



UNIVERSITY OF GOTHENBURG
SCHOOL OF BUSINESS, ECONOMICS AND LAW

WORKING PAPERS IN ECONOMICS

No 378

**The Tax-Spending Nexus: Evidence from a Panel of US State-
Local Governments**

Joakim Westerlund, Saeid Mahdavi and Fathali Firoozi

September 2009

**ISSN 1403-2473 (print)
ISSN 1403-2465 (online)**

THE TAX-SPENDING NEXUS: EVIDENCE FROM A PANEL OF US STATE-LOCAL GOVERNMENTS*

Joakim Westerlund[†]
University of Gothenburg
Sweden

Saeid Mahdavi
University of Texas at San Antonio
United States

Fathali Firoozi
University of Texas at San Antonio
United States

September 10, 2009

Abstract

We re-examine the tax-spending nexus using a panel of 50 US state-local government units between 1963 and 1997. We find that, unlike tax revenues, expenditures adjust to revert back to a long-term equilibrium relationship. The evidence on the short-term dynamics is also consistent with the tax-and-spend hypothesis. One implication of this finding is that the size of the government at the state-local level is not determined by expenditure demand, but rather by resource supply. This is consistent with the fact that many US state and local governments operate under constitutional or legislative limitations that seek to constrain deficits.

JEL Classification: H71; H72; C33.

Keywords: Tax-spend; State and local government; Public finance; Panel unit root; Panel cointegration.

1 Introduction

Persistently large public sector budget deficits have to be eventually corrected through fiscal adjustments in the form of government expenditure cuts and/or tax revenue increases.

*The first author gratefully acknowledges financial support from the Jan Wallander and Tom Hedelius Foundation, research grant number W2006-0068:1. The second author acknowledges with thanks a research grant by the College of Business of the University of Texas at San Antonio.

[†]Corresponding author: Department of Economics, University of Gothenburg, P. O. Box 640, SE-405 30 Gothenburg, Sweden. Telephone: +46 31 786 5251, Fax: +46 31 786 1043, E-mail address: joakim.westerlund@economics.gu.se.

In practice, however, addressing the deficit problem may be complicated by the several issues. One issue is the division of the burden of adjustment between the expenditure and revenue sides of the budget during periods of fiscal retrenchment. A related issue is the temporal causality between taxes and expenditures which is typically discussed in terms of the following four competing hypotheses in the literature.

According to the tax-and-spend hypothesis championed by Friedman (1978), the level of spending adjusts to the level of tax revenues available. Thus, an increase in tax will not lead to lower budget deficits. Friedman therefore favors a reduction in taxes to force subsequent spending cuts. The Buchanan and Wagner (1977) version of this hypothesis states that tax reductions will lead to higher spending through lowering the perceived price of government provided goods and services by the public. To reduce expenditures, the authors suggest limiting the ability of the government to resort to deficit financing.

The so-called spend-and-tax hypothesis maintains that the level of spending is first determined by the government and then tax policy and revenue are adjusted to accommodate the desired level of spending. In this connection, Peacock and Wiseman (1979) argue that temporary increases in expenditures due to a crisis situation are used to justify higher taxes which may then become permanent. Another version of this hypothesis is based on the work of Barro (1979). In his tax smoothing hypothesis, government spending is considered as an exogenous variable to which taxes adjust. Since changes in expenditures drive changes in taxes in this scenario, the preferred approach to fiscal deficit reduction relies on cutting expenditures.

Meltzer and Richard (1981), among others, maintain that voters' choices lead to concurrent changes in taxes and expenditures. The implication of this so-called fiscal synchronization hypothesis is that causal relationship between government revenue and spending is bidirectional.

In contrast, Wildavsky (1988) and others emphasize that separate institutions participate in the budgetary process and that the collapse of a consensus on fundamentals among them may result in an independent determination of the revenue and expenditure sides of the budget. The implication of this institutional separation hypothesis is that taxes and expenditures may be causally independent.

Our main objective is to re-examine this issue of causality between taxes and expenditures. The paper contributes to the existing tax-spending literature in several ways. Firstly,

our empirical evidence is based on a panel of 50 combined US state and local government units, henceforth referred to as state-local governments, and covers over 35 years.¹ Secondly, our empirical model controls for a number of important factors that are likely to affect the relationship between taxes and expenditures. It is also very general in the sense that it accounts not only for the non-stationarity, but also for the panel structure of our data. Thirdly, our approach to causality relies on the fact that if taxes and expenditures are cointegrated, then their levels must be related in the long run with causality running in at least one direction. To exploit this potential channel of causality, we adopt the panel error correction approach of Westerlund (2007a). Finally, we employ alternative variable definitions to check the robustness of our results.

