



Peer-Review Record:

Economic Development and Government Spending: An Exploration of Wagner's Hypothesis during Fifty Years of Growth in East Asia

Hassan Mohammadi and Rati Ram

Economies **2015**, *3*, 150-160, doi:10.3390/economies3040150

Reviewer 1: Anonymous

Reviewer 2: Anonymous

Editor: Ralf Fendel (WHU-Otto Beisheim School of Management, Germany, *Editor-in-Chief* of *Economies*)

Received: 9 June 2015

First Editorial Decision: 26 August 2015

First Revision Received: 10 September 2015

Second Editorial Decision: 15 September 2015

Second Revision Received: 19 September 2015

Accepted: 29 September 2015

Published: 9 October 2015

First Round of Evaluation

Round 1: Reviewer 1 Report

The motivation for testing Wagner's law using data from East Asian countries is fine; the rapid growth should provide some strong evidence.

However, the overall thrust in the paper is somewhat strange. After mechanically, without checking assumptions, doing specification test, checking the results carefully, etc. they conclude that relationships, which are not visible in graphs, cannot be found even with the most sophisticated approach. But who thinks they could, or should, be found? Tests are used to support informal results, such as those obtained in graphs.

I don't really doubt the empirical findings of the paper (which I see as support for the hypothesis in Japan and possibly in Korea), but I dislike the approach of spending a lot of space on various, often pointless, tests. There should be one detailed analysis based on adequate tests, and a discussion about the results. After all, government expenditure is a political decision.

An import question is why are the series for real GDP per capita are taken from PWT 7.0? These series are constructed to be used in cross-country studies, and they are not suitable for the analyses in the paper. See the paper by Johnson *et al.* (2009). This is a major flaw of the paper and the choice should be defended by showing that using the ‘standard’ real GDP series’ would give the same results. Or the GDP series, downloaded from World Development Indicators (local currency, constant prices), should be used.

The dependency ratio has not been included mechanically in most papers, as it is this one. It is supposed to pick up changes in demand for government expenditure, such as schooling and pensions. The general idea is that these have increased government expenditure, both because per-capita schooling and pensions have expanded and because the dependency ratio has increased. In most of the countries studied, as is well-known, the dependency ratio declined during the period studied, while expenditure per person probably increased. Japan is most likely an exception; it should have entered the demographic transition much earlier than the other countries. There is thus no a priori no reason to expect the dependency ratio to play an important role in the Wagner hypothesis in general. It can thus be removed from the analysis, or moved to a footnote. If the author(s) wish to pursue the approach of testing other potentially relevant variables, see Shelton 2007.

Figure 1 tends to hide co-movements between the variables by mixing of scales. I suggest the graphs are re-done using two y-axes and preferably logs.

The correlations in Table 4 are by definition not valid in the cases where the variables are non-stationary. Even high correlations are uninformative in this case. Moreover, most studies that add variables to the models, such as the dependency ratio, argue that Wagner’s law holds conditional on it. This makes the unconditional correlations uninteresting.

The static regressions in Table 5 are used when testing for cointegration (Engle-Granger approach), when the variables have a unit root. In these cases, correction for serial correlation (AR1) just messes up the results. And when the variables are stationary, the validity of AR1 should be tested. Usually it imposes invalid restrictions. The autoregressive distributed lags model is a much better choice in general. But later tests show that the variables are non-stationary.

In Table 5 we can get information about cointegration from the R2 and the *t*-values. Only Japan shows any promise, in all other cases R2 are very low. I would say they are suspiciously low for time series data. Are thy correct?

The discussion about endogeneity can be skipped. In any case, cointegration analysis allows for endogeneity, and it can easily be test when there is cointegration. The instruments used (lags) are not useful since the variables are very persistent. And they should give the same results as OLS.

Unless short run dynamics are an important reason for the lack of cointegration, the Johansen test should give the same results as the static OLS regressions, when there is one cointegrating vector. The authors do not seem to understand they are using different approaches to test for cointegration.

The Johansen approach is suitable for times series with many observations, about 300. When one has fewer than 50, it is not possible to rely on the tests in the way it is done in the paper. Additional information has to be provided. See Juselius (2001) for an instructive example.

Gregory–Hansen is used when one fails to find cointegration with standard tests. There is no point in testing countries such as Japan and Korea, when a superior test shows that they are cointegrated.

The use of the MTAR model is dubious because of the small sample, and it should not be used when cointegration has been established. It is unlikely that there is enough information in the data to capture

asymmetric effects and the standard procedure is to use MTAR when Engle-Granger fails to show cointegration.

