UNR Economics Working Paper Series
Working Paper No. 07-002

The Effect of Federal Government Size on Long-Term Economic Growth in the United States, 1792-2004

Federico Guerrero and Elliott Parker

Department of Economics /030
University of Nevada, Reno
Reno, NV 89557-0207
(775) 784-6850│ Fax (775) 784-4728
email: guerrero@unr.edu; elliottp@unr.edu

August, 2007

Abstract

In this paper, we consider whether there is statistical evidence for a causal relationship between federal government expenditures and growth in real per-capita GDP in the United States, using available data going back to 1792. After studying the time-series properties of these variables for stationarity and cointegration, we investigate Granger causality in detail in the context of a Vector Error Correction Model. While we find causal evidence supporting Wagner's Law, we find no evidence supporting the common assertion that a larger government sector leads to slower economic growth.

JEL Classification: G32, G15, D92, E65, F39.

Keywords: long-term economic growth, federal government size, Wagner's Law, United States, cointegration, Granger causality, vector autoregression, vector error correction model.

The Effect of Federal Government Size on Long-Term Economic Growth in the United States, 1792-2004

Federico Guerrero *
and
Elliott Parker

Assistant Professor and Professor, respectively
Department of Economics
University of Nevada, Reno

July 24, 2007

Abstract: In this paper, we consider whether there is statistical evidence for a causal relationship between federal government expenditures and growth in real per-capita GDP in the United States, using available data going back to 1792. After studying the time-series properties of these variables for stationarity and cointegration, we investigate Granger causality in detail in the context of a Vector Error Correction Model. While we find causal evidence supporting Wagner's Law, we find no evidence supporting the common assertion that a larger government sector leads to slower economic growth.

JEL Classification: G32, G15, D92, E65, F39.

Keywords: long-term economic growth, federal government size, Wagner's Law, United States, cointegration, Granger causality, vector autoregression, vector error correction model.

Contact Information:
Department of Economics /0030
University of Nevada, Reno
Reno, NV 89557-0207 USA
guerrero@unr.edu, eparker@unr.edu

* Corresponding Author

The Effect of Federal Government Size on Long-Term Economic Growth in the United States, 1792-2004

Abstract: In this paper, we consider whether there is statistical evidence for a causal relationship between federal government expenditures and growth in real per-capita GDP in the United States, using available data going back to 1792. After studying the time-series properties of these variables for stationarity and cointegration, we investigate Granger causality in detail in the context of a Vector Error Correction Model. While we find causal evidence supporting Wagner's Law, we find no evidence supporting the common assertion that a larger government sector leads to slower economic growth.

1. Introduction

There are few ideological debates in economics more fundamental than whether or not a larger government is bad for growth, even aside from its more normative effects on economic freedom. In the United States, the Club for Growth, for example, supports candidates who endorse reducing the size and economic role of government, on the assertion that these policies favor faster economic growth. In both theory and in the data, however, the effects are more mixed.

In addition to the short-run fiscal effects of government expenditures on real aggregate demand in an economy with excess capacity, Poot (2000) argued that there are at least seven separate effects of government spending on growth. These include, among other factors, the provision of pure and quasi public goods, the distortionary effect of taxes on resource allocation, and the comparative inefficiency of government control over resources and production, relative to the private sector it replaces. All things considered, Barro (1990) has made a persuasive case that the aggregate relationship between the size of government and economic growth may be shaped like an inverted-U, with low growth resulting from both too little and too much government.

On the empirical front, Landau (1983, 1986) found a negative effect of government consumption on growth, while Ram (1986) found a positive effect, and the international evidence uncovered over the last two decades has remained decidedly mixed. In his survey, Slemrod (1995) argues that the aggregate effect of government involvement is negligible, though some types of taxes affect some behaviors significantly. Engen and Skinner (1996) focus on the effect of taxes, and they find mildly negative effects for some taxes and positive effects for others, but like much of the rest of the literature, the effects of larger government are contradictory, ambiguous, and in the aggregate rather minimal. Plümper and Martin (2003) found a negative effect of government on growth primarily in non-democratic countries, a result generally consistent with the findings of Guseh (1997) and Scully (2001). In a recent survey, Poot (2000) cited 41 studies, with seven finding a positive effect, twelve finding a negative effect, and 23 inconclusive. Even more recently, Lee and Lin (2007) found a negative effect of government that became insignificant once demographic factors were taken into account.

The bivariate causal relationship between government size and economic growth is complicated by the potential for reverse causation, since GDP growth can also lead to increased government spending, an effect that is frequently referred to, in a broad sense, as "Wagner's law." (Wagner, 1893). As Baumol and Bowen (1965) first noted for the example of orchestras, many labor-intensive services have inherently lower rates of productivity improvement, while wages are driven by productivity improvements elsewhere. As a result, a growing economy will tend to have a rising share of income spent on services, many of which are likely to be provided by government, and this leads to a so-called "cost disease" of rising government share. Wagner's law of increasing state spending, the thesis that government's role increases as a country industrializes, has also been interpreted to imply an income elasticity for government expenditures greater than one, although Peacock and Scott (2000) argue that this relationship is usually misspecified. As with the previous causal relationship, there is also a substantial literature testing whether economic growth leads to a larger government. Recent examples include Jackson, et al. (1998), Demirbas (1999), Islam (2001), and Halicioglu (2003).

A common methodological pitfall in the literature that tries to uncover the causal links between government size and long-term economic growth is that they regularly conduct Granger causality tests outside the cointegration framework.[1] As is well-known, this problem may render many of their conclusions invalid (Granger, 1988). Furthermore, the papers that do place their Granger causality analyses within the cointegration framework do not tend to implement a Vector Error Correction (VEC) model, which is the natural follow up in the case in which the variables are cointegrated, given that the definitive test of causality lies with the error correction term (Engle and Granger, 1987; Granger, 1988).[2]

In this paper, we begin in section two with a brief description of government size and economic growth over the long-term for the case of the US, using data going back to 1792. We then estimate an unrestricted vector autoregression (VAR) model and, after checking for the model's stability, we use impulse-response analysis as a way to get first estimates of the inter-relations between government size and economic growth for the case of the United States. In section four, we study the time series properties of the data, testing for stationarity and cointegration; we also perform traditional Granger causality tests outside the cointegration framework as a way to see what our data say when traditional methods are applied and as a way to compare our results with previous contributions. In the fifth section we exploit the results from

[1] Ghali (1998) and Islam (2001) are among the exceptions.
[2] The only exception we could find is Ghali (1998), who implements VEC models in a setup with multiple cross-sections (ten advanced countries), but short time spans (quarterly data for the period 1970-94). Islam (2001) implements Johansen-Juselius' weak exogeneity tests, but reports no VEC model.

cointegration analysis and implement a VEC model that sheds clear light on the issue of causation. We conclude with a brief summary and suggestions for further research.

