THE UNIVERSITY OF TEXAS AT SAN ANTONIO, COLLEGE OF BUSINESS

Working Paper SERIES

Date July 23, 2009

WP # 0101ECO-090-2009

A Re-examination of Wagner's Law Based on Disaggregated U.S. State-Local Government Expenditure

Saeid Mahdavi

Associate Professor of Economics
The University of Texas at San Antonio

UTSA COLLEGE OF BUSINESS KNOWLEDGE FOR A NEW WORLD

ONE UTSA CIRCLE
SAN ANTONIO, TEXAS 78249-0631
210 458-4317 | BUSINESS.UTSA.EDU

A Re-examination of Wagner's Law Based on Disaggregated U.S. State-Local Government Expenditure

Saeid Mahdavi
Associate Professor of Economics
The University of Texas at San Antonio
One UTSA Circle, San Antonio, TX 78249
Tel.: (210) 458-5301, E-mail: smahdavi@utsa.edu

Abstract

We tested the validity of the "Law of Increasing State Activities" or Wagner's Law using time series for the U.S. state-local government (SLG) real expenditure over the period 1957-2006. This period was characterized by rising SLG total expenditure and several of its sub-categories both in absolute terms and relative to state personal income. Cointegration tests of Johansen (1991) and Pesaran, Shin, and Smith (2001) yielded results suggesting that, with the exception of insurance trust benefits (*ins*) and social services and income maintenance (*ssim*) ratios, no other nonstationary expenditure sub-category was cointegrated with state real per capita personal income (*pcpi*) and error-corrected over time. Both *ins* and *ssim* were found to be income elastic. The Toda-Yamamoto (1995) methodology that allows for estimating *level* relationships without pre-testing for unit roots further suggested that *ins* and *ssim* were driven by *pcpi* consistent with a Wagnerian causal ordering.

JEL classification: H7; C22.

Key words: Wagner's Law; state and local governments; public expenditures; cointegration.

I. Introduction

The "Law of Increasing State Activities" enunciated by the German social scientist Adolph Wagner (1893) is one of the earliest hypotheses of the growth of the relative size of the public sector. Generally speaking, Wagner's Law focuses on the nexus between the size of the economy and the size of the public-sector provided goods and services and postulates that the latter grows at a faster pace than the former during the process of industrialization and urbanization. This reflects the increasing expansion of government activities that complement or substitute for private activities. Specifically, Wagner attributed the growth of the public sector to higher expenditures in areas such as enforcing contracts and regulatory activities (necessitated by a higher demand for government intervention in an economy with new layers of externalities and interdependencies), income elastic "cultural and welfare" programs, and public long-term investment and infrastructure projects as well as managing and financing natural monopolies (see, for example, Bird, 1971 and Payne et al., 2006). Based on a more careful reading of Wagner's original writings, Peacock and Scott (2000) and Peacock (2006) maintain that Wagner's definition of "state activities" should include those related to public utilities and enterprises, public provision of health and educational services and a social system of security to protect the (working) population against adverse social consequences of economic transformation.

Many of the activities noted by Wagner have been incorporated into modern theories of the public sector activity. Examples include government interventions to correct externalities (the Neoclassical School), public redistributive expenditures in response to increased demand by the "median voter" for such expenditures (the Public Choice School) and countercyclical government spending (the Keynesian School). However, the evidence supporting the "law" as a frequently observed empirical phenomenon is less than robust. The validity of Wagner's Law has been typically tested based on whether the income (output) elasticity of a measure of government expenditure is statistically significant and greater than one. To this end, a large number of past studies used data from one or a cross-section of countries and employed simple regression analysis and/or standard Granger causality tests to draw conclusions about the elasticity coefficient and the direction of causality in the income-public expenditure nexus.[1] The evidence from such studies may be questionable for the following reasons: Firstly, Wagner's Law focuses on *the long-term* relationship between the size of the public sector and the size of economy (income/output) in the process of economic growth and development in an individual country (or unit of analysis). Thus, multi-country samples typically spanning about three decades may not be long-enough for testing such a relationship. Secondly, it may not be appropriate to pool countries that have a number of social, political and institutional dissimilarities. Thirdly, it is well known that standard regression results may be spurious if the variables employed are nonstationary in levels (Granger and Newbold, 1974). Moreover, if the nonstationary variables are cointegrated, then they have an error-correction representation (Engle and Granger, 1987) requiring the inclusion of an error-correction term into the causality analysis.

With the above points in mind, we investigate the relationship between public expenditure and income using annual observations on the U.S. state-local level of government (SLG). The sample period covers about half-a-century (1957-2006) which may better allow the emergence of a long-term relationship, if any. Moreover, we go beyond the empirical analysis of total government expenditures (as in Islam, 2001) and test the validity of Wagner's Law for various expenditure sub-categories. This feature of our analysis is important in view of the fact that (a) many of the activities of the public sector emphasized by Wagner fall within the realm of SLG responsibilities, (b) various sub-categories of SLG total expenditure have grown at different rates over time and (c) SLG expenditure sub-categories cannot grow much faster than state income or fiscal capacity without creating serious financial stress for sub-national governments whose ability to borrow and run deficits is constrained by balanced budget rules. To our knowledge, there are only a few studies that examined the validity of Wagner's Law using the

[1] The literature on country-level studies is voluminous. For a summary some of major studies see, for example, Peacock and Scott (2000) and Payne *et al*. (2006).