The rest of the paper proceeds as follows. Section 1 provides a theoretical framework. Section 2 describes the empirical methodology and the data. Section 3 presents the results. Section 4 concludes.

2 The theoretical model

While the direction of causality is an empirical question in the final analysis, the use of state-local data may provide prior expectations in that regard. In particular, it is well known that many states and local governments in the US operate under fiscal constraints in the form of budget requirements and debt limits. These constraints, while not strictly binding, may be effective enough to result in revenue-constrained spending decisions. If so, we would expect to obtain results that are consistent with the tax-and-spend hypothesis. Similarly, to the extent that such constraints create causal dependence between revenues and expenditures in either direction, we do not expect to find empirical support for the institutional separation hypothesis.²

¹For a review of the studies published between 1985 and 2002, see Payne (2003). Only a small subset of these was based on US sub-national data. Of these, many employed aggregate US state or local government level data or a single state time series, see for example Ram (1988), Miller and Russek (1990) and Payne (1998). To the best of our knowledge, the only other study comparable to ours was conducted by Holtz-Eakin *et al.* (1989) who applied a panel vector autoregressive model to 171 US municipal governments over the 1972–1980 period. Controlling for federal grants, their results supported the tax-and-spend hypothesis. A later study by Joulfaian and Mookerjee (1990) applied the same panel approach to annual state level data for sixteen countries during the 1955–1986 period.

²This expectation is buttressed by the fact that the divergence of interests, agendas, and decision-making institutions that tend to decouple spending and tax decisions at the federal level is likely to be less pronounced at the state and local levels, see Hoover and Sheffrin (1992).

With these points in mind, we employ a theoretical framework parallel to Sargent's (1987) treatment of the tax smoothing model of Barro (1979), in which the government decision makers, who are assumed to have rational expectations, take the level of spending, henceforth denoted G_t , as exogenous and choose the level of tax revenue, denoted R_t , to minimize tax distortions. As noted by Hoover and Sheffrin (1992), the roles of taxes and spending can be reversed to derive a model in which the path of government spending is smoothed given the path of taxes. This is the behavioral assumption in the model outlined below. More specifically, suppose the spending distortion at time t has the quadratic form $c_1 G_t + \frac{1}{2} c_2 G_t^2$, where c_1 and c_2 are positive constants. The government then chooses the spending path that minimizes the present expectation of discounted sum of all future distortions,

$$\min_{G_t, B_{t+1}} E_t \left(\sum_{t=0}^{\infty} r^t \left(c_1 G_t + \frac{1}{2} c_2 G_t^2 \right) \right) \quad (1)$$

subject to the budget constraint

$$B_{t+1} = (1 + i)(B_t + G_t - R_t), \quad (2)$$

where E_t is the expectation conditional upon the information available at time t , B_t is the government debt stock, i is the interest rate and r is the discount rate. Note that, as in much of the literature, i and r are assumed to be constant over time and all fiscal variables are expressed in real terms. Following the brief steps shown for parallel problems in Sargent (1987), the first order condition requires

$$E_t(G_{t+1}) = -\frac{c_1}{c_2} \left(1 - \frac{i_o}{r} \right) + \frac{i_o}{r} G_t = -c + \frac{i_o}{r} G_t \quad (3)$$

with $i_o = \frac{1}{1+i}$. Derivations parallel to those in Sargent (1987) yield the following first-order solution for the government spending at time t :

$$G_{it} = \frac{c}{i} + \phi R_t + \delta B_t + \delta \left(\sum_{s=1}^{\infty} i_o^s E_t(R_{t+s}) \right) \quad (4)$$

where $\delta = 1 - \frac{i_o^2}{r}$ and $\phi = i_o \delta$. This equation suggests that spending is determined by the expected present value of all future taxes. Also, since i_o^s converges to zero as s rises, the expected taxes in the immediate future periods have a larger impact on than the expected taxes in the distant future. Following Sargent (1987) and Hoover and Sheffrin (1992), we assume that tax is characterized by the following stochastic process:

$$R_t = \bar{R} + u_t, \quad (5)$$

where \bar{R} is the long-term average tax revenue and u_t is a stationary error term. Note that $E_t(R_{t+s}) = \bar{R}$ for all $s \geq 1$, which can be substituted into (4) to obtain

$$G_t = \frac{c}{i} + \delta \bar{R} \left(\frac{i_o}{1 - i_o} \right) + \phi R_t + \delta B_t = \alpha + \phi R_t + \delta B_t. \quad (6)$$