2. Government Size and Economic Growth in the United States

There are many ways to measure the size and scope of government intervention, but the most common metric is the relative amount of government expenditures, including both government purchases and transfers. Federal government expenditures (*G*) as a share of GDP (*Y*) remained low prior to the Great Depression, with exception to wartime: *G/Y* rose to an annual high of five percent of GDP during the War of 1812, 19.5 percent during the Civil War, and 30 percent during World War I. The *G/Y* ratio dramatically increased during FDR's Administration at the federal level. After the Great Depression, *G/Y* peaked at 47.5 percent during World War II, but afterwards it generally ranged from 20 to 25 percent. State and local expenditures, meanwhile, actually fell during the Roosevelt Administration, from over seven percent during the Hoover Administration to under five percent by 1945.

From the Truman Administration through the Reagan Administration, government spending continued to rise, to an average of 22.5 percent of GDP for federal spending in the first half of the 1980s and an average of almost eleven percent of GDP for state and local expenditures. By the first half of the 1990s, total government expenditures in the United States summed to one third of GDP, though this fell somewhat during the Clinton Administration.

How did this secular increase in government size affect growth on the margin in the United States? Gwartney, Lawson, and Holcombe (1998) argued that the size of government in the first half of the 1960s was close to the optimal level, and estimated that U.S. incomes would have been 20 percent higher in 1996 had government stayed the same relative size. But they chose to make their case using the period 1960-65 as their growth baseline, a half-decade with the highest growth rates in postwar American history and, until the Kennedy tax cuts went into effect in 1964, a period characterized by a maximum marginal tax rate of 91 percent on unearned income. They then compared this to the period 1990-95, a half-decade which began with a recession, to argue that growth in government reduced economic growth. Yet even as they wrote their paper, the economy was beginning to once again grow at a pace not seen since, well, the early 1960s.

To more carefully analyze the effect of government spending on growth, we compile annual GDP data provided by Johnston and Williamson (2006), and federal expenditure data provided by Garrett and Rhine (2006). Both sets of data contain annual data spanning the period 1792-2004, thus providing a unique opportunity to investigate the long-run relationship between government size and economic growth. Like Jones and Joulfain (1991), who investigate the relationship between expenditures and revenues prior to the Civil War, we focus on the federal budget and exclude state and local government expenditures from our analysis because of data availability.

Some summary statistics for our dataset are provided in Table 1. Average annual growth rate for real GDP generally declined over time, seemingly lending credence to the argument that the rising size of government after FDR slowed economic growth. But population growth rates were significantly higher before the Great Depression, and the average real growth rate *per capita* was actually faster after 1950, when the government was significantly larger. Furthermore, the standard deviation of growth was smaller, suggesting that the growth of government was not only correlated with faster growth, but also correlated with more economic stability. While *G/Y* was much higher after the Roosevelt Administration than before, it also became much more stable. For the periods of comparison used by Gwartney, et al. (1998), for example, we find that *G/Y* was only slightly higher in the latter period, 22.4 percent of GDP in 1991-1995 versus 21.7 percent in 1961-1965.

Is the United States on the downward-sloping side of Barro's inverted U? In other words, is the marginal effect of government bad for economic growth in the U.S.? For the economies of Western Europe, many economists would argue that the marginal effect of government on growth is negative, but the evidence is weaker that this is also true for the United States.[3] According to Gwartney, et al. (1998), the United States in the mid-1990s had the smallest total government expenditure share of all the OECD countries, and also the smallest growth in government's share during the postwar period. If any developed market economy is on the lower, upward-sloping portion of curve, it seems to follow that it should be the United States.

Islam (2001) recently investigated the relationship between total government expenditures and growth in order to test for Wagner's law, using U.S. data from the U.S. Department of Commerce spanning the period 1929-1996, and found a long-term equilibrium relationship between economic growth and the share of government expenditures, with causality that supported Wagner's Law but not the reverse. Like Islam, we estimate whether or not such a cointegrating relationship exists for the U.S., though we consider a much longer time-series and we investigate both causality issues and the dynamic relationship in more detail. Others have used similar methods to investigate this question for other countries, or in other cases. One of the earlier examples is that of Conte and Darrat (1987), who use Granger causality tests to reject the hypothesis that public sector expansion in postwar OECD economies negatively influenced real GDP growth rates. Ghali (1998) used a multivariate cointegration approach with several

[3] Lindert (2004) provides a compelling case defying the view that economies with larger welfare states grow more slowly.

variables, including GDP and government spending among others, for a sample of OECD countries, while Jones and Joulfain (1991) used cointegration tests and error correction models to find evidence of short-term and long-term causality relationships in the United States between federal revenues and expenditures.

3. A First Approach to the Evidence: An Unrestricted VAR Impulse-Response Analysis

As a first approach to analyzing the effect of *G/Y* on growth in the United States, we estimate an unrestricted VAR model from which we extract impulse-response results. An advantage of using an unrestricted VAR model is that it implies a non-committal approach to the data in which issues of causation, timing and appropriate structural restrictions are temporarily left on hold, awaiting further analysis.

3.1. An Unrestricted VAR model:

We start by defining y as the natural logarithm of U.S. real per-capita GDP, and $g = \ln(G/Y)$, i.e. as the natural logarithm of the ratio of federal government expenditures, including transfers, to real GDP. These data are shown in Figure 1, which strongly suggests that neither variable is stationary in their levels. As is well-known, running regressions involving I(1) variables may give rise to spurious results and multiple inference and interpretations problems, given that the F-statistic does not follow the tabulated values of Fisher's F distribution (Granger & Newbold, 1974). In Figures 2(a) and 2(b), we show the first differences of these variables, and these appear to stationary. A common specification in the literature regresses rate of growth of real GDP, either per-capita or not, on *G/Y* and other variables, but this most likely leads to misspecification, since regressing a stationary dependent variable on independent variables which are not stationary is an approach that Granger (1986: 216) argues "makes no sense as the independent and dependent variables have such vastly different temporal properties." Indeed, the expected coefficient(s) from such a regression would be zero in such a case. Therefore, instead of the commonly-used model $\Delta Y/Y = f(G/Y)$, our estimable VAR model uses both variables in logarithmic first differences and is of the following form:

$$\Delta g_t = c(1) + \sum_{j=1}^{4} a_j \Delta g_{t-j} + \sum_{j=1}^{4} b_j \Delta y_{t-j} \tag{1}$$

$$\Delta y_t = c(2) + \sum_{j=1}^{4} d_j \Delta y_{t-j} + \sum_{j=1}^{4} e_j \Delta g_{t-j} \tag{2}$$

Since only lagged values of the endogenous variables appear on the right-hand side of the equations, simultaneity is not an issue and equation-by-equation OLS yields consistent estimates. Moreover, even though the error terms may be contemporaneously correlated, OLS is efficient and equivalent to GLS since all equations have identical regressors.

