

# Review of: "Government expenditure and economic growth: evidence from the critical sectors in an emerging economy"

Weshah Razzak<sup>1</sup>

<sup>1</sup> Massey University

**Potential competing interests:** No potential competing interests to declare.

In general, this article empirically tests the effects of government expenditures (education, health, agriculture, pensions and gratuities and public debt servicing) on economic growth in Nigeria. The question is a reasonable research question. My initial thought is that G expenditures might have a short-run effect on output but nothing in the long run. However, some sustained G spending on infrastructure such as roads, power, internet and other ICT capital etc. might be argued to have a long-run sustainable effect on output growth.

## The Main Issue

To answer this question, the authors need a formal growth model in order to show how these G expenditures affect growth. In the long-run, growth could be pinned down by population growth. The most interesting bit is the transitional dynamic. In the Solow model (and some updated versions of it such as Mankiw, Romer, and Weil, <https://academic.oup.com/qje/article-abstract/107/2/407/1838296> and Jones <https://www.jstor.org/stable/2118448> for example) Output growth is typically a function of labor (and the quality of labor perhaps), physical capital, human capital (basically average years of schooling), R&D investments, FDI, and research efforts (Razzak, 2022, <https://link.springer.com/article/10.1007/s13132-022-00983-2> ). In the case of Nigeria, I would imagine that oil production is a big part of GDP. All these variables are omitted from the model in this paper and the authors should account for them. These omitted variables lead to a bias and inconsistent parameters estimates, a serious specification problem.

Imagine hypothetically that Nigeria has no roads and no electricity. Measure GDP growth. Now, let the government build roads and provide electricity. Measure growth again; we should find some positive impact on long run GDP growth. It is much more difficult to explain the effects of G expenditures on education, health, etc. on GDP growth because the effects are indirect. There is a literature on the lack of effects of such G spending on the outcomes of education and health. To test these effects you need to show how theoretically first. It is rather G investments (public investments) that are directly affecting output that has positive effects. There is evidence elsewhere where G investments crowds out private investments.

## Secondary Issues

1. VECM requires some theoretical restrictions imposed in order to interpret the impulse response functions (not shown in the paper). Otherwise, the IRF has no causal meaning. I would not use VECM. Perhaps I do OLS, DOLS, FM-OLS, 2SLS (IV estimator), and ML estimator, and analyze all the results.

2. I would always begin with plotting the data. It is more informative to see the data before doing the analysis.
3. For robustness, use all available unit root tests. All commonly used tests are weak against stationary alternatives. They fail to reject the unit root more often. They may not be able to tell the difference between a root of 1, which is how these tests are proved theoretically, and say 0.97. Use more tests, use different specifications: without a constant and trend, with a constant and trend, with a constant only. Without seeing the data I could not tell if there is a break in the data. These unit root tests fit a linear line through the data. They could confuse the break, if any, with unit root. Test for unit root with breaks.
4. If all the data are  $I(1)$  and cointegrated, I would estimate the model in log-levels rather than log-difference because the latter is more about the short-run dynamic. FM-OLS and DOLS are much better estimators in this particular case.
5. The long-run estimates reported in tables (8 and 9) - three only are significant - do not mean much as I explained earlier. We could only take the impulse response functions to convey something, and not even meaningful if the VECM is unrestricted. Impulse response functions should have standard errors around them. If not, the variance decomposition might tell us something.
6. All data must be real **not** nominal.

#### Minor Issues

1. The ADF test's null hypothesis unit root, therefore, it tests for unit root not for stationarity. I know of only one test that tests for stationarity, the KPSS.
2. The ADF paper is not cited, <https://www.jstor.org/stable/2336570>