Note that both δ and ϕ have a positive sign if $i_o^2 < r$, a negative sign if $i_o^2 > r$, and are equal to zero if i_o and r are equal. However, it is typically assumed that $i_o^2 < r$, see for example Sargent (1987, Chapter 6).

3 The empirical model

Based on equation (6), the empirical model that we will consider can be written as

$$G_{it} = \alpha_i + \beta_1 R_{it} + \beta_2' X_{it} + \text{error}, \quad (7)$$

where the index $i = 1, \dots, N$ denote the state-local units, while t again denotes time. Thus, G_{it} is now the spending of state-local government i at time t .

Although we focus on the relationship between G_{it} and R_{it} , these variables cannot be analyzed in isolation. We therefore add X_{it} , a vector of control variables, which includes federal government grants, non-tax revenues, state gross product and of course the debt stock, B_{it} .

A large body of empirical literature has found that grants not only boost the level of spending but do so by an amount which is larger than equal increases in private income.³ On the tax side, grants may create a substitution effect when they replace tax revenues. Accordingly, omission of grants can cause misleading results, for an increase in spending due to an increase in grants may be incorrectly attributed to a change in tax revenues. Non-tax revenues, such as charges and fees, are other sources of funds to state-local governments that have been curiously ignored in much of the empirical literature.⁴ They are expected to have similar qualitative effects on expenditures and tax revenues as grants. Finally, we include a measure of total state output to control for changes in those components of government

³See Hines and Thaler (1995) for a review of the literature.

⁴Data suggest a heavier reliance by state governments on non-tax revenues to finance spending in the past several years. This reflects, among other things, a substitution away from tax revenues, which are constrained by statutory and constitutional limits, and towards non-tax revenues, which are not bound by these limitations, see Skidmore (1999).

spending and taxes that are sensitive to variation in the level of economic activity.⁵

Provided that the variables are integrated of order one and that the regression error is stationary, (7) may be viewed as representing a long-term, or cointegrating, relationship, which can be rewritten as an error correction model. The particular model employed here can be written as

$$\begin{aligned} \Delta G_{it} = & \text{constant} + \rho_{1i}(G_{it-1} - \beta_1 R_{it-1} - \beta_2' X_{it-1}) + \sum_{s=1}^p \delta_{1is} \Delta G_{it-s} + \sum_{s=-p}^p \lambda_{1is} \Delta R_{it-s} \\ & + \sum_{s=-p}^p \gamma'_{1is} \Delta X_{it-s} + \text{error}, \end{aligned} \quad (8)$$

$$\begin{aligned} \Delta R_{it} = & \text{constant} + \rho_{2i}(G_{it-1} - \beta_1 R_{it-1} - \beta_2' X_{it-1}) + \sum_{s=1}^p \delta_{2is} \Delta R_{it-s} + \sum_{s=-p}^p \lambda_{2is} \Delta G_{it-s} \\ & + \sum_{s=-p}^p \gamma'_{2is} \Delta X_{it-s} + \text{error}. \end{aligned} \quad (9)$$

Note that (8) and (9) can be interpreted as two conditional error correction models, one for G_{it} and one for R_{it} . As such, our setup is nothing but a restricted version of the full panel vector error correction model considered by Larsson *et al.* (2001). The idea here is to avoid estimating all the parameters of full model and to make inference based on the conditional models only. In so doing, we assume that the regression errors are independent of ΔX_{it} at all lags and leads. This assumption is not restrictive in the sense that it holds as long as the error correction model in (8) is well specified. If the model is correct, so that all short-run dynamics have been accounted for, then the errors are independent of ΔX_{it} by construction. Apart from this, however, there are basically no restrictions on the two error terms, which may be correlated across both i and t .

