

THE UNIVERSITY OF TEXAS AT SAN ANTONIO, COLLEGE OF BUSINESS

Working Paper SERIES

Date October 7, 2010

WP # 0016ECO-SAEIDMAHDAVI-2010

Fiscal Stringency and Fiscal Sustainability in the American States: Panel Evidence

Saeid Mahdavi
University of Texas at San Antonio
United States

Joakim Westerlund
University of Gothenburg
Sweden

Copyright © 2010, by the author(s). Please do not quote, cite, or reproduce without permission from the author(s).

Fiscal Stringency and Fiscal Sustainability in the American States: Panel Evidence

Saeid Mahdavi†
University of Texas at San Antonio
United States

Joakim Westerlund
University of Gothenburg
Sweden

October 30, 2010

Abstract

Unlike the federal government, most state governments in the U.S. formally operate under statutory or constitutional constraints which limit their ability to run budget deficits and resort to debt financing. A priori, one would expect to find evidence in favor of an intertemporally balanced budget, or fiscal sustainability, among states, especially those that are characterized by a high degree of fiscal stringency. We test this hypothesis in a panel of 47 contiguous states over the period 1961-2006 using four budget balance definitions and subsamples defined on the basis of whether certain balance budget requirements (BBRs) are in place. Our results, obtained from panel estimation techniques that allow for cross-state dependence, suggest that a sufficient condition for “strong” fiscal sustainability is satisfied in most cases. However, we do not find conclusive evidence that in these cases strong (weak) sustainability is due to the presence (absence) of BBRs.

JEL Classifications: H71; H72; C33 .

Keywords: Fiscal sustainability; Fiscal stringency; budget rules; States budget; Public finance;

Panel cointegration.

† Corresponding author. E-mail address: smahdavi@utsa.edu

1. Introduction

Like the federal government, state government revenues and expenditures are affected by cyclical fluctuations in the level of economic activity as well as long-term factors such as structural changes in the economy and changes in citizens' preferences for the public sector provided goods and services. State government budgets, however, are influenced by two additional factors: First, (unfunded) federal spending mandates and, second, statutory or constitutional balanced budget requirements or rules (BBRs) and debt limits under which most states formally operate. Thus, while revenue losses and expenditure increases during economic downturns may lead to large state budget deficits in the short term, BBRs will presumably force subsequent fiscal adjustments to restore the fiscal balance in the long term. If so, then the state budget is expected to be *intertemporally* balanced, or sustainable.

Our review of the literature identifies two major strands of research on this subject. The first strand builds on a study conducted by the US Advisory Commission on Intergovernmental Relations, or ACIR (1987). A common objective of the studies of this strand (see, for example, vonHagen, 1991; Alt and Lowry, 1994; Poterba, 1994; Bohn and Inman, 1996; Endersby and Towle, 1997) is to determine whether BBRs matter in relation to state budget deficits. The main conclusion drawn based on the preponderance of the evidence presented in these studies is that BBRs indeed make a difference. In particular, fiscal adjustments in the form of tax increases and/or spending cuts were found to be larger or quicker, and debts were lower in states with relatively stringent anti-deficit rules, especially no-deficit-carry-over rules when they were accompanied by debt limits. A survey of state constitutions and statutes by Hou and Smith (2006) significantly expands the catalog of state BBRs by identifying several new “technical provisions” in different phases of state budget cycle. More recently, Smith and Hou (2008) and Hou and Smith (2009) present evidence indicating that relatively straightforward and rigid “technical provisions” are more effective in restraining state spending and deficits than “political provisions” that are ambiguous and more easily subject to manipulation.

Several closely related studies examine whether the fiscal policy restrictions imposed by BBRs impede fiscal flexibility and, thus, contribute to state output volatility (see, for example, Bayoumi and Eichengreen, 1995; Alesina and Bayoumi, 1996; Levinson, 1998; Sorenson *et al.*, 2001; Rose, 2005; Fatás and Mihov, 2006; Krol and Svorny, 2007). These studies yield mixed results on the existence of a trade-off between fiscal flexibility and output volatility.