What is the alternative hypothesis of the Pedroni test, that there is cointegration in any country or in all countries? This should be made clear.

The authors provide a very provocative conclusion (see below), but it is not supported by the analyses in the paper. The single-country analysis would provide the same results if done carefully, and the panel data analyses are dubious both because of the small cross-section (six) and the heterogeneity of the sample. I don't mind the use panel analyses they can be presented for illustrative purposes, noting the heterogeneity. However, the time series analyses should be done correctly. Moreover, it is quite obvious that in simple bivariate cases, graphs can be very informative. However, the author(s) do not manage to provide any useful information about the dependency ratio in the 'old fashioned' analyses; they do only in Table 5 when (unknowingly) testing for cointegrating using OLS.

“Methodologically, while the pattern is reflected fairly well in plots and simple correlations, and even in regressions, the widely-used cointegration methodology yields a diverse scenario in four different procedures, and, through a choice of the procedure and the model, it seems as easy to conclude that there is no cointegration, and thus no support for the hypothesis as to conclude that the evidence supports the hypothesis in a majority of the cases. At any rate, it is not obvious that such sophisticated and complex tests of cointegration provide any useful additional insight regarding the empirical status of the hypothesis. More generally, we venture the highly subjective view that despite the immense mathematical sophistication of cointegration tests and their application in perhaps thousands of studies, it is not obvious how much additional substantive insight has been gained in the wide variety of contexts in which these tests have been used.”

These statements should be explained. Is the second one a rejection of the last 40 years of advances in times series analysis?

Despite Biehl's (1998) insightful essay, almost all empirical research on the topic has interpreted the Wagnerian proposition

As Ram (1998, Pp. 149-150) has explained, considerations of “nonstationarity” of the variables or “spuriousness” of the correlations or regression do not have much bearing on such tests.

References:

Cameron A. Shelton (2007) “The size and composition of government expenditure” *Journal of Public Economics*, 91 (11–12)

Simon Johnson, William Larson, Chris Papageorgiou, Arvind Subramanian. (2009) “Is Newer Better? Penn World Table Revisions and Their Impact on Growth Estimates” NBER Working Paper No. 15455

Juselius, K. (2001). Unit roots and the demand for cigarettes in Turkey: pitfalls and possibilities. University of Copenhagen Dept. of Economics Discussion Paper, (01-02).

Round 1: Author Response to Reviewer 1

1. Following the reviewer's comments, we have now focused on one detailed analysis based on adequate tests and a discussion of the results. The revised version drops Descriptive statistics in old Table 2, rates of increase given in the old Table 3, correlations reported in old Table 4, the regressions in old Table 5, and pooled regressions of old Table 6. The discussion based on these Tables has consequently been omitted from the revision. The unit-root-test Tables have been moved from the Appendix to the text as Table 2, and in view of the evidence in favor of unit roots, rates of growth of

GDP and government variables have been shown in Table 3 somewhat like the old Table 3. These are the most major changes.

2. The authors are aware of the Johnson et al. paper (JME 2013) and other critiques of the PWT data. However, while the LCU numbers are appropriate for single-country analysis, these are not suitable for panel-data format which is a significant part of this study. Therefore, we use the PWT numbers throughout and hope that single-country analysis with these numbers is a reasonable approximation to the scenario indicated by LCU data.

3. The dependency variable has been removed from the analysis.

4. Figure 1 has been redone with separate scales for logarithms of real GDP per capita and government share.

5. The correlations of old Table 4, the static regressions of Table 5, and the discussion of endogeneity have been omitted.

6. Since our data cover a long period of nearly 50 years, we are inclined to rely on Johansen tests although the number of data points is smaller than 100. The Gregory-Hansen test is done for Japan and Korea also just for completeness, and might be of some methodological interest since it does not show cointegration even for Japan and Korea. The reviewer's observation has been acknowledged in note 2.

7. We agree with the reviewer about the limited usefulness of MTAR in small samples and for Japan and Korea. However, it is included in the hope it would indicate whether lack of cointegration for Malaysia, Philippines, Singapore and Thailand might be due to asymmetric adjustment. The reviewer's observation has been acknowledged in note 3.

8. We have highlighted the alternative hypothesis in Pedroni's tests on Page 8. As Pedroni (1999) clearly points out, the alternative hypothesis rejects the null of no-cointegration across all cross-sections in both within- and between-dimension tests.