The key parameters in (8) and (9) are ρ_{1i} and ρ_{2i} , which measures the extent of the error correction. If $\rho_{1i} < 0$ and/or $\rho_{2i} < 0$, then there is error correction, which implies that G_{it} , R_{it} and X_{it} are cointegrated, whereas if $\rho_{1i} = \rho_{2i} = 0$, then there is no error correction and thus no cointegration. Note that this interpretation of rests on two key assumptions. The first one is that there can be at most one cointegrating relationship, suggesting that the elements of R_{it} and X_{it} cannot be cointegrated among themselves. Although clearly an important assumption, being testable, it is not very restrictive. The second one is that the extent of

⁵In this connection, note that deterioration in the state of the economy can reduce tax revenues and increase some expenditure, at the same time. If output, as the factor that derives both tax revenues and expenditures, is omitted from the estimating equation, the inverse relationship between the two variables may be incorrectly interpreted as support for the Buchanan and Wagner (1977) hypothesis.

cointegration can be inferred by looking at ρ_{1i} and ρ_{2i} alone, which means that X_{it} cannot be error correcting. In other words, the regressors contained in X_{it} must be weakly exogenous with respect to ρ_{1i} and ρ_{2i} . This assumption can be tested by performing a test for error correction in a reverse regression with for example ΔB_{it} as the dependent variable.

Finally, note that weak exogeneity of a variable does not preclude the possibility of dependence between that variable and other variables in the system. To test if a particular variable is strictly exogenous with respect to the other variables in the system we also need to test if the lags and leads of the first differences of the other variables are zero in the regression corresponding to the variable we want to test.

4 Empirical Results

4.1 Data

Our data set consists of a panel of 50 US state-local government units covering the period 1963–1997. The sample period was determined by availability of consistent data on state gross product (see the data appendix for details and data sources). All variables are expressed in log real per capita terms. This obviates the need for adding population as an additional variable to our model to control for changes in taxes and spending that are due to changes in the size of state population.

There are several advantages associated with our data set. First, the fact that the data has a panel structure fills a gap that exists in the empirical literature between studies that have used time series from individual states and those that have used aggregate state or state-local level data. Second, unlike cross-national data, the data from US states enjoy a relatively high degree of homogeneity in dimensions that range from definition and measurement of variables to fiscal and political institutions, processes, and constraints. Third, there is significant degree of variation in the levels of the variables across the state-local government units, which may improve the precision of the estimated parameters of the model. This variation will not be exploited if the cross-sectional units are pooled as in Joulfaian and Mookerjee (1990), or when individual time series are used as in Payne (1998). Fourth, the use of panel data addresses the well-known problem of low power of conventional time series unit root and cointegration tests, as it increases the sample size considerably.

4.2 Unit root tests

We begin the empirical analysis by testing the variables for unit roots, employing the recently developed bootstrap tests of Smith *et al.* (2004). The tests use a sieve sampling scheme to account for error dependence across both the time series and cross-section dimensions of the panel. We consider four tests denoted \bar{t} , \overline{LM} , \overline{max} and \overline{min} , which are all constructed with a unit root under the null hypothesis and heterogeneous autoregressive roots under the alternative. A rejection of the null should therefore be taken as evidence in favor of stationarity for at least one state. The order of the sieve is permitted to increase with T at the rate $4(T/100)^{\frac{2}{9}}$ and so is the lag length of the individual unit root test regressions.⁶ As none of the series seems to be trending, each test regression is fitted with an intercept but no trend. The bootstrapped p -values are based on 1,000 replications.

Table 1: Unit root test results.

Variable	Test values				p -values			
	\bar{t}	\overline{LM}	\overline{max}	\overline{min}	\bar{t}	\overline{LM}	\overline{max}	\overline{min}
Expenditures	4.893	-4.684	9.773	-5.183	1.000	1.000	1.000	1.000
Tax revenues	6.809	-5.746	8.254	-5.645	1.000	1.000	1.000	1.000
Federal grants	-10.496	8.684	20.350	1.198	0.001	0.002	1.000	0.045
Non-tax revenues	5.517	-5.035	14.782	-4.136	1.000	1.000	1.000	0.999
Debt	-2.254	-0.204	9.174	-4.945	0.763	0.847	1.000	1.000
Output	10.304	-6.770	9.557	-5.648	1.000	1.000	1.000	1.000

Notes: The Smith *et al.* (2004) tests take a unit root as the null hypothesis. The test regression is fitted with an intercept and $4(T/100)^{\frac{2}{9}}$ lags. The p -values are based on 1,000 bootstrap replications.