U.S. sub-national government data. However, all of these studies specified models that were estimated using the OLS technique without pre-testing the level variables for stationarity.[2]

In Section II of what follows, we briefly describe our data and examine changes in the level and composition of SLG expenditure. In Section III, we first specify the empirical model corresponding to the version of Wagner's Law tested in this paper. We then explain the empirical methodology and discuss the results. The methodology emphasizes the questions of cointegration and causality between SLG expenditure and income. In addition to standard tests, we employ the bounds testing of cointegration proposed by Pesaran, Shin, and Smith (2001) and an estimation approach to drawing Granger non-causality inferences proposed by Toda and Yamamoto (1995) to shed light on these questions. These time-series techniques are particularly helpful when the order of integration of the series cannot be conclusively determined based on the results of alternative unit root tests. The final section summarizes our findings and notes their implications for the federal relief funds to fiscally stressed sub-national governments.

II. Data: The level and composition of U.S. state-local expenditure

Our data set consists of annual observations on the U.S. SLG total expenditure and eight of its major sub-categories (*source: State and local Government Finances, U.S. Census Bureau*) and aggregate state personal income (*source: Regional Economic Accounts, U.S. Bureau of Economic Analysis*) over the period 1957-2006. The time series for the expenditure and personal income variables at this intermediate level of aggregation were deflated using the price indexes

---

[2] Yousefi and Abizadeh (1992) used time series for thirty randomly selected states (1950-85) and a multivariate model. They found income elasticity estimates consistent with Wagner's Law for twenty-one states. Eberts and Gronberg (1992) employed the share in state output of total government spending as well as several of its sub-categories (1964-86). Controlling for the sectoral composition of state output and state age, they found statistically significant and *negative* per capita output elasticity of the share variables in pooled data which refuted Wagner's Law. However, they also reported supporting evidence in relation to "protective services" (in eight states) and "public welfare" (in fifteen states) when individual state time-series data were used instead. Grand (1998) studied fifty states (1945-1984) and regressed the share of state spending in total state output on lagged levels of state personal income and state population. His results supported Wagner's Law in only nine states. See Henrekson (1993) for a criticism of using of possibly nonstationary level variables in level equations in this context.

for total state and local government consumption and gross investment expenditure and U.S. GDP (2000=100), respectively (*source: U.S. Bureau of Economic Analysis, Table 3.9.4*).[3] The use of separate price indexes allows for differential changes in the prices of goods and services purchased by the SLGs versus those that are included in a larger basket that constitutes GDP.

Table 1 summarizes the changes in real SLG total expenditure and its major sub-categories as well as real aggregate state personal income over the period 1957-2006. Several observations can be made based on the table. Firstly, real total SLG expenditure increased at an average annual rate of 3.56 percent during this period which is only slightly higher than the corresponding 3.38 percent growth rate of real state personal income. This fairly small growth rate differential, however, translated into roughly 1.7 percent increase in the ratio of SLG total expenditure to personal income during the fifty years of the sample. Secondly, the expenditure sub-categories exhibited three distinct patterns of change: (a) highway expenditure grew at a rate well below those of real total expenditure and real state personal income. As a result, it lost in terms of its shares of total expenditure and income. (b) Police and fire protection, education, and utility were among the sub-categories that grew at rates roughly equal to the growth rates of total spending and income. This explains their remarkably stable expenditure and income shares during the sample period. (c) Increases in insurance trust benefits (especially its employment retirement component), social services and income maintenance (especially its public welfare component), interest on general debt, and financial administration and general control expenditure sub-categories significantly outpaced increases in total expenditure and personal income. As a consequence, these sub-categories gained shares.[4]

---

[3] The choice of the sample period and the variables were based on data availability on a consistent basis. In this connection, note that gross state product (GSP) that one might have preferred to state personal income was available based on a methodology and industrial classification that changed after 1997.

[4] Daly (2003) pointed out that states have a tendency to rapidly expand funding to various programs during economic booms and, in so doing, sharply depart from standard spending rules of constant real spending relative to income or constant real spending on a per capita basis. As noted above, in the longer-term, this departure is more pronounced in relation to only a few sub-categories.

These observations may be interpreted as *prima facie* evidence in favor Wagner's Law in relation to SLG total expenditure and some of its sub-categories. The patterns noted above also suggest that formal analyses Wagner's Law may be more informative if they cover various types of public expenditure.

III. Empirical Methodology and Results

Since Wagner emphasized the growth of the public sector in a relative sense, we choose the following popular formulation due to Musgrave (1969) to test his hypothesis:[5]

$$\ln\left(\frac{G}{Y}\right)_t = \alpha + \beta \ln\left(\frac{Y}{P}\right)_t + \mu_t \qquad (1)$$

where ln denotes the natural logarithm, G represents a measure of government expenditure, Y is a measure of the size of the (private) economy, P is the size of population, and μ is a white noise error term. Note that if the elasticity coefficient β is positive and statistically significant, then we have statistical evidence consistent with Wagner's Law.[6] For the purpose of our analysis, we use SLG real total expenditure and its major sub-categories as alternative measures of public expenditure and real aggregate state personal income (PI) a proxy for the size of the economy (see Table 1). Next, we estimate the following long-term relationship for the combined SLG unit in year t:

$$\ln(g_j)_t = \alpha_j + \beta_j \ln(pcpi)_t + \varepsilon_{j,t} \qquad (2)$$

where, g_j is the jth real expenditure sub-category scaled by real personal income and *pcpi* is real per capita personal income. Note that while parsimonious, the model incorporates the effects of changes in population, price level, and the size of the economy on SLG expenditure.

[5] There are a number of other proposed bivariate specifications that employ various combinations of the expenditure and income variables expressed in level terms, or scaled by population and/or output. See, for example, Peacock and Scott (2000) and Payne *et al*. (2006).