9. The "provocative conclusion" reflected one author's subjective perception and was stated as such. He shares the perspective noted by Bennett McCallum's in his "Is the spurious regression problem spurious?" (Economics Letters, 2010). However, the segment cited by the reviewer has been deleted since it is a more general proposition and is not necessary for the limited analysis reported in our paper.

10. We completely share the reviewer's view that despite Biehl's PF/FP essay, almost everyone follows the standard approach, and we do the same. Our observation is just a restatement of the proposition that our approach is traditional.

Round 1: Reviewer 2 Report

The paper analyzes the validity of Wagner's law in six East Asian countries during the period 1960-2008. While interesting, generally well written and scientifically sound, several suggestions should be considered before further consideration. The suggestions are listed below in no particular order of importance:

1. List of references is scant and limited. It is necessary to add more recent works on the subject with alternative ways of testing for Wagner's Law. Suggestions include "Testing the Validity of Wagner's Law in Bolivia: A Cointegration and Causality Analysis with Disaggregated Data" by A. Bojanic (2013); "Wagner's Law and Italian Disaggregated Public Spending: Some Empirical Evidences" by C. Magazzino (2010); and "Panel Data, Cointegration, Causality and Wagner's Law: Empirical Evidence from Chinese Provinces" by Narayan, P.K., *et al.* (2008).
2. The paper only lists one way of testing for Wagner's hypothesis (*i.e.*, analyzing the relation

between the share on GDP of government spending and real GDP per capita) when in fact there are many ways of looking at the hypothesis. The authors should include at least a couple of recent ways of testing this law, in addition to the traditional way of testing for it.

3. The variables used do not have the same units, which may clog the analysis of the paper. Specifically, real GDP per capita is measured in 2005 dollars while the share of government consumption on GDP is measured in current prices. Either both should be in real terms or both in current prices, but not as is currently the case.
4. The descriptive stats in Table 2 reflect stats for the whole period? If so, the table should explicitly state this fact.
5. Why are there 3 graphs for the Philippines in Figure 1? By the way, if the variables were measured in either current prices or utilizing the same base year, these graphs would likely be quite different (same goes for all statistical analysis, it is likely to vary when the two variables are measured the same way)
6. What is the point of the dependency rate? The authors should expand on the need to add this variable and what exactly is it trying to capture. In the same vein, what do the authors mean by “however, it seems that the fall in young-dependency ratio was greater than the rise in the old-dependency rate” on p. 4?

In Table 4, delete the “generated by SAS” note.

Round 1: Author Response to Reviewer 2

1. Several new references, including those kindly mentioned by the reviewer, have been added, and there is a concise description of these in note 1.

2. Several formats in which Wagner’s hypothesis has been formulated and tested have been listed on page 5 of the revision. That paragraph includes six different approaches.

3. Real GDP per capita is a fairly standard proxy for the “level of development” in the context of Wagner’s hypothesis. It might not be appropriate to use nominal GDP per capita. For the government share, it seems better to take the current-price ratio which indicates what fraction of the current output is used for government activities. The deflators for GDP and government spending are different and “real” government-share may not be a good indicator of the resource-use by the government during the current year.

4. The old Table 2 has been deleted. Those descriptive statistics were for the entire period. The growth rates in the new Table 3 are also for the entire period, and that has been indicated.

5. The redundant graphs for Philippines have been deleted.

6. The dependency-rate variable has been dropped.

7. The old Table 4 has been deleted along with the remark about p-values being taken from SAS.

Second Round of Evaluation

Round 2: Reviewer 1 Report

The paper is much better and I only have some minor comments.

First, the authors should provide some tentative explanation for their findings in the conclusion section based politics, the characteristics of the economies, *etc.* Now the paper is mainly an econometric

exercise. Even if many studies have failed to that Wagner's law holds, there is a common perception that it holds over periods when countries become developed.

The graphs in Figure 1 are still misleading, probably because of the scaling of the axes. It is obvious for Japan, where there is no discernible evidence for cointegration. With proper scaling, the lines should cross each other four times, providing some (but weak) evidence for cointegration.

The x -axes in Figure 1 should have years.

Round 2: Author Response to Reviewer 1

Please note the following changes in revised version: (a) the plots in Figure 1 have been rescaled and now there are cross-overs in the plot for Japan, (b) years have been shown on the horizontal axis in Figure 1, (c) a sizable addition has been made toward the end of Section 3 to address the reviewer's observations about possible reasons for lack of cointegration in most countries, and (d) a few minor editorial alterations have been made.

Round 2: Reviewer 2 Report

My recommendation is to accept the paper for publication.

Round 2: Author Response to Reviewer 2

Thank you.

© 2015 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).