The results reported in Table 1 suggest that the unit root null cannot be rejected at any conventional significance level for any of the variables. The only exception is federal grants, for which the null must be rejected at the 1% level when using the \bar{t} and \overline{LM} tests. However, since the rejections are quite marginal, we chose to proceed as if all six variables are indeed non-stationary.⁷

⁶The idea is that the serial correlation of the data can be approximated arbitrarily well by an autoregressive model of increasing order. To also preserve the cross-sectional dependence, the bootstrap innovations are drawn from the joint cross-sectional distribution on the estimated residuals.

⁷The conclusion that the variables are non-stationary is reinforced by the fact that if we permit for the possibility of a linear trend, the null cannot be rejected for any of the variables.

4.3 Cointegration tests

Given that the variables appear to be non-stationary, we now proceed to test for cointegration. The approach used for this purpose is taken from Westerlund (2007a), who develops four tests based on the error correction models in (8) and (9). All four tests take no error correction as the null hypothesis, but differ in the way the alternative is formulated. Two of the tests, P_α and P_τ , assume that the error correction coefficient of for example equation (8) is equal for all state-local units, in which case the alternative is formulated as that $\rho_{1i} = \rho_1 < 0$ for all i . The second pair, G_α and G_τ , do not require ρ_{1i} to be equal, which means that the alternative is formulated as that $\rho_{1i} < 0$ for at least some i . Thus, while a rejection by the first two tests provides evidence in favor of cointegration for all states, this is not the case for the other two. Similar to the Smith *et al.* (2004) unit root tests, the error correction tests use a sieve type sampling scheme that accounts for both the time series and cross-sectional dependencies of the regression error.⁸

Table 2: Cointegration test results.

Test	Expenditures		Tax revenues	
	Value	<i>p</i> -value	Value	<i>p</i> -value
G_τ	-19.958	0.002	-10.189	0.412
G_α	-1.835	0.367	0.330	0.987
P_τ	-10.305	0.027	-8.794	0.241
P_α	-4.808	0.061	-3.470	0.534

Notes: The Westerlund (2007a) tests take no cointegration as the null hypothesis. The test regression is fitted with an intercept and $4(T/100)^{\frac{2}{5}}$ lags and leads. The *p*-values are based 1,000 bootstrap replications.

The computed values of the test statistics are presented in Table 2 along with the bootstrapped *p*-values based on 1,000 replications. We begin by examining the results from equation (8) with expenditures as the dependent variable. As can be seen, except for G_α , the no cointegration null is rejected at least at the 10% level, which we take as evidence in favor of cointegration. There is no difference depending on whether ρ_{1i} is restricted to be homogeneous or not, suggesting that the whole panel is cointegrated. The fact that G_α has such a large *p*-value is strange, but consistent with its relatively poor power properties in small samples,

⁸As with the Smith *et al.* (2004) tests, we set the order of the sieve approximation equal to $4(T/100)^{\frac{2}{5}}$. The same rule is used for selecting *p*, the number of lags and leads used in the estimation of (8) and (9).

as documented by Westerlund (2007a). We therefore choose to interpret these results as evidence in favor of cointegration.⁹

As pointed out by Westerlund (2007a), violations of the assumption of weakly exogenous regressors are only problematic to the extent that the tests are unable to reject the null of no cointegration, in which case we do not know whether there is no cointegration at all, or if there is cointegration, but it is only R_{it} or X_{it} that are error correcting. In other words, our finding of cointegration is not going to be altered even if some of the regressors happen to be non-weakly exogenous. Nonetheless, in order to shed at least some light on the appropriateness of this assumption, we performed a series of reverse regression tests. As explained in Section 3, if regressors in X_{it} are indeed weakly exogenous, then they should not be error correcting, and this is exactly what we find. In fact even tax revenues seem to pass the weak exogeneity test. This is shown in rightmost panel of Table 2, which reports the results from (9) with tax revenues as the dependent variable. Note that, consistent with the notion of weak exogeneity, the null of no error correction cannot be rejected. In other words, there seem to be no serious violations of the weak exogeneity assumption.