[6] Equation (1) can be rewritten to express all the variables in level terms. Then, it can be easily shown that the elasticity coefficient E=$\partial \ln G / \partial \ln Y = 1 + \beta$. If β>0, then E >1 which means that the percentage change in the level of government spending is larger than the percentage change in the level of income.

Since it is generally understood that Wagner's Law is most relevant to economies in the early stages of economic development (Peacock and Scott, 2000), the sole focus on estimating the income elasticity coefficient in relation to mature economies may be somewhat misplaced. A particularly important and relevant empirical question that we focus on in this paper is whether the government expenditure and income variables are cointegrated in such economies.[7] If so, they do not drift too far apart over time even if their levels are nonstationary; for the two cointegrated variables have an error-correction representation (Engle and Granger, 1987). Thus, a shock to a variable that causes it to deviate from the long-term equilibrium (or cointegrating) relationship will be corrected over time so that the equilibrium relationship is restored. The existence of a cointegrating relationship and an error-correction mechanism is particularly important to explore in the case of SLG expenditure sub-categories such as insurance trust benefits, social services and income maintenance and interest on debt whose high growth rates , even on an inflation adjusted basis, have raised public concerns. Fortunately, the questions of elasticity and cointegration can be both investigated within the framework of an error-correction model (EMC).

For the jth spending category in year t, the ECM may be written as follows:

$$\Delta \ln(g_j)_t = c_j + \sum_{i=1}^{p} \phi_{j,i} \Delta \ln(g_j)_{t-i} + \sum_{i=1}^{p} \varphi_{j,i} \Delta \ln(pcpi)_{t-i} + \lambda_j [\ln(g_j)_{t-1} - \beta_j \ln(pcpi)_{t-1}] + \xi_{j,t} \quad (3)$$

According to Equation (3), short-term changes in g_j reflect an adjustment to deviations from its long-term (cointegrating) relationship with *pcpi*. The magnitude of this adjustment in each period is given by the error-correction coefficient λ_j. Therefore, contrary to Peacock and

[7] In this connection, Peacock and Scott (2000, p.10) contended that: "Interestingly, the cointegration approach, which attempts to identify a long-term connection among variables which may not be apparent in a multiple regression test, could be regarded as consistent with Wagner's view that there was not necessarily a cause and effect relationship between economic development and government activity; this is because the existence of cointegration does not imply causality." They also suggested that "Wagner's writings imply that he would have been satisfied with cointegration." We will return to the issue of the link between cointegration and causality later in this section.

Scott's (2000) claim, cointegration between g_j and *pcpi* does imply causality in at least one direction (Engle and Granger, 1987). A statistically significant and negative value of λ_j is taken as evidence of a causal relationship running from *pcpi* to g_j. A similar inference can be made regarding causality from g_j to *pcpi* based on the error-correction term in a model with *pcpi* as the dependent variable.

We first use the maximum likelihood approach of Johansen and Juselius (1990) and Johansen (1991) to determine whether g_j and *pcpi* are cointegrated and then estimate the cointegrating vector, if any. This approach requires pre-testing of the variables in order to determine whether they are first-differenced stationary, or integrated of order one (I(1)).

a. Unit root tests

We examine the time series property of each variable using the ADF (Dickey and Fuller, 1979), PP (Phillips and Perron, 1988), and KPSS (Kwiatkowski, Phillips, Schmidt, and Shin, 1992) unit root tests. While the null hypothesis of the ADF and PP tests is that the variable in question has a unit root or is nonstationary in level, the KPPS test has the null of stationary level. Our unit root test results are considered robust if we can reject the null hypothesis in the ADF and PP tests *and* fail to reject it in the KPPS test.

Table 2 presents the results of the unit root tests. For each variable, the tests were carried out with a constant term (C) and with and without a trend variable (T) in the test equation. As can be seen, there is some evidence of nonstationarity of the level of all the variables with the exception of total expenditure (*te*). Note the KPPS test results contradict the finding of the existence of a unit root based on the ADF and PP tests in some cases. In these cases, we refer to our test results as inconclusive. Further tests indicated that all the variables became stationary after first-differencing (results are not shown).

b. Cointegration tests

We now proceed to perform the cointegration tests given that, except for the total expenditure variable (*te*), there is *some* evidence of level nonsationarity and first-differenced stationarity for all the variables. Table 2 summarizes the cointegration rank test results using the procedure suggested by Johansen and Juselius (1990) and Johansen (1991).[8] According to both the trace and maximum eigenvalue test statistics the null hypothesis of no cointegrating relationship (r=0) between expenditure and *pcpi* can be rejected in favor of the alternative hypothesis of at most one cointegrating relationship ($r \leq 1$) only for the insurance trust benefits (*ins*) and the social services and income maintenance (*ssim*) variables. Furthermore, the estimated coefficients of *pcpi* in the cointegrating equations for *ins* and *ssim* are positive (0.425 and 0.169, respectively) and significant. This implies that the corresponding expenditure levels are income elastic as hypothesized by Wagner (see footnote 5). The results of our formal statistical analysis are, thus, consistent with Wagner's Law and corroborate the informal evidence presented in Table 1 for these two sub-categories.[9]

Having tested for cointegration, we turn to the question of the direction of short-term causality based on the estimated error-correction coefficient (λ). As can be seen, λ has the correct negative sign and is statistically significant for both *ins* and *ssim*. Its magnitude indicates the fraction the deviation of each variable from its long-term relationship with *pcpi* that is corrected each year (0.433 and 0.180, respectively). Since the error-correction terms in similar ECMs with *pcpi* as the dependent variable was statistically insignificant (results are not shown

[8] Since tests results are sensitive to specification of the cointegrating equation, we compared different models based on the Schwarz information criterion. In all cases, the model with an intercept term only was selected over the model including both intercept and trend terms.