To determine the optimal number of lags in the VAR model, which is important to make sure that the residuals are uncorrelated and homoskedastic across time periods, we try a battery of selection criteria, with each test performed at the five percent significance level. These include a sequentially modified Likelihood Ratio (LR) test, a Final Prediction Error (FPE) test, an Akaike Information Criterion (AIC) test, the Hannan-Quinn (HQ) Information Criterion test, and the Schwartz Information Criterion (SIC) test. As is shown in Table 2, the first three indicators suggest the use of four lags, whereas the last two criteria suggest the use of one lag. Given that a Wald test on parameter restrictions rejects the null hypothesis that there are no lags of any order higher than unity, we opted for a four-lag specification in line with the first three criteria. This selection will be further justified by Granger causality tests in section 4.2. below.

Table 3 contains the estimates of our VAR(4) model. The first column reports the estimate of the growth in the size of the government equation, and the second column reports the estimate of the economic growth equation. As usual with macroeconomic series, the autoregressive components are important statistical determinants of both series in both columns. Both the first and the second lags of economic growth per capita are statistically significant explanatory factors for the size of the federal government, while none of the lags of government size growth are statistically significant in explaining economic growth per capita in the second equation.

In order for the VAR to be stationary, all the inverse roots of the characteristic AR polynomial must lie inside the unit circle. If this is not the case, impulse-response inferences are not valid. In this case, the modulus values are 0.80, 0.64, 0.43, and 0.16, and so the VAR is stationary and we can proceed to the impulse-response analysis.

3.2. Impulse Response Analysis:

To clarify the overall effects of innovations to both endogenous variables over a long time horizon, we report accumulated impulse-response graphs over a ten-year window in Figures 3(a) and Figure 3(b). Economic growth appears to have a statistically significant accumulated effect on the growth in the size of the government in year 4, a finding that lends support to Wagner's law, while the size of the federal government has statistically-insignificant effects on economic growth at all lags.

The orthogonalization of impulses to identify shocks is known to be sensitive to the order in which the variables are declared. Figures 3(a) and 3(b) show the case in which g is entered first (i.e., assumed to be less endogenous), but the results are almost identical if we reverse the order.

3.3. Block ExogeneityTests:

Table 4 contains standard block exogeneity tests (pairwise Granger causality tests) for government and economic growth per capita. Results are unambiguous: the lags of economic growth display block exogeneity, but the lags of government size growth do not. The number of lags included in the tests displayed in Table 4 ranges from 4 to 10. No block exogeneity can be identified if less than 4 lags are included in the tests, suggesting either mis-specification due to insufficient lags or a lack of cointegration, an issue that will be tackled in the next section. Including a larger set of lags (we tried with up to 20 lags) does not change results: the lags of economic growth display block exogeneity but the lags of government size growth do not. This result is later confirmed when causality is tackled by means of a VEC model.

This initial approach undertaken so far is not free of problems. Problems with the unrestricted VAR model approach include inefficient estimation due to over-parameterization (see Zellner, 1988, among others) and mis-specification when the data are first-differenced and the variables are cointegrated (Engle and Granger, 1987), which is potentially the case here. Hence, more analysis is needed.

4. Stationarity and Cointegration

A more precise and rigorous statistical analysis of the inter-relations between government and economic growth requires a careful study of the time series properties of the series at hand. We proceed in three steps. We consider the issue of stationarity (or lack thereof) first. After having studied the stationarity properties of the series, we start our search for valid restrictions to our VAR(4) model. Our search starts in the last part of this section with cointegration analysis. Cointegration analysis is important because provided both variables (y and g) are non-stationary, if a long-term equilibrium relationship between government and economic growth exists, the data can be modeled by means of a Vector Error Correction (VEC) model, which helps improve the econometric fit by tying the short-run dynamics to the long-run relationship between both variables. On the other hand, if the government and economic growth series turn out to not be cointegrated, then VEC restrictions would not be appropriate (Engle and Granger, 1987).

4.1. Stationarity Issues:

We searched for unit roots using regressions that included a constant, one lag of the variable in levels and one lag of the variable in first differences as augmentation terms. Alternatively, a linear trend was also added to the specifications. All tests provide a uniform message: both series have unit roots in their levels (i.e., they are I(1) processes) and both series are I(0) processes in their first differences.[4] This confirms the intuition provided by Figures 1 and 2, and confirms that it would have been inappropriate to postulate the unrestricted VAR model in the levels of the variables y and g. Table 5 provides the tests without a linear trend (i.e., including only a constant and the augmentation terms).[5]

[4] For a more detailed discussion on unit roots in US GDP, the reader is referred to Murray and Nelson (1998) and the literature that follows them.

[5] Like all other results in this paper which are not reported in detail, results which include a linear trend are available on request.

4.2. Cointegration Analysis:

If these series are cointegrated, causality tests conducted outside the cointegration analysis framework may lead to incorrect causal inferences, since the error correction term is omitted in the specifications used to test for Granger causality (see Granger, 1988 for additional discussion). Hence, to account for that difficulty we follow a two step procedure in what follows. First, we check if the series y and g, which have been shown to be I(1), are cointegrated. Second, if the hypothesis that the series are not cointegrated is rejected, we implement a VEC model in order to double check the results on block exogeneity reported in section 3.3. If cointegration test results show the presence of only one cointegrating equation, the one in which g is the dependent variable and y the explanatory one, and the VECM model contains only one statistically significant error correction term in the dynamic equations then, and only then, the block exogeneity test results reported in section 3.3 are valid (Granger, 1988).

In general, if two variables such as g and y are both I(1), any combination of these variables, such as $z = g - a\ y$ will also be I(1), where z is called the equilibrium error term. However, there may exist a singularity a^*, such that g - $a^*\ y$ is I(0). If such a singularity exists, g and y are said to be cointegrated. This implies that in the long-run, although g and y can be arbitrarily high or low, they must be proportional to each other, with a factor of proportionality a^*. It is clearly possible for more than one equilibrium relation to govern the joint behavior of the variables.

Johansen test statistics are shown in Table 6. Following Johansen's methodology, the existence of a deterministic linear trend in the data was tested and could not be rejected. Therefore, the test specification includes such a trend. Both the Trace and the Maximum Eigenvalue tests reject the hypothesis on no cointegration between g and y, but only at the six percent level. A similar problem affects Islam's (2001) cointegration analysis, in which he used a shorter sample spanning the period 1929-1996 and got statistically significant results only at the ten percent level. More analysis on cointegration is thus needed and is conducted below in the context of a VEC model.

5. An Error Correction Model

Engle and Granger (1987) have shown that if two variables are cointegrated, then there must exist a VEC linking these variables. Furthermore, the VEC representation of the bivariate system of cointegrated variables sheds light on the direction of causation between those variables (Granger, 1988).