Finally, to test the validity of the assumption that the regressors in (8) cannot be cointegrated amongst themselves, we tested the rank of (R_{it}, X_{it}) using the trace test of Johansen (1988). The results indicate that in only six out of the 50 cases do we end up rejecting the null hypothesis of full rank at the 1% significance level, which means that the regressors can be considered as roughly non-cointegrated.¹⁰ Similar results were obtained for the regression in (9).

4.4 Cointegration estimation

It is well known that the presence of endogeneity and cross-sectional dependence makes the least squares estimator inefficient and biased. A common approach to alleviate this problem is to use seemingly unrelated regressions techniques. However, since this approach is not feasible when $N > T$, in this paper we instead apply the newly developed estimator of Westerlund (2007b), which is based on modelling the cross-sectional dependence by means of a small number of common factors. The estimator, which can be seen as a factor augmented

⁹Note that in addition to being more powerful than conventional time series tests the panel tests applied here have a great operational advantage in that they do not require tabulation and evaluation of the individual tests, which is not practical in the typical panel where N is relatively large.

¹⁰Note that with 50 states, we expect the full rank null to be rejected a certain number of times just by chance.

version of the more conventional bias-adjusted estimator of Kao and Chiang (2000), is implemented in two steps. The first step involves estimating the common factors using the method of principal components.¹¹ In the second step, the cointegration vector is estimated by least squares conditional upon the resulting first-step factor estimates.

Table 3: Cointegration estimation results.

Variable	LS			Bias-adjusted LS		
	β	SE	p -value	β	SE	p -value
Tax revenues	0.498	0.006	0.000	0.524	0.018	0.000
Federal grants	0.076	0.002	0.000	0.058	0.007	0.000
Non-tax revenues	0.032	0.004	0.000	0.040	0.015	0.006
Debt	0.016	0.002	0.000	0.031	0.006	0.000
Output	-0.030	0.007	0.000	-0.076	0.020	0.000

Notes: The value β refers to the estimated cointegrating slope, SE refers to the Newey and West (1994) robust standard error and LS refers to the least squares estimator. The bias-adjusted LS estimator is that of Westerlund (2007b). The results are based on an intercept and the p -values are for a double-sided test of a zero slope.

For comparison, the bias-adjusted estimation results are reported along with their unadjusted least squares counterparts in Table 3. It is seen that both estimators produce very similar results, and that all five right-hand side variables are highly significant. Note also the positive sign of the estimated slope coefficients of the first four variables, which corroborates the notion that expenditures at the state level are resource constrained. The positive sign of the tax variable is of particular interest as it provides support in favor of the tax-and-spend hypothesis.

The fact that the standard errors of the bias-adjusted estimator are larger than those of the least squares estimator can be due to computational differences, but it can also be due to the least squares bias in the presence of cross-section dependence.

4.5 Exogeneity tests

We have already established that R_{it} and X_{it} appear to be weakly exogenous. To determine whether they are also strictly exogenous, we now proceed to test the significance of the first

¹¹The number of common factors is determined using the IC_1 information criterion recommended by Bai and Ng (2004). The maximum number of factors was set to five but the IC_1 criterion suggested that four factors should be enough to model the cross-section dependence.

differenced variables. Since these variables are stationary, the exogeneity test is implemented as an ordinary F -test of the null hypothesis that the lags and leads of each element in (8) and (9) are jointly zero.¹² The problem is that there is not just one, but N regressions to consider for each choice of dependent variable. To facilitate inference at the overall panel level, we propose combining the p -values of the individual F -tests, henceforth denoted as p_i , in the following way:

$$P_m = -\frac{1}{\sqrt{N}} \sum_{i=1}^N (\ln(p_i) + 1).$$

As shown by Choi (2001), given that the individual tests are independent across i , then P_m converges to the standard normal distribution as N grows large. As already noted, however, the assumption of cross-sectional independence is unlikely to hold in our data. To allow for violations of this assumption, we further propose bootstrapping the individual F -tests under the null hypothesis of short-run exogeneity. The resulting p -values can then be used in place of p_i in the formula above, and P_m should again converge to the standard normal distribution.

Table 4: Short-run exogeneity test results.

Variable	Expenditures		Tax revenues	
	P_m	p -value	P_m	p -value
Tax revenues/Expenditures	2.991	0.001	2.052	0.020
Federal grants	1.598	0.055	0.679	0.249
Non-tax revenues	3.704	0.000	-0.341	0.633
Debt	1.948	0.026	1.255	0.105
Output	2.686	0.004	5.621	0.000

Notes: The value P_m refers to the p -value test based on the individual F -tests for short-run exogeneity. The test regression is fitted with an intercept and $4(T/100)^{\frac{2}{9}}$ lags and leads. The p -values are based on the normal distribution.