[9] A somewhat different interpretation of the *ssim* result is to suggest that an increase in *pcpi* increases public spending on social welfare programs, because it represents a higher level of state fiscal capacity (U.S. Department of Health and Human Services, 2004). Also note that the *ssim* result is consistent with Eberts and Gronberg's (1992) finding in relation to "public welfare" expenditure noted earlier.

here) the causality implication is that *pcpi* drives both *ins* and *ssim*, but the reverse does not hold true.

An alternative approach to cointegration is the "bounds testing" in the context of an autoregressive distributed lag (ARDL) model proposed by Pesaran *et al.* (2001). The advantage of this approach is that it does not require pre-testing of the variables to establish the order of integration. For this reason, it can be used to examine the existence of a level relationship irrespective of whether the variables of interest are purely I(0) or purely I(1). This is an important advantage in view of the well-known problem of the low power of the unit root tests in small samples and the fact that the results obtained from alternative unit root tests do not always agree. The ARDL estimated in this approach is essentially an unrestricted ECM which for the jth expenditure variable is specified as follows:

$$\Delta \ln(g_j)_t = d_j + \sum_{i=1}^{m} \gamma_{j,i} \Delta \ln(g_j)_{t-i} + \sum_{i=1}^{m} \eta_{j,i} \Delta \ln(pcpi)_{t-i} + \pi_{j,1} \ln(g_j)_{t-1} + \pi_{j,2} \ln(pcpi)_{t-1} + \psi_{j,t} \quad (4)$$

Equation (4) is estimated using OLS and the joint hypothesis of $\pi_{j,1} = \pi_{j,2} = 0$ is then tested using an F-test (or a Wald test). The rejection of the null is taken as evidence in favor of a cointegrating relationship. Since the asymptotic distribution of the *F*-statistic is non-standard under the null hypothesis and the variables can be either I(0) or I(1), Pesaran *et al*. (2001) provide two sets of asymptotic critical values (associated with different number of variables in the cointegrating space) and use them as critical value bounds for testing.

More specifically, the critical value with all the variables assumed to be purely I(0) constitutes the lower bound (CV_L) and that associated with all the variables assumed to be purely I(1) the upper bound (CV_U). If the computed *F*-statistic> CV_U, then the null hypothesis of no cointegration is rejected. If the *F*-statistic<CV_U, then the null hypothesis of no cointegation can not be rejected. Finally, if CV_U<*F*-statistic<CV_U, then the test result is inconclusive.

Table 4 summarizes the bounds test results.[10] As before, we find evidence of cointegration only in the case of *ins* and *ssim* variables. For other expenditure variables, the test results are either inconclusive (*te, fagc* and *pf*), or they reject cointegration (*ed, ins, int, hy,* and *ut*). Thus, the bounds cointegration test results confirm our earlier conclusions based on Table 3.

c. Toda-Yamamoto non-causality tests

If one interprets Wagner's Law as suggesting that the process of economic growth and development *necessitates* the expansion of the public sector, then there is an implication a causal effect running from the level of (per capita) income to government spending. On the other hand, within the Keynesian framework causality may run in the opposite direction.[11]

As we have before, the ECM provided some evidence of causality from *pcpi* to *ins* and *ssim*. For the sake of completeness, we proceed to conduct further causality tests employing a methodology suggested by Toda and Yamamoto (1995). Here, unlike standard Granger causality tests, a *level* VAR is specified and estimated without having to pre-test the variables for the degree of integration and/or cointegration rank(s). This approach has been employed to make inferences regarding *long-term* causality among variables in a level relationship like Equation (2).[12] The steps in Toda-Yamamoto (T-Y) procedure are briefly described below:

[10] The bounds are valid if the residuals are serially uncorrelated. Thus, the lag orders in Equation (3) must be sufficiently long to ensure that is the case. On the other hand, the specified model should be parsimonious. Given these two important considerations, we first chose the optimal lag lengths using the Schwarz Bayesian Information Criterion (SBC) and then tested the residuals for serial correlation using the Breusch-Godfrey LM test. With one exception, we could not reject the null hypothesis of no serial correlation. The exception was the model for total expenditure (*te*) whose residuals were serially correlated at the SBC selected lag orders of (1,1). We thus reestimated the model with lag orders (2, 2) to remove serial correlation.

[11] It should be pointed out that inferring a causal relationship and its direction from Wagner's writings has been questioned by some authors. Peacock and Scott (2006, p. 9), for example, argue that "A more fundamental difficulty arises which is indicated by Wagner's repeated emphasis on the modesty of his claims, remembering that he suggested an association between the growth of G and Y rather than some firm causation. In fact, at times he writes as if the chain of causation could be opposite by the authors, because a *pre-requisite* of economic growth must be growth in infrastructure." See also footnote 7.

[12] See, for example, Worthington and Higgs (2003) for an application.

a. Specify a VAR model in terms of the level of the variables and use a lag-order selection criterion to determine the optimal lag order (*k*).

b. Augment *k*, as determined before, by the *maximal* order of integration, d_{max}, in the system (usually 1 for most economic time series). Then estimate a VAR($k + d_{max}$) model using OLS.

c. Test zero parameter restrictions on the first *k* lags only in the VAR($k + d_{max}$) model (ignoring the rest) using a Wald test. This test statistic has an asymptotic χ^2 distribution under the null hypothesis. The over-parameterized VAR ensures that the asymptotical critical values are applicable regardless of the integration properties of the variables in the system.