5.1. Methodological preliminaries:

Following Engle and Granger (1987) the VEC model is formulated as follows:

$$\begin{aligned}(1-L)g_t = c(1) + b(1)\,E(1)_{t-1} + \sum_{j=1}^{T} c_{1,j}\,(1-L)\,g_{t-j} \\ + \sum_{j=1}^{T} d_{1,j}\,(1-L)\,y_{t-j} + u_{1,t}\end{aligned} \qquad (3)$$

$$\begin{aligned}(1-L)y_t = c(2) + b(2)E(2)_{t-1} + \sum_{j=1}^{T} c_{2,j}\,(1-L)\,y_{t-j} \\ + \sum_{j=1}^{T} d_{2,j}\,(1-L)\,g_{t-j} + u_{2,t}\end{aligned} \qquad (4)$$

where L is the lag operator, T is the number of lags to be included, and the error correction terms are given by $E(j)_{t-1}$ for $j = 1, 2$, which are the residuals from the OLS static regressions of g on y and vice versa, respectively. In equations (3) and (4), the VEC model allows for the finding that g Granger-causes y, or vice versa, so long as the corresponding error correction term carries a statistically significant coefficient, even if the estimated d_j coefficients are not jointly statistically significant (Granger, 1988). If g and y are cointegrated, then the error correction terms are stationary, I(0) processes. Conversely, if the residuals from the static regressions involving g and y are I(0), then g and y are cointegrated (Engle and Granger, 1987).

The estimation of the VEC model proceeds in two steps. First, the static equations in the variables' levels are estimated, providing estimates of the cointegrating parameters of the long-term relationship between g and y and the residuals that will enter into equations (3) and (4). Second, the residuals stemming from the static equation estimates are then lagged one period and used in the second step to estimate (3) and (4).

5.2. Estimation of the Static Equations:

The first part of Table 7 shows the OLS estimate for the first static equation, followed by several stationarity tests on the residuals from that regression. Non-stationarity can be soundly rejected in all tests at the one percent level, and the main economic result is that the long-run income elasticity for the relative size of the federal government is roughly 0.8, which is much higher than the 0.4 value previously estimated by Islam (2001) for the period 1929-1996. The second part shows the OLS estimate for the second static equation, followed again by stationarity tests for the residuals. Again, non-stationarity can be soundly rejected at the one percent level in all cases but the DF-GLS test, in which it could only be rejected at the five percent level, still our chosen level for critical values for all tests in both equations. Non-stationarity of the residuals in these static VEC equations is thus rejected, confirming the previous result obtained with the Johansen cointegration tests at the six percent level.

5.3. Estimation of the Dynamic Equations:

Two questions remain unanswered. What is the number of cointegrating relations, and what is the direction of causation? A major practical problem in the estimation of the above VEC equations is the determination of the number of lags. We followed Hendry's "General-to-Specific" Methodology (Gilbert, 1986, provides an excellent exposition), setting $T = 10$ in equations (3) and (4) and proceed to simplify the model toward a parsimonious form by means of likelihood ratio and F tests to eliminate redundant variables. We start with a relatively large number of lags to ensure the absence of significant autocorrelation in the residuals from the dynamic equations (Hendry, 1986: 88).

We report the estimates of the parsimonious model in Table 8. The first column shows the estimates for the first dynamic equation (the rate of growth in government size). The second column displays the estimates for the second dynamic equation (the rate of growth of per capita real GDP). The most important conclusion is that causation runs from economic growth to government size growth, but not the reverse. Indeed, Wald tests for the hypothesis that the cross-terms in the second dynamic equation are simultaneously equal to zero cannot be rejected at conventional levels. Similar Wald tests on the joint significance of the cross-terms in the first dynamic equation produce very different results: the null hypothesis that the cross-terms are jointly statistically insignificant can be rejected at the 1 percent level. These results provide confirmation of the block exogeneity tests presented before in section 3.3.

The insignificance of the lagged residual term in the second dynamic equation and its significance in the first equation conclusively settles the causality question, regardless of the joint significance of the cross-terms (Granger, 1988). There is only one cointegrating equation (static equation 1) and Granger-causality runs in only one direction. Economic growth leads to a larger share of federal government expenditures, but not the reverse. For the United States, at least, the conclusion is that previous evidence that government is bad for growth must be based on poorly specified models.

The Durbin-Watson statistic associated to both dynamic equations' estimates suggests that there is no significant first order autocorrelation problem in the residuals, the number one criterion for model selection in Hendry's methodology (Hendry, 1986: 88). A wide variety of tests were applied to both series of residuals to check for the validity of the stationarity hypothesis. While results are not displayed here, the unit root hypotheses are rejected in all four tests (ADF, DF-GLS, Phillips-Perron, and KPSS) at the one percent level. Figure 4(a) and Figure 4(b) display the graphs of the residuals and confirm the message from the statistical tests: residuals do not seem to be non-stationary.

6. Conclusion

A number of studies over the past two decades have considered whether a larger government is good or bad for growth, but the results have not been conclusive. In this paper, we use annual data for federal government expenditures and real per-capita GDP for the United States, going back to 1792, and we carefully test the time-series properties of these variables for stationarity, cointegration, and Granger causality, an series of steps some studies have begun but not completed. Standard Granger causality tests conducted outside the cointegration framework for up to ten lags reveal that Granger causality is unidirectional, running from economic growth to government size growth. The log levels of these variables are both non-stationary and are cointegrated.

After a careful study of the issue of causation within the cointegration framework, we find that faster growth appears to cause a larger government in the long run, but we don't find significant evidence supporting the hypothesis that the relative size of federal government expenditures affects growth either up or down. Our results confirm some of the results reported by Islam (2001), even though our sample size was much larger, and our investigation and our focus was on uncovering the relationship running from government size to economic growth rather than the reverse. Because the U.S. is a country with high marks for economic freedom and democratic institutions, our results are also consistent with several cross-country studies that found the negative effects of government size on economic growth were much larger in countries with less freedom and democracy.

It remains to be seen, however, whether these results can be generalized to other countries. Relative to other countries at similar levels of development, the U.S. is somewhat of an outlier in that the relative size and role of government is less and the growth in government's size has been much less, as well. Our study did not consider the effects of state and local government, nor did we specify a growth model that considered other factors. Much work remains to determine whether or not the assertion that more government causes lower growth applies elsewhere.

References

Barro, R.J. (1990). Government spending in a simple model of endogenous growth, *Journal of Political Economy* 95(5): S103-125.

Baumol, W.J., & W.G. Bowen (1965). On the performing arts: the anatomy of their economic problems, *American Economic Review* 55(1/2): 495-502.

Conte, M.A., & A.F. Darrat (1988). Economic growth and the expanding public sector: a reexamination, *The Review of Economics and Statistics* 70(2): 322-330.

Demirbas, S. (1999). Cointegration Analysis-Causality Testing and Wagner's Law: The Case of Turkey, 1950-90. Paper presented at the 1999 Annual Meeting of the European Public Choice Society, Lisbon, April 7-10.

Engen, E., & J. Skinner (1996). Taxation and economic growth, *National Tax Journal* 49(4): 617-42.

Engle, R.F. & C.W.J. Granger (1987). Co-Integration and error correction: representation, estimation, and testing, *Econometrica* 55(2): 251-76.

Garrett, T.A., & R.M. Rhine (2006). On the size and growth of government, *Federal Reserve Bank of St. Louis Review* 88(1): 13-30.

Ghali, K.H. (1998). Government size and economic growth: evidence from a multivariate cointegration analysis, *Applied Economics* 31(8): 975-987.

Gilbert, C.L. (1986). Professor Hendry's econometric methodology, *Oxford Bulletin of Economics and Statistics* 48(3): 283-307.