The results from are reported in Table 4 and may be summarized as follows. Firstly, we see that the first differences of all five explanatory variables enter (8) significantly, at least at the 10% level. Thus, expenditures react not only to deviations from the long-run relationship, but also to short-run movements in the rest of the system, including tax revenues. At the

¹²Although insignificant for R_{it} and X_{it} , for simplicity all test statistics are computed with an unrestricted error correction term.

1% level, however, only tax revenues, non-tax revenues, and output pass the short-term exogeneity test.

Secondly, if we look at the tax revenue equation, then there are only two significant variables at the 10% level, expenditures and output. At the 1% level only output remain significant. Thus, since the error correction term was also insignificant in (9), tax revenues may be considered as roughly strictly exogenous.

Finally, based on the combined results of the effects of the short-run dynamics, we cannot reject the hypothesis that there is a bidirectional relationship between expenditures and tax revenues at the more conventional 5% level of significance. These results are consistent with the findings in favor of the tax-and-spend hypothesis at the subnational level using different data sets and level of aggregation reported by Holtz-Eakin *et al.* (1989), and Payne (1998), as well as those favoring the fiscal synchronization hypothesis reported by Miller and Russek (1990) among others.

4.6 Robustness checks

To check the robustness of our initial results, we first dropped Alaska and Hawaii from the original sample, because they are not continental states and/or have an atypical fiscal structure making them outliers. Also, since there is no consensus about the measures of taxes and spending variables in the empirical literature (see for example Baghestani and McNown, 1994), we repeated our analysis using the full sample and replacing the real per capita variables with real total variables and nominal total variables scaled by state output.

Our original findings seem to be quite robust to the change in the composition of the sample and to the use of alternative variable definitions. In particular, based on the bootstrapped p -values, the evidence of error correction was consistently stronger for expenditures than for tax revenues. In fact, the coefficient of the error correction term in the tax revenue equation was not even statistically significant in the sample of the 48 contiguous states. The evidence consistent with bidirectional causality was also confirmed by the new results.¹³

5 Concluding remarks

In this paper, we re-examined the tax-spending nexus using, for the first time, a panel of 50 US states-local government units over a period of roughly three and a half decades. The

¹³All results discussed in this section are available from the corresponding author upon request.

statistical evidence suggests that while taxes are rather exogenously set, expenditures adjust not only to deviations from the long-term equilibrium relationship but also to the short-run changes in taxes, other funding sources and output. Stated differently, expenditures seem to bear the adjustment burden in response to budgetary disequilibria. An implication of this finding is that the size of the government at the state-local level is not determined by expenditure demand, but rather by resource supply, such as taxes and grants. It is hard not to conclude that these results, at least in part, reflect the constitutional or legislative limitations that seek to constrain deficits under which many state and local governments operate in the US. These include submission of balanced budgets, limiting appropriations to estimated revenues, and/or requiring revenue shortfalls to be matched by spending cuts.

That expenditures seem to depend on taxes both in the long and short terms underscores the important role of taxes in controlling government deficits at the state-local level. In this connection, reductions in the federal commitment to existing entitlement and mandatory programs and/or introduction of new unfunded mandates will result in fiscal imbalances through cost shifts to state and local governments. To avoid confronting these governments with the unpleasant choices of raising taxes or cutting other expenditures, stricter adherence to provisions of the Federal Unfunded Mandates Reform Act of 1995 is necessary.