For our purpose, the level VAR($k + d_{max}$) for the jth spending category is written as follows:

$$\ln(g_j)_t = \varpi_{j,1} + \sum_{i=1}^{k+d_{\max}} \theta_{j,i} \ln(g_j)_{t-i} + \sum_{i=1}^{k+d_{\max}} \rho_{j,i} \ln(pcpi)_{t-i} + \upsilon_{j,t} \tag{5}$$

$$\ln(pcpi)_t = \varpi_{j,2} + \sum_{i=1}^{k+d_{\max}} \kappa_{j,i} \ln(pcpi)_{t-i} + \sum_{i=1}^{k+d_{\max}} \omega_{j,i} \ln(g_j)_{t-i} + \tau_{j,t} \tag{6}$$

where d_{max}=1 in both equations given that all the variables involved are all I(1). The null hypotheses tested are: $(H_0)_5$: ρ_j s for the first k lags only are jointly equal to zero in Equation (5) and $(H_0)_6$: ω_j s for the first k lags only are jointly equal to zero in Equation (6). If we fail to reject $(H_0)_5$, then the Granger non-causality from *pcpi* to g_j is rejected (stated differently, *pcpi* is said to Granger cause g_j). Similarly, the Granger non-causality from g_j to *pcpi* is rejected if we fail to reject $(H_0)_6$.

As shown in Table 5, there is evidence of a unidirectional causal relationship from *pcpi* to *te*, *ed, ins* and *ssim.* The variables *pf*, *int* and *ut* appear to be independent of *pcpi*. Evidence supporting unidirectional causality from spending to *pcpi* is observed in relation to *fagc* only; although the evidence is relatively weak (the Chi-squared test *p*-value=0.08). Finally, the variable *hy* has a bidirectional causal relationship with *pcpi* based on the results of the Chi-squared test.

Accordingly, the causality direction from income to the relative size of public spending implied by Wagner's Law receives support in relation to SLG total expenditure and expenditures on education, social security and income maintenance (including public welfare and health and hospital), insurance benefits (including employment retirement), and highways. Interestingly enough, the aforementioned four income-driven expenditure sub-categories are the top four in Table 1 in terms of their relative size. The causality direction from public expenditure to income emphasized by Keynesians is observed in relation highway and financial administration and general control expenditure sub-categories. As noted earlier, the statistical evidence in these cases is relatively weak and inconsistent.

IV. Summary and Concluding Remarks

We investigated whether Wagner's Law was supported by the U.S. state-local expenditure data (1957-2006) emphasizing various expenditure types and the concept of cointegration. Our informal analysis of the data indicated that the level of total expenditure and several of its sub-categories, particularly insurance trust benefits and social services and income maintenance grew at rates (significantly) above the rate of growth of aggregate state personal income on an inflation adjusted basis. This observation, which we interpreted as *prima facie* evidence supporting Wagner's Law, is not surprising in view of a shift in public spending emphasis from the more traditional areas such as police and fire protection and highway to "social insurance" expenditures demanded by the electorate (Overbye, 1995) in more mature economies. What is perhaps surprising are the results of our formal analysis which indicated that, with the exception expenditures on insurance trust benefits and social services and income maintenance scaled by personal income (*ins* and *ssim*, respectively), no other nonstationary spending ratio was part of a cointegrating relationship with real per capita personal income (*pcpi*) as implied by Wagner's Law. Other statistical evidence in favor of Wagner's Law was fairly consistent in relation to these two sub-categories: Granger non-causality tests identified a causal ordering from *pcpi* to *ins* and *ssim*. Moreover, both *ins* and *ssim* were income elastic.

The apparent long-term decoupling of some fast growing and nonstationary spending sub-categories (such as interest on debt and financial administration and general control) and *pcpi* is disconcerting, for *pcpi* is also an important indicator of fiscal capacity. It lends credence to public anxiety about runaway government spending and higher attendant budget deficits and debt. From this perspective, the evidence suggesting that SLG total expenditure relative to personal income, while not cointegrated with *pcpi*, was stationary and *ins* and *ssim*, while nonstationary, error-corrected is somewhat reassuring. One may speculate that the error-correcting behavior of *ssmi*, at least to some extent, reflects the change in state funding responsibility for welfare spending subsequent to the welfare reforms introduced in 1996.

Research indicates that the U.S. states "public welfare" expenditures tend to change in countercyclical manner (Dye and McGuire, 2004). However, the finding that such expenditures relative to personal income can not deviate too much from their long-term relationship with *pcpi* suggests that the ability to increase them during the current economic downturn and falling income is somewhat limited. If anything, spending cuts may be necessary especially in states with stringent balance budget rules. In this case, the discretionary "social services" component is likely to bear the brunt these cuts (Gais, 2009). Thus, the injection of additional federal funds in the context of the American Recovery and Reinvestment Act of 2009 may be a welcome relief to many fiscally stressed state and local governments.

References

Abizadeh, S., M. Yousefi (1992). “Wagner’s Law: New Evidence (Anthology).” *Atlantic Economic Journal*, June, 20, 2, p. 100

Bird, R. M. (1971). “Wagner’s law’s of expanding state activity.” *Public Finance*, 26, 1- 26.

Daly, Mary. (2003). “Understanding State Budget Troubles.” FRBSF Economic Letter, The Federal Reserve Bank of San Francisco, August 15..

Dickey, D.A. and W.A. Fuller (1979). “Distribution of the Estimators for Autoregressive Time Series with a Unit Root.” *Journal of the American Statistical Association*, 74, 427–431.

Dye, Richard F. and Therese J. McGuire. (2004). “State Fiscal Systems and Business Cycles: Implications for State Welfare Spending When the Next Recession Occurs.” Assessing the New Federalism; *An Urban Institute Program to Assess Changing Social Policies* 99–4.

Eberts, Randall W. and Timothy J. Gronberg (1992). “Wagner’s Hypothesis: A Local Perspective.” Working paper 9202. The Federal Reserve Bank of Cleveland.