Granger, C.W.J. & P. Newbold (1974). Spurious regressions in econometrics, *Journal of Econometrics* 2(2): 111-120.

Granger, C.W.J. (1986). Developments in the study of cointegrated economic variables, *Oxford Bulletin of Economics and Statistics* 48(3): 213-228.

Granger, C.W.J. (1988). Some recent developments in a concept of causality, *Journal of Econometrics* 39(1-2): 199-211.

Guseh, J.S. (1997). Government size and economic growth in developing countries: A political-economy framework, *Journal of Macroeconomics* 19(1): 175-92.

Gwartney, J., R. Lawson, & R. Holcombe (1998). The size and functions of government and economic growth (Washington, D.C., Joint Economic Committee, U.S. Congress).

Halicioglu, F. (2003). Testing Wagner's law for Turkey, 1960-2000, *Review of Middle East Economics and Finance* 1(2): 129-40.

Hendry, D.F. (1986). Using PC-GIVE in econometrics teaching, *Oxford Bulletin of Economics and Statistics* 48(3): 87-98.

Islam, A.M. (2001). Wagner's law revisited: cointegration and exogeneity tests for the USA, *Applied Economics Letters* 8(8): 509-515.

Jackson, P. M., M. D. Fehti and S. Fehti (1998), Cointegration, Causality and Wagner's Law: A Test for Northern Cyprus. Paper presented at the Second International Congress on Cyprus Studies.

Johnston, L.D., & S.H. Williamson (2006), The annual real and nominal GDP for the United States, 1790 – Present, Economic History Services, http://eh.net/hmit/gdp.

Jones, J.D., & D. Joulfaian (1991). Federal government expenditures and revenues in the early years of the American Republic: evidence from 1792-1860, *Journal of Macroeconomics* 13(1): 133-55.

Landau, D. (1983). Government expenditure and economic growth: a cross-country study, *Southern Economic Journal* 49(3): 783-92.

Landau, D. (1986). Government and economic growth in the less developed countries: an empirical study for 1960-80, *Economic Development and Cultural Change* 35(1): 35-75.

Lee, B.S., & S. Lin (1994). Government size, demographic changes, and economic growth, *International Economic Journal* 8(1): 91-108.

Lindert (2004). *Growing Public: Social Spending and Economic Growth Since the Eighteenth Century,* vols. 1-2 (Cambridge University Press).

Murray, C.J., & C.R. Nelson (1998), The uncertain trend in U.S. GDP. Unpublished manuscript, University of Washington, Seattle.

Peacock, A., & A. Scott (2000). The curious attraction of Wagner's Law, *Public Choice* 102(1-2): 1-17.

Plümper, T., & C.W. Martin (2003). Democracy, government spending, and economic growth: A political-economic explanation of the Barro-effect, *Public Choice* 117(1-2): 27-50.

Poot, J. (2000). A synthesis of empirical research on the impact of government on long-run growth, *Growth and Change* 2000 31(4): 516-546.

Ram, R. (1986). Government size and economic growth: a new framework and some evidence from cross-section and time-series data, *American Economic Review* 76(1): 191-203.

Scully, G.W. (2002). Economic freedom, government policy and the trade-off between equity and economic growth, *Public Choice* 113(1-2): 77-96.

Slemrod, J. (1995). What do cross-country studies teach about government involvement, prosperity, and economic growth?, *Brookings Papers on Economic Activity* 1995(2): 373-431.

Wagner, A. (1893). *Grundlegung der Politischen Okonomie*, 3rd Edition (Leipzig: C.F. Winter'sche Verlagshandlung).

Zellner, A. (1988). Causality and causal laws in economics, *Journal of Econometrics* 39(1,2): 7-28.

Table 1: Average Growth and Federal Government Expenditures in the United States, 1792-2004

Time Period	Real GDP Growth	Real Per-Capita GDP Growth	Standard Deviation	Federal Government Expenditures (Share of GDP)	Standard Deviation
1792-1860	4.3%	1.3%	2.5%	2.5%	0.7%
1861-1870	3.3%	0.9%	4.3%	9.4%	5.4%
1871-1915	3.7%	1.5%	5.0%	3.0%	0.7%
1916-1929	3.9%	2.5%	5.3%	7.8%	8.3%
1792-1929	**4.0%**	**1.5%**	**3.9%**	**3.7%**	**3.7%**
1930-1950	3.5%	2.4%	9.7%	18.0%	13.4%
1951-1970	3.8%	2.3%	2.7%	21.4%	1.8%
1971-1990	3.2%	2.2%	2.4%	23.8%	1.1%
1991-2004	3.0%	1.8%	1.4%	20.3%	1.8%
1950-2094	**3.4%**	**2.1%**	**2.3%**	**21.9%**	**2.2%**

Table 2: VAR Lag Order Selection Criteria						
Lag	LogL	LR	FPE	AIC	SC	HQ
0	316.18	N/A	0.000153	-3.111	-3.078	-3.097
1	327.86	23.00	0.000142	-3.187	-3.088*	-3.147*
2	332.86	9.76	0.000140	-3.197	-3.033	-3.130
3	334.99	4.11	0.000143	-3.178	-2.949	-3.085
4	347.41	23.74*	0.000130*	-3.262*	-2.967	-3.142
5	350.11	5.09	0.000133	-3.249	-2.888	-3.103
6	352.42	4.34	0.000135	-3.232	-2.806	-3.060
7	353.64	2.25	0.000139	-3.204	-2.713	-3.006
8	356.21	4.71	0.000141	-3.190	-2.633	-2.965
9	357.59	2.51	0.000145	-3.164	-2.542	-2.912
10	361.79	7.51	0.000145	-3.166	-2.478	-2.888

* indicates lag order selected by the criterion

LogL: Log Likelihood

LR: sequential modified LR test statistic (each test at 5% level)

FPE: Final prediction error

AIC: Akaike information criterion

SC: Schwarz information criterion

HQ: Hannan-Quinn information criterion

Table 3: Vector Autoregression Estimates		
	Equation 1	Equation 2
	Δg	Δy
Δg_{t-1}	0.143	0.014
	[2.11]	[1.26]
Δg_{t-2}	-0.207	-0.010
	[-3.01]	[-0.90]
Δg_{t-3}	0.0235	-0.014
	[0.34]	[-1.23]
Δg_{t-4}	-0.123	-0.013
	[-1.81]	[-1.19]
Δy_{t-1}	0.376	0.262
	[0.87]	[3.77]
Δy_{t-2}	0.319	-0.020
	[0.72]	[-0.28]
Δy_{t-3}	0.639	-0.029
	[1.44]	[-0.40]
Δy_{t-4}	-1.500	-0.169
	[-3.53]	[-2.46]
Constant	0.016	0.016
	[0.75]	[4.69]
R-squared	0.142	0.155
Adj. R-squared	0.107	0.121
Sum sq. resides	13.219	0.346
S.E. equation	0.258	0.042
F-statistic	4.10	4.55
Log likelihood	-8.53	370.43
Akaike AIC	0.169	-3.475
Schwarz SC	0.313	-3.331
[t-statistics in brackets]		

