system.

Table 3: Parameter stability tests in VAR(1) model

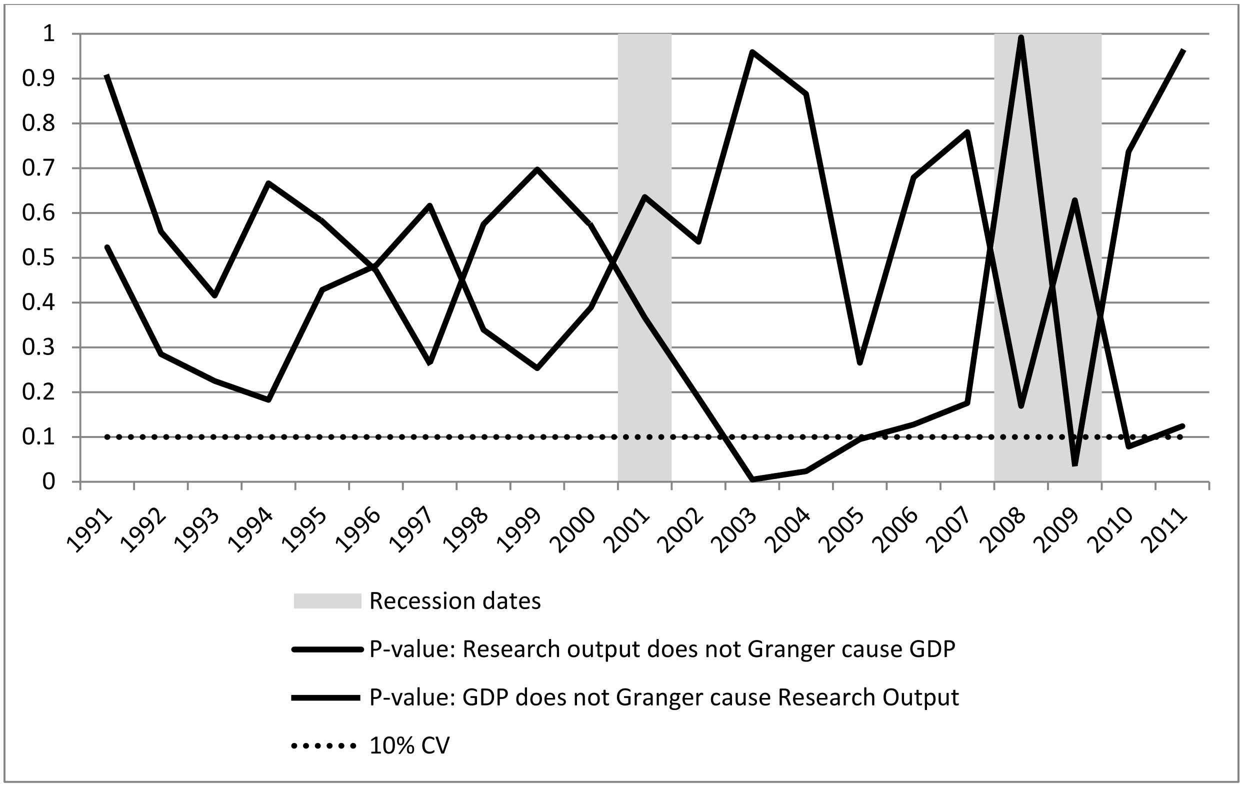
	GDP equation		**Research output Equation**		**VAR(1) system**	
	Statistics	**Bootstrap *p*-value**	**Statistics**	**Bootstrap *p*-value**	**Statistics**	**Bootstrap *p*-value**
Mean-F	8.21	0.01	32.33	<0.01	6.665	0.33
Sup-F	29.24	<0.01	285.59	<0.01	16.28	0.12
Exp-F	12.06	<0.01	139.66	1.00	5.56	0.13

All in all, there is clear evidence of parameter instability.