References

- Baghestani, H., R. McNown. (1994). Do Revenue or Expenditures Respond to Budgetary Disequilibria? *Southern Economic Journal* **61**, 311–322.
- Bai, J., S. Ng (2004). A Panic Attack on Unit Roots and Cointegration. *Econometrica* **72**, 1127–1177.
- Barro, R. J. (1979). On the Determination of the Public Debt. *Journal of Political Economy* **81**, 940–971.
- Buchanan, J. M., R. W. Wegner. (1977). *Democracy in Deficit*. Academic Press, New York.
- Choi, I. (2001). Unit Root Tests for Panel Data. *Journal of International Money and Finance* **20**, 249–272.
- Friedman, M. (1978). The Limitations of Tax Limitation. *Policy Review* **5**, 7–14.
- Johansen, S. (1988) Statistical Analysis of Cointegration Vectors. *Journal of Economic Dynamics and Control* **12**, 231-254
- Hines, J. R., R. H. Thaler (1995). The Flaypaper Effect. *Journal of Economic Perspectives* **9**, 217–26.
- Holtz-Eakin, D., W. Newey, H. S. Rosen (1989). The Revenues-Expenditures Nexus: Evidence from Local Government Data. *International Economic Review* **30**, 415–429.
- Hoover, K. D., S. M. Sheffrin (1992). Causation, Spending, and Taxes: Sand in the Sandbox or Tax Collector for the Welfare State? *The American Economic Review* **82**, 225–248.
- Joulfaian, D., R. Mookerjee (1990). The Intertemporal Relationship Between State and Local Government Revenues and Expenditures: Evidence from OECD Countries. *Public Finance* **45**, 109–117.
- Kao, C., M.-H. Chiang. (2000). On Estimation and Inference of a Cointegrated Regression in Panel Data. *Advances in Econometrics* **15**, 179–222.
- Larsson, R., J. Lyhagen, M. Lötgren (2001). Likelihood-Based Cointegration Test in Heterogeneous Panels. *Econometrics Journal* **4**, 109–142.

- Meltzer, A. H., S. P. Richard (1981). A Rational Theory of the Size of Government. *Journal of Political Economy* **89**, 914–927.
- Miller, S. M., F. S. Russek (1990). Co-Integration and Error-Correction Models: The Temporal Causality between Government Taxes and Spending. *Southern Economic Journal* **57**, 221–229.
- Newey, W., K. West (1994). Autocovariance Lag Selection in Covariance Matrix Estimation. *Review of Economic Studies* **61**, 631–653.
- Payne, J. E. (1998). The Tax-Spend Debate: Time Series Evidence from State Budgets. *Public Choice* **95**, 307–320.
- Payne, J. E. (2003). A Survey of the International Evidence on the Tax-Spend Debate. *Public Finance Review* **31**, 302–324.
- Peacock, S. M., J. Wiseman (1979). Approaches to the Analysis of Government Expenditures Growth. *Public Finance Quarterly* **7**, 3–23.
- Ram, R. (1988). Additional Evidence on Causality between Government Revenue and Government Expenditure. *Southern Economic Journal* **54**, 763–769.
- Sargent, T. J. (1987). *Macroeconomic Theory*. Academic Press.
- Skidmore, M. (1999). Tax and Expenditure Limitations and Fiscal Relationships between State and Local Governments. *Public Choice* **99**, 77–102.
- Smith, V. L., S. Leybourne, T-H. Kim, P. Newbold (2004). More Powerful Panel Unit Root Tests with an Application to the Mean Reversion in Real Exchange Rates. *Journal of Applied Econometrics* **19**, 147–170.
- Wildavsky, A. (1988). *The New Politics of Budgetary Process*. Glenview, Scott Foresman.
- Westerlund, J. (2007a). Testing for Error Correction in Panel Data. *Oxford Bulletin of Economics and Statistics* **69**, 709–748.
- Westerlund, J. (2007b). Unbiased Estimation of Cointegrated Panel Regressions with Cross Section Dependence. *Journal of Financial Econometrics* **5**, 491–522.

Data appendix

The data for state-local expenditures, tax revenues, non-tax revenues, total debt outstanding, and federal intergovernmental revenues (grants-in-aid to state and local governments) were extracted from the computer files provided by the US Census Bureau, <http://www.census.gov/>.

The data for state gross product were obtained from the US Department of Commerce, Bureau of Economic Analysis, see <http://www.bea.gov/regional/gsp/>. The series used were computed based on the Standard Industrial Classification industry definitions which covered the years 1963 to 1997. The Bureau of Economic Analysis switched to the North American Industrial Classification System industry definitions after 1997 and renamed the series as gross domestic product. This change created discontinuity in the data due to differences in source data and different estimation methodologies. Since the Bureau of Economic Analysis explicitly cautioned against appending the two series, the end year of our sample was 1997.

In the absence of an appropriate state-specific deflator we used a price index for aggregate state-local government consumption expenditures and gross investment to convert nominal values into real values. The source of the data for this series is the US Bureau of Economic Analysis, see <http://www.bea.gov/national/nipaweb>.