Table 4: Block-Exogeneity Tests			
Pairwise Granger F-Tests			
A) H0: Lags of Δy are NOT block-exogenous			
	Obs	F-stat	p-value
4 LAGS	208	4.01	0.0037
5 LAGS	207	2.96	0.0135
6 LAGS	206	2.90	0.0100
7 LAGS	205	2.72	0.0105
8 LAGS	204	2.47	0.0143
9 LAGS	203	2.21	0.0230
10 LAGS	202	2.17	0.0212
B) H0: Lags of Δg are NOT Block-exogenous			
	Obs	F-stat	p-value
4 LAGS	208	1.58	0.1812
5 LAGS	207	1.79	0.1174
6 LAGS	206	1.48	0.1872
7 LAGS	205	1.18	0.3179
8 LAGS	204	1.32	0.2381
9 LAGS	203	1.04	0.4078
10 LAGS	202	1.40	0.1820

Table 5: Unit Root Tests for both *y* and *g*				
Variables	ADF /1	DF-GLS /2	Phillips-Perron /3	KPSS /4
Levels				
y	0.49	3.04	0.83	1.85
g	-1.86	-1.51	-1.51	1.47
1st Differences				
y	-10.7	-8.83	-10.27	0.16
g	-11.42	-3.31	-13.95	0.07

Notes:

/1: The null hypothesis is that the series contains an autoregressive unit root.

ADF is the t-ratio corresponding to the Augmented Dickey-Fuller test.

The critical values for 1%, 5%, and 10% risk levels are, respectively: -3.46, -2.87, -2.57

/2: The null hypothesis is that the series contains an autoregressive unit root

DF-GLS is the t-ratio corresponding to the Dickey-Fuller test applied on a GLS regression

The critical values for 1%, 5%, and 10% risk levels are, respectively: -2.57, -1.94, -1.61

/3: The null hypothesis is that the series contains an autoregressive unit root.

Phillips-Perron is the t-ratio stemming from an autoregression of the series with no lagged first diff.

The critical values for 1%, 5%, and 10% risk levels are, respectively: -3.46, -2.87, -2.57

/4: The null hypothesis is that the series is stationary (i.e., no autoregressive unit root exists)

KPSS is the Lagrange Multiplier, LM statistic.

The critical values for 1%, 5%, and 10% risk levels are, respectively: 0.74, 0.46, 0.35

Table 6: Johansen's cointegrations tests	
Four lags in Differenced Variables Deterministic linear trend allowed in data	
H0: No cointegration equation exists	
Trace Stat	Max Eigen Stat
14.63	13.70
Critical values (at 5% level)	
Trace Stat	Max Eigen Stat
15.49	14.27
p-values	
Trace Test	Max Eigen Test
0.07	0.06
H0: No more than 1 cointegration eq.exists	
Trace Stat	Max Eigen Stat
0.92	0.92
Critical values (at 5% level)	
Trace Stat	Max Eigen Stat
3.84	3.84
p-values	
Trace Test	Max Eigen Test
0.34	0.34

Table 7: Error Correction Model, Static Equations

FIRST STATIC EQUATION			
Dependent Variable: g			
	C	y	
coefficients	-9.544	0.791	
t-stat	-31.85	22.59	
p-value	0.0000	0.0000	
R-squared	0.707	F-statistic	510.24
Adjusted R-squared	0.706	Prob(F-statistic)	0.0000
S.E. of regression	0.543	Durbin-Watson stat	0.25
Included observation 213			
Stationarity tests applied to Residuals of Static Equation 1			
	stat	critical value (5%)	
ADF test	-4.26	-2.88	
DF GLS test	-2.58	-1.94	
PP test	-4.03	-2.88	
KPSS test	0.17	0.46	
SECOND STATIC EQUATION			
Dependent Variable: y			
	C	g	
coefficients	11.015	0.8940	
t-stat	92.82	22.5885	
p-value	0.0000	0.0000	
R-squared	0.707	F-statistic	510.24
Adjusted R-squared	0.706	Prob(F-statistic)	0.0000
S.E. of regression	0.577	Durbin-Watson stat	0.18
Included observations 213			
Stationarity tests applied to Residuals of Static Equation 2			
	stat	critical value (5%)	
ADF test	-3.195	-2.875	
DF GLS test	-3.616	-1.942	
PP test	-3.318	-2.875	
KPSS test	0.882	0.463	
Notes:			
H0 is Non-stationarity of residuals for ADF, DF-GLS and PP tests			
H0 is Stationarity of residuals for KPSS test			

Table 8: Error Correction Model, Dynamic Equations		
	First Dynamic Equation	Second Dynamic Equation
	Dependent Variable: Δg	Dependent Variable: Δy
Δg_{t-1}	0.199	0.016
	[2.98]*	[1.44]
Δg_{t-2}	-0.107	-0.011
	[-1.57]	[-0.98]
Δy_{t-1}	0.558	0.268
	[1.38]	[4.00]*
Δy_{t-4}	-1.229	-0.198
	[-3.08]*	[-3.02]*
Constant	0.022	0.016
	[1.07]	[4.66]*
$E(1)_{t-1}$	-0.117	
	[-3.32]*	
$E(2)_{t-1}$		-0.001
		[-0.11]
R-squared	0.160	0.137
Adjusted R-squared	0.139	0.116
S.E. of regression	0.253	0.0418
F-statistic	7.69*	6.43*
Durbin-Watson stat	1.96	2.00
t-statistics in [.] below coefficients; * = significant at 5% level		