e) Bootstrap rolling causality tests

Given the existence of parameter instability, full-sample causality tests can be relied upon and hence, Figure 1 illustrates the bootstrap *p*-values of the rolling test statistics based on a fixed window-size of 10. According to Balcilar et al. (2010) the choice of the window size is an important aspect to consider as it determines the number of rolling estimates. They state that the larger the window size, the greater the precision of estimates although in the presence of heterogeneity there may be less representativeness of parameters. On the other hand, a smaller fixed-window size may increase representativeness and heterogeneity but may lead to large standard errors which result in biased parameter estimates. When choosing our window size, *l*, we have to take into account these two aspects and try and establish a balance between accuracy and representativeness. We opt for a smaller window size of 10 to guard against heterogeneity. For a small window size, the bootstrap method applied to 21 sub-sample-based causality tests tends to produce more precise estimates.[4] Also a 10 year window allows us to include periods immediately after the recession in 1990.

Figure 1: Granger causality test *p*-values: Rolling-window estimates



Note: The grey areas in the graph denote the recession dates as they are reported by the National Bureau of Economic Research (NBER).

The null hypothesis is that the GDP does not Granger cause Research output and vice versa. The non-causality hypothesis is tested at the 10% percent level of significance. The *p*-values change over the whole sample. The null hypothesis that Research output does not Granger cause GDP cannot be rejected for almost the entire sample, with the exception of 2010. On the other hand, the hypothesis that GDP does not Granger cause Research output can be rejected for the period 2003 to 2005 and 2009, at the 10% significance level.

[4] Our conclusions are unchanged with a window size of 15. The graph of this analysis presenting the Granger causality test p-values can be found in the Appendix.

The US experienced a short recession in 2001 causing a collapse of the stock market and its impact to GDP induced an overall global downturn in economic activity. The September 11 attacks also contributed to have a negative effect on US and global markets. After this recession period, our results showed that GDP Granger caused research output. It was the period that higher growth again stimulated research activities in an effort to improve the quality of human capital. Also at the end of the 2007-2010 recession periods, the findings showed that research output does Granger cause GDP. All in all, we can observe that the lack of causality between research output and economic growth is affected after periods of recession experienced in the US economy.

f) Regime switching and state dependent causality tests

Rolling estimation results indicate significant time variation in the parameter estimates and causality relationships. The time variation in the causality relationships does however vary with the state of the GDP growth. We find causality in recession periods from GDP growth to research output and also from research output to GDP growth in the beginning of the recovery period after 2007-2010 recessions. This strongly points out the regime-switching (nonlinear) nature of the series distort the causality test results from linear models. We further observe that it is the switching in and out of recessions that describes the causality shifts, implying causality depends on the state of the economy.[5]

In order check causality in a nonlinear causality in the full sample we employ a vector smooth transition autoregressive (VSTAR) model. VSTAR models are regime switching models where the state of the economy is determined be a state or transition variable. Recent empirical studies

[5] We thank an anonymous referee for pointing out the state dependent causality and suggesting a nonlinear causality test.

show that (VSTAR) models can successfully model economic time series that move smoothly between two or more regimes (e.g., recession to expansion). When considering the joint dynamic properties of the real GDP and the research output, it is natural to consider vector STAR (VSTAR) models. Recent applications (e.g., Rothman, *et al.*, 2001; Psaradakis, *et al.*, 2005; Tsay, 1998; De Gooijer and Vidiella-i-Anguera, 2004) find that VSTAR models successfully model nonlinear economic time-series data.

In our case, we specify the two-regime two-dimensional VSTAR model as follows:

$$z_t = (\Phi_{1,0} + \sum_{j=1}^{p} \Phi_{1,j} z_{t-j}) + (\Phi_{2,0} + \sum_{j=1}^{p} \Phi_{2,j} z_{t-j}) G(s_t; \gamma, c) + \varepsilon_t, \qquad (6)$$

where $\Phi_{i,0}$, $i = 1,2$, are (2×1) vectors, $\Phi_{i,j}$, $i = 1,2$, $j = 1,2,..., p$, are (2×2) matrices, and $\varepsilon_t = (\varepsilon_{rt}, \varepsilon_{yt})$ is a *k*-dimensional vector of white noise processes with zero mean and nonsingular covariance matrix Ω, $G(\cdot)$ is the transition function that controls smooth moves between the two states (regimes), and s_t is the transition variable.