Engle, R. F. and C.W.J. Granger (1987). “Cointegration and error correction: representation, estimation, and testing.” *Econometrica*, 55, 251-76.

Gais, Thomas. (2009). “Stretched Net: The Retrenchment of State and Local Social Welfare Spending Before the Recession.” *Publius:TheJournal of Federalism*, 39 (3), 557-79.

Garand, James C. (1998). “Explaining Government Growth in the U.S. States.” *American Political Science Review*, 82 (3).

Granger, C. W. J. and Newbold, P. (1974). “Spurious Regressions in Econometrics.”*Journal of Econometrics*, 2, 111- 20.

Henrekson, M. (1993). “Wagner’s Law: A Spurious Relationship?” *Public Finance*, 48(3), 406-15.

Islam, Anisul M. (2001). “Wagner’s Law Revisited: Cointegration and Exogeneity Tests for the USA.” *Applied Economics*, 8, 509-515.

Johansen, Søren and Katarina Juselius (1990). “Maximum Likelihood Estimation and Inferences on Cointegration—with applications to the demand for money.” *Oxford Bulletin of Economics and Statistics*, 52, 169–210.

Johansen, Søren, (1991). “Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models.” *Econometrica*, 59, 1551-1580.

Kwiatkowski, Denis, Peter C. B. Phillips, Peter Schmidt & Yongcheol Shin (1992). “Testing the Null Hypothesis of Stationary against the Alternative of a Unit Root.” *Journal of Econometrics*, 54, 159-178.

MacKinnon, James G. (1996). “Numerical distribution functions for unit root and cointegration tests.” *Journal of Applied Econometrics*, *11*, 601–618

MacKinnon, James G., Alfred A. Haug, and Leo Michelis (1999). “Numerical Distribution Functions of Likelihood Ratio Tests for Cointegration.” *Journal of Applied Econometrics*, 14, 563-577.

Musgrave, R. A. (1969). Fiscal Systems, Yale University Press, New Haven and London.

Narayan, P.K. (2005). “The Saving and Investment Nexus for China: Evidence from Cointegration Tests.” *Applied Economics*, 2005, 37, 1979–1990.

Overbye, Einar. (1995). “Explaining welfare spending.” *Public Choice*, 85, 313-35.

Payne James E., Bradley T. Ewing and Hassan Mohammadi (2006). “Wagner’s Hypothesis: New Evidence from Using the Bounds Testing Approach,” in The Elgar Companion to Public Economics, Attiat F. Ott and Richard J. Cebula (eds.) Edgar Elgar.

Peacock, Alan and Alex Scott (2000). “The Curious Attraction of Wagner’s Law.” *Public Choice*, 102, 1-17.

Peacock, Alan (2006). “Wagner’s Law of Increasing Expansion of Public Activities,” in The Elgar Companion to Public Economics, Attiat Ott and Richard J. Cebula (eds.) Edgar Elgar.

Pesaran, M.H., Y. Shin and S. Smith (2001). “Bounds Testing Approach to the Analysis of Level Relationships.” *Journal of Applied Econometrics*, Vol. 16, 289-326.

Phillips, P.C.B. and P. Perron (1988). “Testing for a Unit Root in Time Series Regression.” *Biometrika*, 75, 335–346.

Toda, H. Y. and Yamamoto, T. (1995). “Statistical Inference in Vector Autoregressions with Possibly of Integrated Process.” *Journal of Econometrics*, 66, 225-250.

U.S. Department of Health and Human Services (2004). “Spending on Social Welfare Programs in Rich and Poor States.” http://aspe.hhs.gov/hsp/social-welfare-spending04/report.pdf

Wagner, A. (1893) *Grundlegung der politischen Okonomi*

Worthington, Andrew C. and Helen Higgs (2003). “Art as an investment: Short and long-term comovements in major painting markets.” *Empirical Economics*, 28:649–668

Table 1. Changes in US Total State-Local Expenditure, Expenditure Categories and Personal Income (constant dollars, 1957-2006)

Expenditure	1957 Constant $1000	1957 % total expenditure	1957 % personal income	2006 Constant $1000	2006 % total expenditure	2006 % personal income	Average annual growth rate (1957-2006)
Total expenditure (TE)	339,980,975	100	19.06	1,957,132,296	100	20.80	3.56
• Police and fire protection-direct (PF)	16,290,634	4.79	0.91	88,394,014	4.52	0.94	3.44
• Education-direct (ED)	101,052,363	29.72	5.66	568,280,396	29.04	6.04	3.51
• Utility (UT)	24,981,483	7.35	1.40	132,279,964	6.76	1.41	3.39
• Insurance trust benefits (INS)	19,655,058	5.78	1.10	159,218,007	8.14	1.69	4.27
Unemployment compensation (UC)	10,720,834	3.15	0.60	21,933,902	1.12	0.23	1.44
Employment retirement (ER)	6,739,901	1.93	0.38	121,919,986	6.23	1.30	5.96
• Social services and income maintenance(SSIM)	47,439,501	13.95	2.66	430,827,460	22.01	4.58	4.51
Public welfare-direct (PW)	24,333,738	7.16	1.36	289,090,791	14.77	3.07	5.07
Health and hospital-direct (HH)	22,882,391	6.73	1.28	141,736,669	7.24	1.51	3.71
• Highway-direct (HY)	55,882,484	16.44	3.13	105,707,826	5.40	1.12	1.28
• Financial admin. and general control (FAGC)	12,335,919	3.63	0.69	77,463,946	3.96	0.82	3.74
• Interest on general debt (INT)	7,904,426	2.32	0.44	66,869,852	3.42	0.71	4.36
Personal income (PI)	1,783,884,288			9,409,007,000			3.38

Notes:

Percents do not add up to 100, because not all expenditure categories are presented.
Real expenditure and personal income figures were calculated using the state and local government consumption and investment spending deflator and U.S. GDP deflator, respectively.