The second strand focuses on deriving the conditions for fiscal sustainability using an intertemporal budget constraint (IBC) framework. The empirical tests of these conditions usually require examining the integration and cointegration properties of fiscal variables such as budget deficit, debt, revenue and expenditure. Almost all U.S. studies in this line of research use time series data on the federal government fiscal variables (see, for example, Hakkio and Rush, 1991; Trehan and Walsh, 1991; McDonald, 1992; Tanner and Liu, 1994; Ahmed and Rogers, 1995; Quintos, 1995; Haug, 1995; Martin, 2000). With a few exceptions, their results support the sustainability of fiscal policy at the federal level. Bohn (1998) proposes an alternative test of fiscal sustainability based on whether the reaction of primary surplus to the stock of debt (both expressed as a ratio of output) is positive. His results also support federal fiscal sustainability.¹

This paper synthesizes some aspects the two strands of research reviewed above to explore whether (a) state government revenue and expenditure variables behave in a way that satisfies a sufficient condition for fiscal sustainability, and (b) the evidence in favor such behavior is more pronounced in states that are characterized by a relatively high degree of fiscal stringency. Several features of our empirical analysis are worth noting. Firstly, we use a panel data set that includes 47 contiguous states over the period 1961-2006. The panel data, through combining the information from the cross-section and time series dimensions, address the problem of the low power of statistical tests that is common to many previous studies that tested sustainability based on time series data. Secondly, we employ some recently developed panel estimation techniques that allow for cross-sectional dependence and heterogeneity among states.

¹ Also, see Goyal *et al* (2004) and Kia (2008) for some recent country level studies.

This is crucial to drawing correct statistical inferences based on a sample such as ours, because states may experience common shocks and each state's fiscal condition is likely to be affected by the fiscal conditions of its neighboring states. Thirdly, in addition to testing for sustainability in subsamples that are defined based on different indicators of fiscal stringency, we use alternative definitions of state revenue and expenditure (or fiscal balance) to assess the sensitivity of fiscal sustainability to the definition of the fiscal deficit (surplus).

To our knowledge, this is the first study that empirically investigates the question of fiscal sustainability in a panel of American states and explores fiscal sustainability-fiscal stringency linkage. Our results may shed light on the state governments' long-term fiscal health which may be masked by the budgetary imbalances they experience during economic downturns. They can also be informative as to whether BBRs matter in this connection. The rest of this paper proceeds as follows. Section 2 presents an outline of the IBC framework for testing fiscal sustainability. Section 3 briefly discusses various political and technical components of state BBRs. Section 4 describes the data, empirical methodology and presents the results. The last section summarizes our main findings and discusses their implications.

1. Fiscal Sustainability

Generally speaking, fiscal sustainability requires that the government budget to be intertemporally balanced. A brief outline of a formal framework for deriving sufficient condition for sustainability follows.² The starting point is the following one-period budget constraint:

$$B_{it} - B_{it-1}(1+r) = G_{it} - R_{it} \quad (1)$$

where R_{it} is government revenue for the i^{th} state in period t , G_{it} is government expenditure inclusive of interest payments, B_{it} is the stock of public debt and r is the mean of the real interest rate on that debt, r_{it} . By using forward substitution and then taking first-differences, (1) becomes

² See Quintos (1995) for a more detailed discussion.

$$\Delta B_{it} = G_{it} - R_{it} = \sum_{s=0}^{\infty} \left(\frac{1}{1+r} \right)^{s+1} [(\Delta R_{it+s} - \Delta G_{it+s}) - \Delta(r_{it+s} - r)B_{it+s-1}]$$

$$+ \lim_{s \rightarrow \infty} \left(\frac{1}{1+r} \right)^{s+1} \Delta B_{it+s} \quad (2)$$

If the government obeys its intertemporal budget constraint, then the expected present value of future primary surpluses must be equal to the current value of debt, which implies

$$\lim_{s \rightarrow \infty} \left(\frac{1}{1+r} \right)^{s+1} E_t(\Delta B_{it+s}) = 0, \quad (3)$$

where E_t is the expectation operator conditional on the information available at time t . The most common way to test the sustainability hypothesis has been to test whether debt is stationary, or I(0), in its first-differences. Alternatively, this can be implemented as a test for cointegration in the following regression:

$$R_{it} = \alpha_i + \beta G_{it} + \varepsilon_{it}, \quad (4)$$

where ε_{it} is a mean zero error term. This, together with (2), implies that the first-differenced debt can be written as

$$\Delta B_{it} = G_{it} - R_{it} = (1 - \beta)G_{it} - \alpha_i - \varepsilon_{it}. \quad (5)$$