The VSTAR model in equation (6) defines for two states, one associated with $G(s_t; \gamma, c) = 0$ and another associated with $G(s_t; \gamma, c) = 1$. The transition from one state to the other occurs smoothly, depending on the shape of the $G(\cdot)$ function. In this paper, we consider a logistic transition function

$$G(s_t; \gamma, c) = \frac{1}{1 + \exp\{-\gamma(s_t - c)/\hat{\sigma}_s\}}, \qquad \gamma > 0, \qquad (7)$$

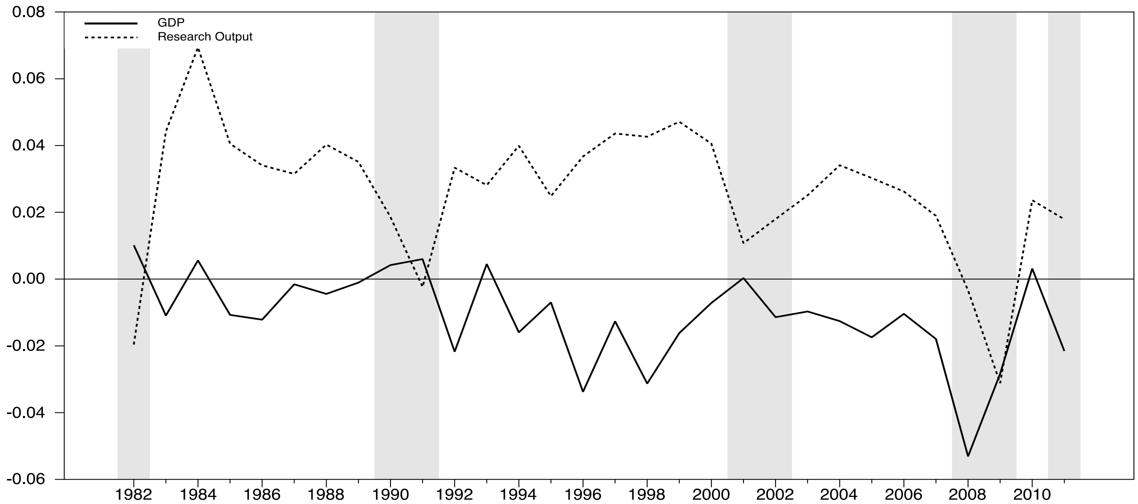
where $\hat{\sigma}_s$ is the estimate of the standard deviation of transition variable s_t. The threshold parameter c determines the midpoint between two regimes at $G(c; \gamma, c) = 0.5$. The parameter γ determines the speed of transition between the regimes with higher values corresponding to faster transition.

To specify the VSTAR model, we follow the procedure presented in Terasvirta (1998) (see, also, Van Dijk, *et al*. 2002; Lundbergh and Terasvirta, 2002). First, we specify the lag order of p =1, selected by the BIC. Second, we test linearity against the VSTAR alternative. Since the VSTAR model contains parameters not identified under the alternative, we follow the approach of Luukkonen *et al*. (1988) and replace the transition function $G(\times)$ with a suitable Taylor approximation to overcome the nuisance parameter problem. The testing procedure selects a logistic VSTAR model with a single threshold, which we maintain for the univariate case as well.

Third, we select the transition variable s_t. To identify the appropriate transition variable, we run the linearity tests for several candidates, $s_{1t}, s_{2t}, ..., s_{mt}$, and select the one that gives the smallest *p*-value for the test statistic. Here, we consider lagged values of both variables for lags 1 to 2 as the candidate transition variable. Let $s_t = x_{i,t-d}$, where x is any of the two variables $\{z_r, z_y\}$. We test linearity with these variables for delays $d = 1,2$. We obtain the smallest *p*-value with $s_t = z_{y,t-d}$ and $d = 1$. Given the selections $s_t = z_{y,t-1}$and *p*=1, we estimate the parameter of the VSTAR model given in (6)-(7) using nonlinear least squares.

Figure 2 shows the GDP growth rate and research output growth rate along with the classification of the states based on estimates of the threshold parameter *c*. State 1 corresponds to a low growth recessions regime and gray shaded periods are estimated as recession states. We observe that the VSTAR model consistently and accurately estimates the rescession and expansion states of the US economy.