The average annual growth rate(AAGR) is calculated using the formula:

$$AAGR = \left[\frac{(value)_{2006}}{(value)_{1957}} \right]^{\frac{1}{50}} - 1$$

Table 2. Unit Root Test Results

	ADF test (H0:Unit Root)		PP test (H0: Unit Root)		KPSS test (Ho: No Unit Root)	Level (non)stationary[d]
Variable [a]	t-statistic	P-value [b]	Adj. t-statistic	P-value [b]	LM-Statistic [c]	
ln(pcpi)						
{1, C, T}	**-1.318**	0.872	**-1.533**	0.805	**0.195****	Nonstationary
{1, C, .}	**-1.751**	0.400	**-1.186**	0.673	**0.931*****	
ln(te)						
{1, C, T}	-3.399	0.063	-3.736	0.029	0.067	Stationary
{1, C, .}	-3.417	0.015	-3.944	0.004	0.139	
ln(ed)						
{1, C, T}	**-2.102**	0.532	**-3.131**	0.111	0.104	Inconclusive
{0, C, .}	**-1.776**	0.388	**-2.072**	0.257	0.340	
ln(fagc)						
{0, C, T}	**-0.136**	0.993	**-0.676**	0.969	**0.192****	Nonstationary
{0, C, .}	**-1.247**	0.647	**-1.075**	0.719	**0.899*****	
ln(ins)						
{0, C, T}	**-3.097**	0.118	-3.360	0.069	0.068	Inconclusive
{1, C, .}	-2.863	0.057	-3.038	0.038	**0.493****	
ln(int)						
{1, C, T}	-3.438	0.058	-3.375	0.067	0.080	Inconclusive
{1, C, .}	**-0.262**	0.923	**0.495**	0.985	0.886	
ln(pf)						
{1, C, T}	**-2.977**	0.149	**-2.921**	0.165	0.059	Inconclusive
{1, C, .}	-3.011	0.041	-2.996	0.042	0.063	
ln(ssim)						
{1, C, T}	**-3.044**	0.132	**-2.254**	0.450	0.076	Nonstationary
{1, C, .}	**-1.566**	0.492	**-1.714**	0.418	**0.803*****	
ln(hy)						
{2, C, T}	**-0.546**	0.978	**-1.065**	0.925	**0.213****	Nonstationary
{2, C, .}	**-2.137**	0.232	**-0.994**	0.749	**0.861*****	
ln(ut)						
{0, C, T}	**-1.299**	0.877	**-1.522**	0.808	0.116	Inconclusive
{0, C, .}	**-1.369**	0.590	**-1.549**	0.501	0.281	

Note:

Bold face test values indicate evidence in favor of the existence of a unit root in the series.
All variables are first-difference stationary or I(1). Results not reported.
a. {L,C,T} represents the specification of the ADF unit root test equation. L denotes the optimal lag length selected by the Schwarz Bayesian Information Criterion (SBC) with a maximum lag length of 5. C and T indicate the presence of a constant and a trend term, respectively, in the three unit root tests.

b. MacKinnon (1996) one-sided p-values.

c. Critical values for the LM statistic are from Kwiatkowski-Phillips-Schmidt-Shin (1992, Table 1)

***,** and * denote significance at the 1, 5 and 10 percent levels, respectively.

Table 3: Johansen Cointegration Test Results

Hypothesized No. of cointegration, r	Eigenvalue	Trace Statistic	Critical Value† (5 percent)	Prob.†	Max Eigen Statistic	Critical Value† (5 percent)	Prob.†	No. of r at 5%	Estimated income Elasticity of Spending (SE)‡	Error-correction Coefficient (SE)
ln(ed) and ln(pcpi)										
None	0.134	8.165	15.495	0.448	6.930	14.265	0.498	0	-0.303	
At most 1	0.025	1.235	3.841	0.266	1.235	3.841	0.266		(0.112)	
ln(ins) and ln(pcpi										
None§	0.439	31.462	15.495	0.001	27.78	14.265	0.002	**1**	**0.169**	**-0.433**
At most 1	0.074	3.683	3.841	0.055	3.683	3.841	0.055		**(0.086)**	**(0.078)**
ln(fagc) and ln(pcpi)										
None	0.108	8.041	15.495	0.461	5.480	14.265	0.608	0	-1.525	
At most 1	0.052	2.561	3.841	0.109	2.561	3.841	0.109		(0.144)	
ln(int) and ln(pcpi)										
None	0.189	11.020	15.495	0.210	10.064	14.265	0.208	0	-1.415	
At most 1	0.020	0.956	3.841	0.328	0.956	3.841	0.328		(0.133)	
ln(pf) and ln(pcpi)										
None	0.203	13.491	15.495	0.098	10.862	14.265	0.161	0	-0.0300	
At most 1	0.053	2.628	3.841	0.105	2.628	3.841	0.105		(0.035)	
ln(ssim) and ln(pcpi)										
None §	0.313	21.537	15.495	0.005	18.007	14.265	0.012	**1**	**0.425**	**-0.180**
At most 1	0.071	3.530	3.841	0.060	3.530	3.841	0.060		**(0.058)**	**(0.050)**
ln(hy) and ln(pcpi)										
None	0.079	6.637	15.495	0.620	3.951	14.265	0.864	0	-0.215	
At most 1	0.054	2.686	3.841	0.101	2.686	3.841	0.101		(0.471)	
ln(ut) and ln(pcpi)										
None	0.077	6.228	15.495	0.669	3.863	14.265	0.874	0	3.575	
At most 1	0.048	2.365	3.841	0.124	2.365	3.841	0.124		(1.789)	