Quintos (1995) assumes that debt is integrated of at most order two, or I(2), in which case the sustainability can be of two types:

1. If ΔB_{it} is I(0), then the sustainability is said to be “strong.” Equation (5) implies that for this to hold β must be equal to one *and* ε_{it} must be I(0). Hence, in this case debt is I(1), and G_{it} and R_{it} are cointegrated and one-to-one.
2. Strong sustainability is consistent with the idea that deficits cannot be persistent. However, as Quintos (1995) shows, sustainability holds even if ΔB_{it} is I(1), so that debt is I(2). But since the government is now spending more than it receives, it will

eventually run into difficulties in marketing its debt, and the sustainability is therefore said to be “weak.” Thus, the only requirement here is that $0 < \beta < 1$.

The above sustainability conditions are based on the assumption that debt is at most $I(2)$. While empirically rather irrelevant, in theory it is, of course, possible to conceive cases in which debt is integrated of even higher order. In this connection, Bohn (2007) shows that sustainability holds if debt is integrated of any finite order. In terms of the terminology of Quintos (1995), this type of sustainability might be referred to as “absurdly weak.”

3. State Balanced Budget Requirements and Fiscal Stringency

While virtually all states have some form of BBRs, the requirements are not equally stringent across states. To assess the effects of variations in fiscal constraints on fiscal sustainability, we identify subsamples on the basis of alternative indicators of the degree of fiscal stringency. A frequently employed summary measure is the one constructed by the ACIR (1987). ACIR defines five categories of BBRs (henceforth referred to as ACIR-BBRs) and assigns points to each category depending on the perceived degree fiscal stringency as follows:

1. The governor has to submit a balanced budget (1 point);
2. the legislature has to pass a balanced budget (2 points);
3. the state may carry over a deficit but must correct it in the subsequent budget period (4 points);
4. the state may not carry over a deficit into the next budget period (6 points);
5. the state may not carry over a deficit into the next fiscal year (8 points).

For each state, the ACIR fiscal stringency index (ACIRFSI) is equal to the state’s point(s) from its highest ranked category plus one point if the budget requirement is statutory, or two points if it is constitutional. The maximum value for the index is 10 and its minimum value is 0.

In their critical review of the state BBR literature, Hou and Smith (2006) point out that the popular ACIRFSI and similar summary measures “are grounded in data that at least partially

reflect personal perception, judicial interpretation, and other nonstatutory and nonconstitutional considerations” (p. 27). The authors further emphasize that “BBRs are systems that are best categorized in a framework of interrelated rules of a political and/or technical nature governing the executive preparation, legislative review, and implementation phase of the budget cycle” (p. 27). Based on their in-depth survey of state constitutions and statutes the authors identify nine rules and taxonomized them along a “political-technical” continuum. Political rules govern the budgetary procedure while technical rules are more relevant to the substance of the budgetary process. Moreover, unlike technical rules, political rules are relatively easier to circumvent and manipulate and are, thus, more ambiguous. For this reason, the authors argue that political rules must be substantiated by technical rules to be effective in sealing any possible leakage.

The nine political and technical BBRs identified in the survey of Hou and Smith (2006) (hereafter referred to as HS-BBRs) are as follows:

1. Governor must submit a balanced budget (political).
2. Own-source revenue must match (meet or exceed) expenditure (technical).
3. Own-source revenue and general obligation (or unspecified) debt (or debt in anticipation of revenue) must match (meet or exceed) expenditures (technical).
4. Legislature must pass a balanced budget (political).
5. A limit is in place on the amount of debt that may be assumed for purpose of deficit reduction (technical).
6. Governor must sign a balanced budget (political).
7. Controls are in place for supplementary appropriations (technical).
8. Within fiscal-year controls are in place to avoid deficit (technical).
9. No deficit may be carried over to the next fiscal year or biennium (technical).