Figure 2. GDP Growth, Research Output Growth and Estimates of Recession States



VSTAR framework allows us to test three full sample non-causality hypotheses for each of the variables:

i) z_i does not Granger cause z_j in the first state, $i,j=r,y$ and $i \neq j$: $H_0: \phi_{ij,1}(L)=0,\ i \neq j$

ii) z_i does not Granger cause z_j in the second state, $i,j=r,y$ and $i \neq j$: $H_0: \phi_{ij,2}(L)=0,\ i \neq j$

iii) z_i does not Granger cause z_j in the first and second states jointly, $i,j=r,y$ and $i \neq j$: $H_0: \phi_{ij,1}(L)=\phi_{ij,2}(L)=0,\ i \neq j$

These restrictions are only imposed on the piecewise linear components of the model and can be calculates as Wald tests. Since, we have a small sample size we obtain the *p*-values of the tests using 1000 parametric bootstrap, where the residuals are sampled from the models under the null and the bootstrap samples are generated using the parameter estimates under the alternative. Test results are given in Table 5.

Table 5: State Dependent Full Sample Granger Causality Tests

	Wald Statistic	Bootstrap *p*-value
State 1: Recession state		
H_0: Research output does not Granger cause real GDP	10.0690***	0.0015
H_0: Real GDP does not Granger cause research output	2.0972	0.1476
State 1: Recovery state		
H_0: Research output does not Granger cause real GDP	3.0529*	0.0806
H_0: Real GDP does not Granger cause research output	0.5091	0.4755
State 1 & 2: Recession & recovery states jointly		
H_0: Research output does not Granger cause real GDP	18.7940***	0.0001
H_0: Real GDP does not Granger cause research output	4.0355	0.1330

Note: * (**) [***] denotes 10% (5%) [1%] level of significance.

Test results in Table 5 finds Granger causality from research output to GDP in all states. The null hypothesis of no Granger causality from research output to GDP is rejected at 1% level in the recession state and 10% in the expansion state, while join non causality in both states is rejected at 1% level. On the other hand, results show no Granger causality from GDP to research output. Therefore, the nonlinear VSTAR model with two states finds unidirectional causality from research output to GDP in the full sample.

5. Conclusion

The paper investigated the causal relationship between economic growth and research output for the US economy for the period 1981 to 2011. Employing a bivariate VAR, stationarity and cointegration were tested and indicated that economic growth and research output are integrated of order one and that no long run relationship exists between the two series. Full-sample Granger causality test found absence of a causal relationship between research output and economic growth. Stability of parameter estimates detected instability in the short-run (as well as the long-run) parameters which, led us to investigate time-varying (fixed-window i.e., rolling causality between output and research, as results from full-sample Granger causality tests cannot be relied upon.

The bootstrap rolling-window tests revealed show that during sub-periods 2003-2005 and 2009, GDP Granger caused research output while in 2010, the causality ran in the opposite direction. In general, our results support the line of thinking of Lee et al. (2011). The authors concluded that the relationship between GDP and research output are in general weaker in developed than in developing countries. Interestingly, we detect causality for the US, after or during periods of recessions, i.e., when the US economy was in a weak state or was recovering. Within a two state VSTAR regime-switching model we find unidirectional state dependent Granger causality from research output to GDP in the full sample.

The overall lack of a relationship between the academic research output and economic growth might also be attributed to the type of research conducted, the specific fields and whether the findings of important research are transferred either as knowledge or skills to the rest of the economy (Nelson and Romer, 1996). For that reason, it is considered that universities have the capacity to revitalise the relationship with growth by aligning their research with the industries current needs and also promote the transfer of knowledge to new graduates that will hopefully extend it beyond the existing limits.

To boost the research levels and their relevance to enhance the economy, Salter and Martin (2001) mention that scientists and institutions constantly argue that more funding is needed. However, for policy makers the benefits associated with government spending on infrastructure or education are more observable by the public and so they get preference. But, as shown by our findings, after recessionary periods, research output becomes an important factor until the economy stabilizes, and the market possibly start investing less than the optimum in basic research (Nelson in Pavitt, 1991).