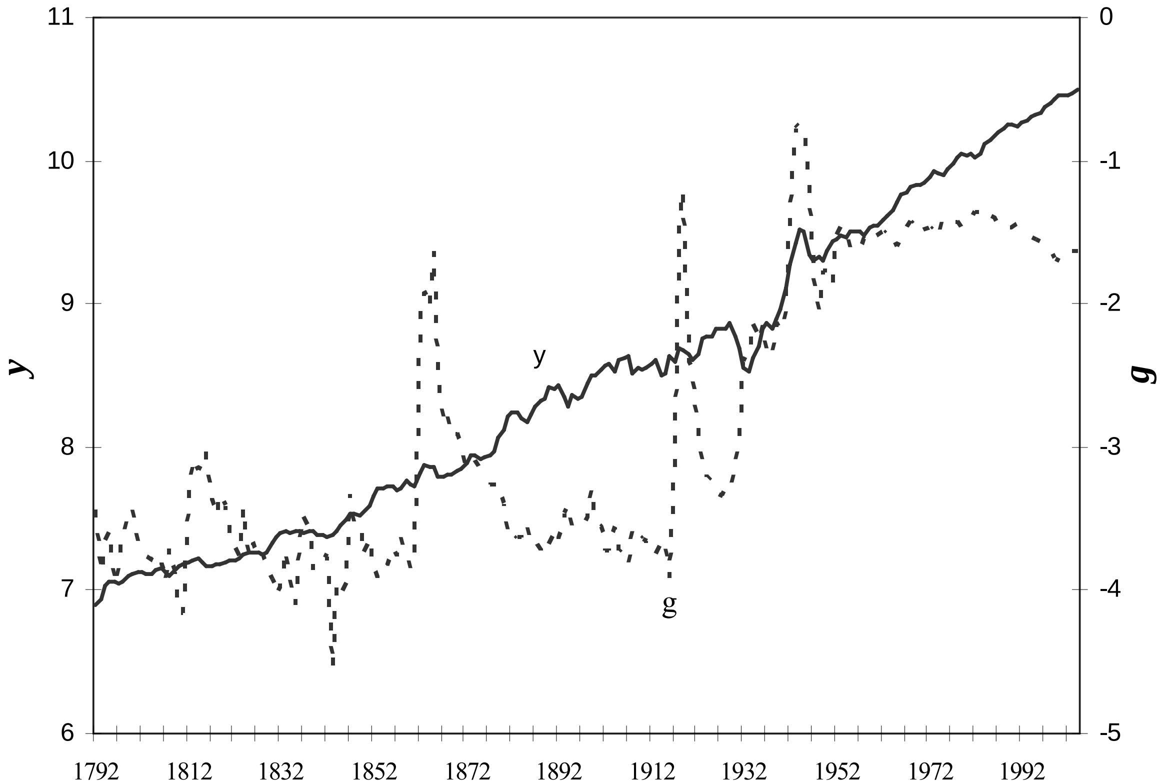

Figure 1: Per-Capita lncome and Government Expenditures, USA

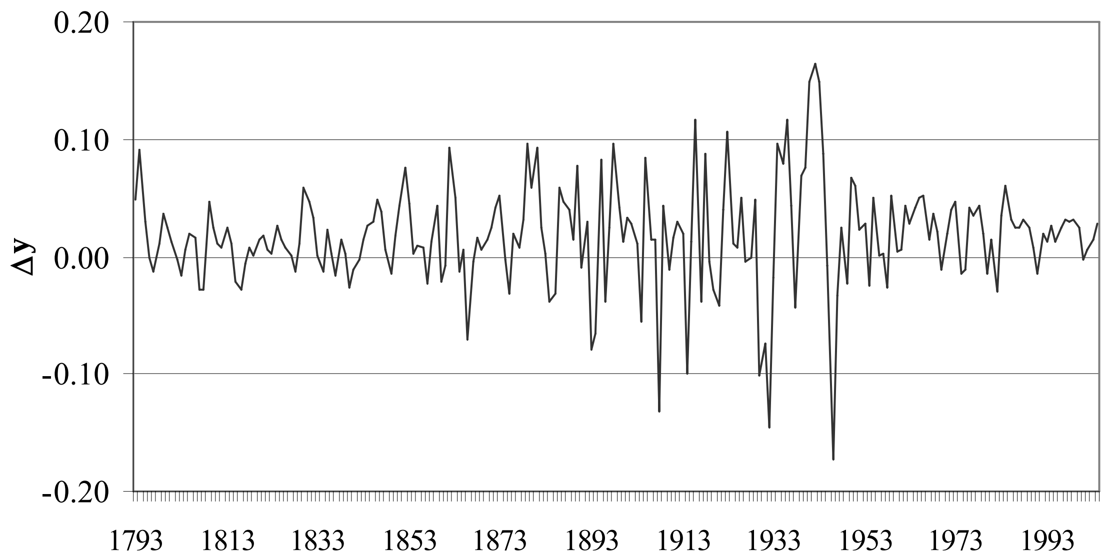

Figure 2(a): First Difference of y

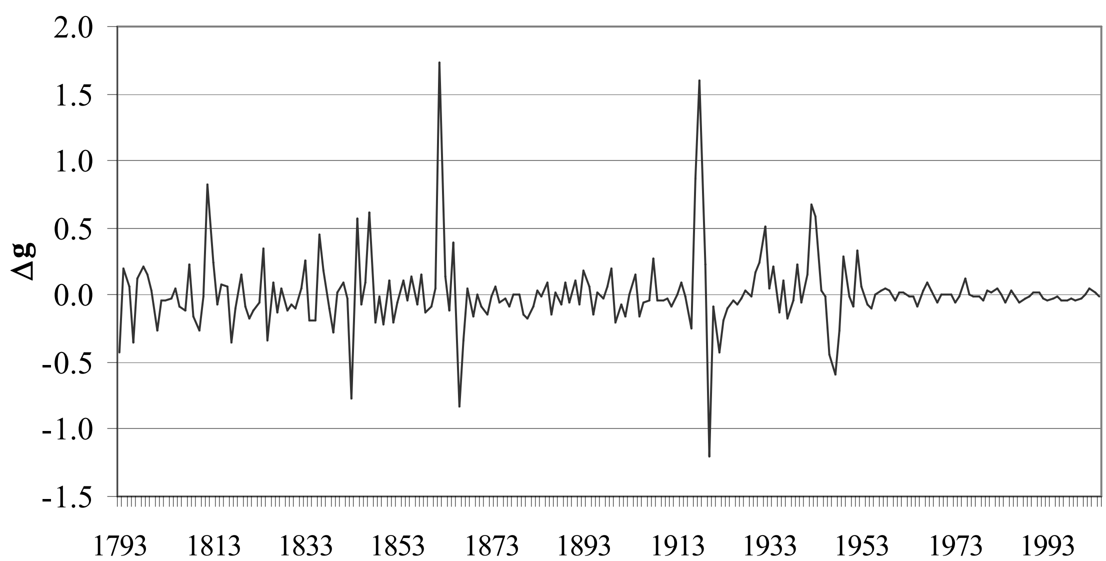

Figure 2(b): First Difference of g

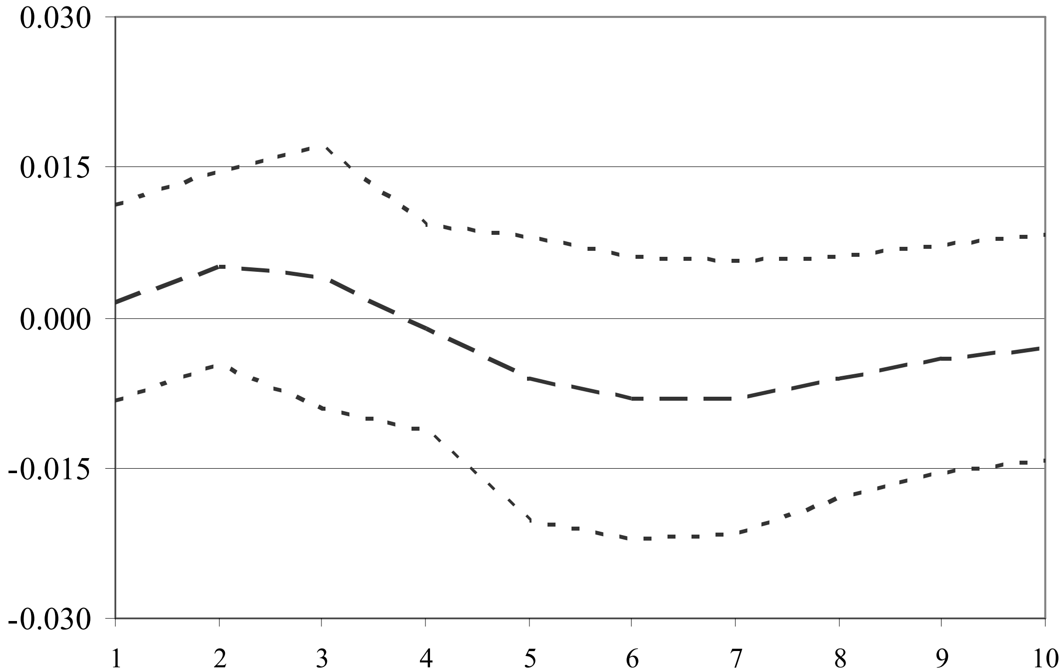