Note that HS-BBRs include four new technical rules, but have three political rules (the submission, passage, and signing of a balanced budget) and one technical rule (no deficit carryover) in common with ACIR-BBRs. In contrast to ACIR-BBRs, however, HS-BBRs do not

make a distinction between types of no-deficit-carryover requirement depending on the length of the budget cycle, nor do they give weight to the constitutional-statutory distinction to infer the degree of fiscal stringency associated with BBRs. Finally, there are discrepancies in the two sets of BBRs with respect to the number of states cataloged under some provisions (Rule 6, for example). Perhaps the most significant discrepancy concerns the state of Vermont, which has been long considered in the literature as the only one with no BBRs. However, according to HS-BBRs, Vermont should be replaced by North Dakota.

Table 1 provides a summary of state BBR systems using the data from Hou and Smith (2005) and ACIR (1987). In Columns 2-9 of the table, we use a binary variable that takes the value one for a state if there is a provision for the corresponding BBR in the state's constitution and/or statutes and zero otherwise. In the last two column of the table, we record the values of ACIRFSI and a newly constructed variable SUM789 (to be discussed later), respectively.

Several points are worth emphasizing in relation to the BBRs described above. Firstly, as noted earlier, BBRs are prevalent but substantially differ in terms of scope and nature across the states. For instance, while 40 states require submission of a balanced budget by the governor (the weakest anti-deficit rule) only two explicitly require the governor to sign a balanced budget and only eight have an explicit “no-deficit-carry-over” provision (the most strict anti-deficit rule). Secondly, some BBRs may not be strictly binding due to lack of a formal enforcement mechanism, the use of “creative accounting” and “gimmicks” by states.³ Thirdly, the definitions of revenues and expenditures (deficits) subject to BBRs are not uniform across states. However, it is commonly understood that the BBRs do not generally apply to capital expenditure and special (trust) funds (Snell, 2004). On the revenue side, as noted by Hou and Smith (2006), several states

³ These include changing the timing of expenditures and revenues, transferring of funds among accounts (“fund sweeping”), issuing debts through state agencies, and shifting on budget items to off budget (see, for example, Briffault, 1996). Relative to spending cuts and tax increases, however, the effects of these accounting changes on deficit reductions are fairly small (Poterba, 1996).

have made their BBR systems more flexible by allowing the submission of balanced budgets that include “other resources.”

4. Data and Empirical Methodology

4.1 Data

We collected data on several alternative definitions of state expenditure and revenue variables to conduct tests of fiscal sustainability for the American states. Our sample includes the data for 47 contiguous states for the period 1961-2006.⁴ As noted before, the exact set of expenditure categories that are subject to BBRs is rather unclear. Moreover, a particular BBR may not equally apply to all expenditure categories (for example, borrowing restrictions for capital expenditures may be much less strict than those for current expenditures). For these reasons, we define the following four state budget balances (ranging roughly from broad to narrow) to estimate the slope parameter in the regression in (4).

- B1. Total revenues less total expenditures: The revenue and expenditure measures included in this balance definition are all encompassing. As such, the possibility of circumventing BBRs through “fund sweeping” is nonexistent for the whole balance, although the possibility of “off-budgeting” and borrowing cannot be ruled out.
- B2. General revenues less general expenditures: This balance definition is narrower compared with B2 in that it excludes the revenue and expenditure items related to insurance trust funds, utilities and liquor stores. Since these items are not part of the official budget they provide an opportunity for states to shift deficits to them.⁵
- B3. General revenues less current (non-capital) general expenditures: This balance definition is the same as B2 except that it excludes capital expenditures from the

⁴ Following Bohn and Inman (1996) and others, we exclude Alaska, Hawaii, and Wyoming whose “unique” fiscal structures make them potential outliers.

⁵ Sorensen *et al.* (2001), however, do not find statistical evidence that states with more stringent budget rules shift the deficit from the general budget to pension funds to circumvent BBRs.

spending side of B2. Capital spending is typically outside the state operating budgets and it is financed by issuing debt. Where BBRs directly affect capital spending, it may be targeted for significant reduction, especially when a portion of it has to be paid out of current revenues, for it is less politically costly to cut.