References

Andrews, D.W.K. and Ploberger,W. 1994. Optimal tests when a nuisance parameter is present only under the alternative. Econometrica 62 (6), 1383-1414.

Andrews, D.W. K. 1993. Tests for parameter instability and structural change with unknown change point. Econometrica, 61, 821-856.

Balcilar, M., Gupta, R. and Miller, S.M. 2013. Housing and the great depression. Department of Economics, University of Pretoria, Working Paper No. 201308.

Balcilar, M., Ozdemir, Z.A. and Arslanturk, Y. 2010. Economic growth and energy consumption causal nexus viewed through a bootstrap rolling window. Energy Economics, 32, 1398-1410.

Barnett, A., Mumtaz, H., and Theodoridis, K. 2012. Forecasting UK GDP growth, inflation and interest rates under structural change: a comparison of models with time-varying parameters, Bank of England working papers 450, Bank of England.

De Gooije, G., and Vidiella-i-Anguera, A. 2004. Forecasting threshold cointegrated systems. International Journal of Forecasting, 20, 237-253.

De Moya-Anegon, F. and Herrero-Solana, V. 1999. Science in America Latina: A comparison of bibliometric and scientific-technical indicators, Scientometrics, 46, 299–320.

Dolado and Lutkepohl 1996. Making wald tests work for cointegrated VAR systems. Econometric reviews 15 (4), 369-389.

Efron,B. 1979. Bootstrap methods: another look at the jacknife. The annals of statistics 7(1), 1-26.

Fedderke, J. W. and Schirmer, S. 2006. The R&D performance of the South African manufacturing sector, 1970–1983. Economic Change, 39, 125–151.

Fedderke, J.W. 2005. Technology, human capital and growth. University of Cape Town, School of economics. Working paper No 27.

Granger, C.W.J. 1996. Can we improve the perceived quality of economic forecasts. Journal of Applied Econometrics 11(5), 455-473.

Groenewold, N. and Fraser, P. 1999. Time-varying estimates of CAPM betas, Mathematics and Computers in Simulation, 48, 531-539.

Hacker, R.S. and Hatemi-J,A. 2006. Test for causality between integrated variables using asymptotic and bootstrap distribution: theory and application. Applied Economics 38(3), 1489-1500.

Hansen, B.E. 1992. Testing for parameter stability in linear models. Journal of Policy modelling 14(4), 517-533.

Horowitz, J.L. 1994. Bootstrap-based critical values for the information matrix test. Journal of Econometrics 61(2), 395-411.

Inglesi-Lotz, R. and Pouris, A. 2011. Scientometric impact assessment of a research policy instrument: the case of rating researchers on scientific outputs in South Africa. Scientometrics, 88, 747–760.

Inglesi-Lotz, R. and Pouris, A. 2013. The influence of scientific research output of academics on economic growth in South Africa: an autoregressive distributed lag (ARDL) application. Scientometrics 95 (1), 129-139.

Inglesi-Lotz, R., Chang, T., Gupta, R., 2013. Causality between research output and economic growth in BRICS countries. Working paper 201337. Department of Economics, University of Pretoria, South Africa.

Johansen, S. 1991. Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models. *Econometrica* (59), 1551–1580.

Kealey, T. 1996. The economic laws of scientific research. New York: St. Martin's Press.

King, D. A. 2004. The scientific impact of nations. What different countries get for their research spending. Nature, 430, 311–316.

Lee, L.-C., Lin, P.-H., Chuang, Y.-W., & Lee, Y.-Y. 2011. Research output and economic productivity: a Granger causality test. Scientometrics, 89, 465–478.

Lucas, R. E. 1988. On the mechanics of economic development. Journal of Monetary Economics, 22, 3–42.

Lundbergh, S., and Terasvirta, T. 2002. Forecasting with smooth transition autoregressive models. In: Clements, M. P., and Hendry, D. F., (eds.). *A Companion to Economic Forecasting.* Oxford: Blackwell, 485–509.

Luukkonen, R., Saikkonen, P., and Terasvirta, T. 1988. Testing linearity against smooth transition autoregressive models. Biometrika, 75, 491–499.