Figure 3(a): Accumulated Response of Δy to Δg

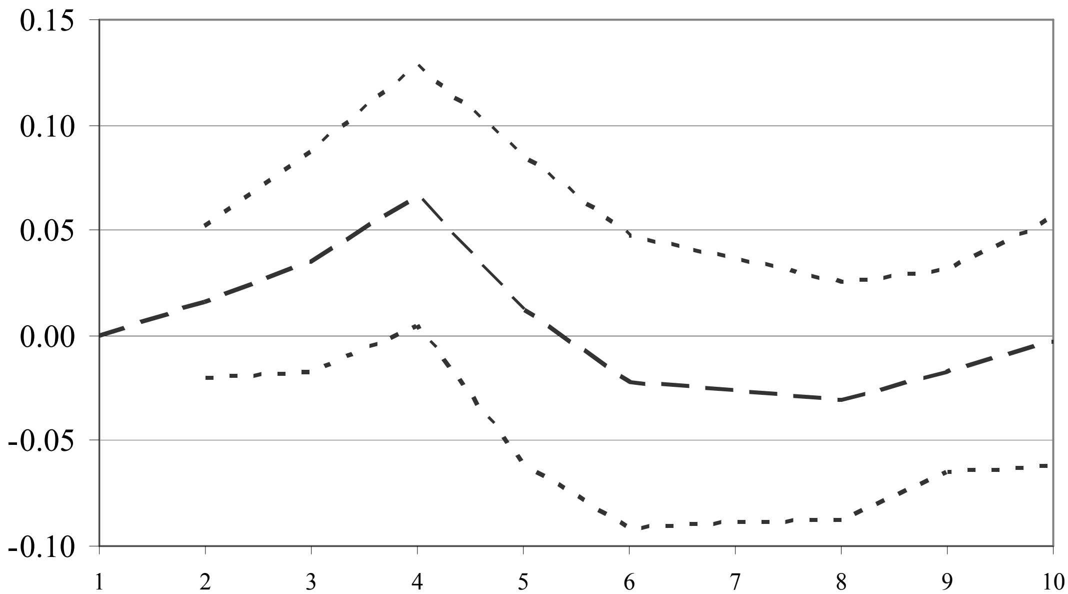

Figure 3(b): Accumulated Response of Δg to Δy

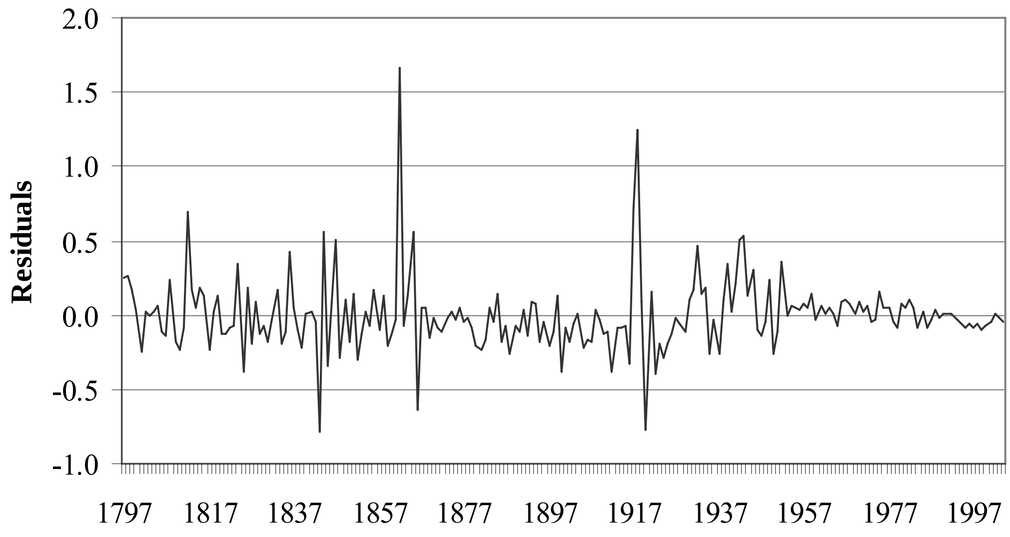

Figure 4(a): Residuals of the First Dynamic Equation

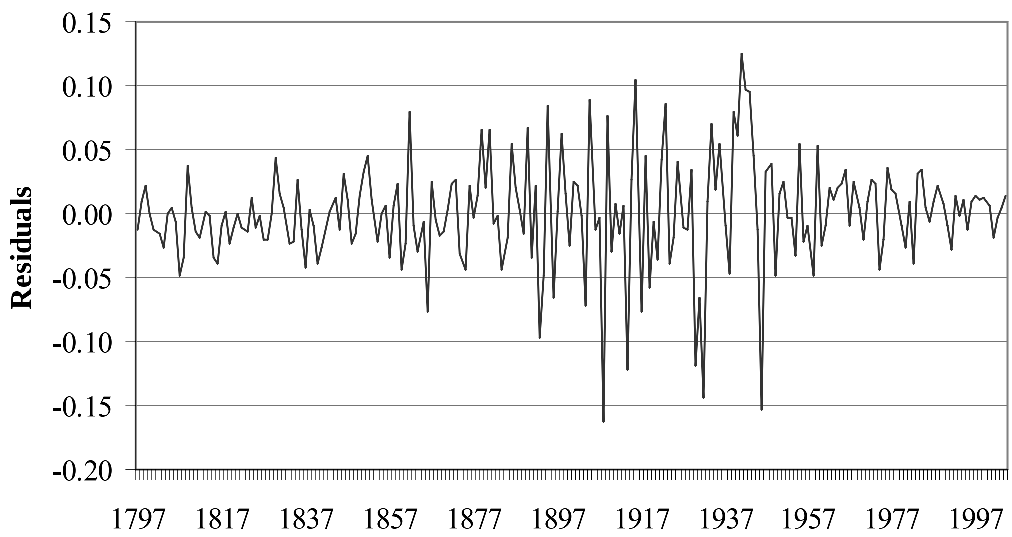

Figure 4(b): Residuals of the Second Dynamic Equation