Notes:
In all cases, the cointegration analysis was based on the assumption of a linear deterministic trend in data, an intercept in the cointegrating equation and specification of first-difference lag interval of (1, 1).
§ Denotes rejection of the null at the 0.05 level.
† MacKinnon, Haug and Michelis (1999) p-values.
‡Not strictly valid when r=0

Table 4. Bounds Testing of Cointegration

Variable Pairs	Optimal Lag Length (m)	N*R-squared	Prob. Chi-squared	$F_{(p,q)}$	Inference
ln(te) and ln(pcpi)	2, 2	2.32	0.31	$CV_L<4.26_{(2,40)}<CV_U$	Inconclusive
ln(ed) and ln(pcpi)	3, 3	3.74	0.15	$3.99_{(2,37)}< CV_L$	No cointegration
ln(ins) and ln(pcpi)	1, 1	1.06	0.58	$\mathbf{15.28_{(2,43)}> CV_U}$	**Cointegration**
ln(fagc) and ln(pcpi)	1, 1	2.24	0.33	$CV_L<4.47_{(2,43)}<CV_U$	Inconclusive
ln(int) and ln(pcpi)	1, 1	0.45	0.80	$4.06_{(2,43)}<CV_L$	No cointegration
ln(pf) and ln(pcpi)	1,1	0.75	0.69	$CV_L<4.53_{(2,43)}<CV_U$	Inconclusive
ln(ssim) and ln(pcpi)	2, 2	2.76	0.25	$\mathbf{5.63_{(2,40)} > CV_U}$	**Cointegration**
ln(hy) and ln(pcpi)	2, 2	0.74	0.69	$3.39_{(2,40)} < CV_L$	No cointegration
ln(ut) and ln(pcpi)	1, 1	4.40	0.11	$1.09_{(2,43)}< CV_L$	No cointegration

Notes:

Optimal lag lengths (p) were selected based on the Schwarz Bayesian Information Criterion (SBC)

N*R-squared is the Breusch-Godfrey Serial Correlation LM test statistic. The statistic has an asymptotic chi-square distribution under the null hypothesis of no serial correlation.

Small sample (N=50) lower bound (CV_L) and upper bound (CV_U) critical values at the five percent level are 4.07 and 5.19, respectively (Narayan, 2005, p.1988).

Table 5: Toda-Yamamoto Granger Non Causality Tests

Variable Pairs	Optimal VAR Lag Length (k)[a]	F(k, l)[b]	*p*-value	$(\chi^2_k)^c$	*p*-value	Granger Causality Direction
ln(te) and ln(pcpi)	2	3.03 (2,40)	**0.060**	6.63	**0.036**	*ln(pcpi)* ⇒ *ln(te)*
		1.57 (2,40)	0.220	3.56	0.169	*ln(pcpi)* ← *ln(te)*
ln(ed) and ln(pcpi)	2	2.85 (2,40)	**0.070**	6.26	**0.044**	*ln(pcpi)* ⇒ *ln(ed)*
		1.20 (2,40)	0.287	2.94	0.230	*ln(pcpi)* ← *ln(ed)*
ln(fagc) and ln(pcpi)	2	0.160 (2,40)	0.853	0.37	0.830	*ln(pcpi)* → *ln(fagc)*
		2.196 (2,40)	0.124	4.90	**0.086**	*ln(pcpi)* ⇐ *ln(fagc)*
ln(ins) and ln(pcpi)	2	13.90 (2,40)	**0.000**	24.79	**0.000**	*ln(pcpi)* ⇒ *ln(ins)*
		0.325 (2,40)	0.724	0.758	0.684	*ln(pcpi)* ← *ln(ins)*
ln(int) and ln(pcpi)	2	0.420 (2,40)	0.660	0.980	0.614	*ln(pcpi)* → *ln(int)*
		1.32 (2,40)	0.279	3.00	0.223	*ln(pcpi)* ← *ln(int)*
ln(pf) and ln(pcpi)	2	1.42 (2,40)	0.254	3.22	1.990	*ln(pcpi)* → *ln(pf)*
		1.49 (2,40)	0.237	3.38	0.185	*ln(pcpi)* ← *ln(pf)*
ln(ssim) and ln(pcpi)	2	3.30 (2,40)	**0.047**	7.18	**0.028**	*ln(pcpi)* ⇒ *ln(ssim)*
		1.060 (2,40)	0.355	2.429	0.299	*ln(pcpi)* ← *ln(ssim)*
ln(hy) and ln(pcpi)	3	2.22 (3,37)	0.102	7.62	**0.054**	*ln(pcpi)* ⇒ *ln(hy)*
		1.91 (3,37)	0.144	6.63	**0.085**	*ln(pcpi)* ⇐ *ln(hy)*
ln(ut) and ln(pcpi)	1	0.438 (1,43)	0.511	0.486	0.485	*ln(pcpi)* → *ln(ut)*
		0.022 (1,43)	0.883	0.024	0.875	*ln(pcpi)* ← *ln(ut)*

Notes:
a. Selected based on the Schwarz Bayesian Information Criterion (SBC).
b. Test statistics correspond to the F-test of redundancy of the first k lags of the causal variable in Equations (5) and (6).
c. Test statistics correspond to the log likelihood ratio test of redundancy of the first k lags of the causal variable in Equations (5) and (6).
⇒ indicates rejection of the non-causality null in the direction shown. → indicates failure to reject the non-causality null in the direction shown.