Mantalos, P. 2000. A graphical investigation of the size and power of the Granger-causality tests in integrated-cointegrated VAR systems. Studies in Non-linear Dynamics and Econometrics, 4, 17–33.

Mantalos, P., Shukur, G. 1998. Size and power of the error correction model cointegration test. A bootstrap approach. Oxford Bulletin of Economics and Statistics, 60, 249–255.

Nelson, R.R. and Romer, P.M. 1996. Science, Economic growth and Public policy. Challenge 39(2), 9-21.

Pavitt, K. 1991. What makes basic research economically useful. Research Policy 20, 109-119.

Phillips, P.C. 1987. Time series regression with a unit root. Econometrica 55 (2), 277-301.

Phillips, P.C. and Perron, P. 1988. Testing for a unit root in time series regression. Biometrika, 75, 335–346.

Pouris, A., & Pouris, A. 2009. The state of science and technology in Africa (2000–2004): a scientometric assessment. International Journal of Scientometrics, 79, 297–309.

Price, D. S. 1978. Toward a model for science indicators. In Y. Elkana, G. J. Lederber, R. K. Merton, A. Thackray, & H. Zuckerman (Eds.), Toward a metric of science: the advent of science indicators. New York: Wiley.

Psaradakis, Z., Ravn M. O., and Sola, M. 2005. Markov switching causality and the money—output relationship. Journal of Applied Econometrics, 20, 665-683.

Romer, P. M. 1986. Increasing returns and long-run growth. Journal of Political Economy, 94, 1002–1037.

Rothman, P., van Dijk, D., and Hans, P. 2001. Multivariate STAR analysis of money-output relationship. Macroeconomic Dynamics, 5, 506-532.

Salter, A.J. and Martin, B.R. 2001. The economic benefits of publicly funded basic research: a critical review. Research Policy 30, 509-532.

Schumpeter, J.A. 2000. Entrepreneurship as innovation. Entrepreneurship: The social science view 51-75.

Shelton, R.D. & Leydesdorff, L. 2011. Bibliometric Evidence for Funding Trade-Offs in National Funding Strategies, 13th International Conference on Scientometrics and Informetrics, Durban, July. 2011. Available at http://itri2.org/s/

Shukur, G. and Mantalos , P. 1997a. Size and power of the RESET test as applied to systems of equations: A boot strap approach. Working paper 1997:3, Department of Statistics, University of Lund, Sweden.

Shukur, G. and Mantalos , P. 1997b. Tests for Granger causality in integrated-cointegrated VAR systems. Working paper 1998:1, Department of Statistics, University of Lund, Sweden.

Shukur, G. and Mantalos, P. 2000. A simple investigation of the Granger-causality test in integrated-cointegrated VAR systems. *Journal of Applied Statistics*, 27, 1021–1031.

Stock, J.H. and Watson, M.W. 1996. Evidence on Structural Instability in Macroeconomic Time Series Relations, Journal of Business & Economic Statistics, 14, 11-30.

Tamura, R.F. 1991. Fertility, human capital and the ‘‘Wealth of Nations’’. University of Iowa, Department of Economics. Working paper No 91-17.

Toda, H.Y. and Yamamoto, T. 1995. Statistical inference in vector autoregressions with possibly integrated processes. *Journal of Econometrics*, 66, 225–250.

Tsay, R. S. 1998. Testing and modeling multivariate threshold models. Journal of the American Statistical Association, 93, 1188-1202.

Terasvirta, T. 1998. Modelling economic relationships with smooth transition regressions. In: Ullah, A., and Giles, D. E. A., (eds.). *Handbook of Applied Economic Statistics*. New York: Marcel Dekker, 507-552.

Van Dijk, D., Terasvirta, T., and Franses, P. H. 2002. Smooth transition autoregressive models - A survey of recent developments. Econometric Reviews, 21, 1-47.

Vinkler, P. 2008. Correlation between the structure of scientific research, scientometric indicators and GDP in EU and non-EU countries. Scientometrics, 74, 237–254.

Appendix

Figure A2: Granger causality test *p*-values: Rolling-window estimates