- B4. Own general revenues less current (non-capital) general expenditures: This balance further excludes inter governmental revenues (mainly federal grants-in-aid) from the revenue side of B3. Federal grants tend to change counter cyclically with respect to state output fluctuations. Their inclusion in any balance makes the size of the associated surplus (deficit) larger (smaller) than may have otherwise been the case.⁶

In summary, the *possibility* of fund sweeping and the proportion of the funds subject to BBRs tend to generally increase as we move from B1 to B4.

We apply the cointegration tests of fiscal sustainability to the full sample as well as subsamples defined based on the value of ACIRFSI and the binary variables corresponding to HS-BBRs shown in Table 1. For reasons discussed earlier, the criteria for splitting the full sample emphasize several technical provisions. They are briefly explained below.

- C1. This criterion emphasizes whether a state has a no-deficit-carry-over provision (annual or biennium) in its *constitution*. This provision is fairly rigid and the most difficult to circumvent. Thus, we define two subsamples with relatively “tight” or “soft” BRRs, using $ACIRFSI = 10$ and $ACIRFSI < 10$ values, respectively.
- C2. This criterion is intended to serve as a quantitative indicator of the presence of the technical safeguards against deficit spending in the implementation phase of the budget cycle (Smith and Hou, 2009). We construct a variable called SUM789 by summing the values corresponding to BBR7-BBR9 in Table 1. SUM789 has a

⁶ Snell (1996, p.6) notes that “Almost all federal reimbursements or grants in aid to a state are committed to specific purposes, and the governor and legislature have little discretion over the use of most federal funds.” Consequently, the focus of budget balancing efforts is on the smaller general fund.

maximum value of three and a minimum value of zero. We associate a higher degree of fiscal stringency with the presence of at least two technical rules; so a state's degree of fiscal stringency is considered as relatively high if $SUM789 \geq 2$.⁷

- C3. This criterion splits the sample on the basis of whether there is a provision impelling the state to balance the budget based on “own-source” revenues alone. A relatively high degree of fiscal stringency in this case is associated with $BBR2 = 1$.
- C4. This criterion emphasizes the role of debt limits. States with such limits have presumably a more difficult time to use borrowing as a way of covering chronic deficits and, therefore, need to be more fiscally disciplined. According to this criterion, the degree of fiscal stringency is considered as high if $BBR5 = 1$.

The data for all state fiscal variables are taken from the computer files of the *Census Bureau* (State Government Finances). The variables were adjusted for inflation using the “state and local government consumption and investment” price index of the *Bureau of Economic Analysis* (US Department of Commerce). The ACIRFSI values were taken from ACIR (1987). Finally, the information on the nine individual BBR provisions was extracted from Hou and Smith (2006, Table 2).

4.2 Empirical Methodology

Our empirical test of the sustainability hypothesis is rooted in Quintos (1995), and proceeds as follows. We begin by testing revenue and expenditure variables for unit roots. If both variables are found to be $I(0)$, we conclude that sustainability holds, and proceed no further. If the variables are found to be $I(1)$, however, the testing proceeds by estimating the potentially cointegrated regression in (4). Additional hypotheses are then sequentially tested in relation to the estimated slope coefficient as follows:

⁷ Note that $SUM789$ includes the no-deficit-carry-over provision ($BBR9$) irrespective of whether the provision is constitutional or statutory.

First, we test the null hypothesis that $\beta = 0$ versus the one-sided alternative that $\beta > 0$. If the null is accepted, the deficit is growing faster than the economy (as measured by the mean interest rate) and is thus not sustainable, whereas if it is rejected, then we test $\beta = 1$ versus $\beta \neq 1$. If the null is rejected to the left, then $0 < \beta < 1$ and therefore the sustainability is weak, whereas if the null is rejected to the right, then revenues are growing faster than expenditures.

Only if the null of $\beta = 1$ is accepted is the condition of cointegration going to be important. If cointegration holds the sustainability is strong, but if it fails to hold the sustainability is weak.

As we noted earlier, the empirical literature on fiscal sustainability in the U.S. is almost solely based on the time series data at the federal government level. A natural way to expand the literature using the state data is to estimate (4) repeatedly for each state in the sample and then test for sustainability. The problem with this state-by-state approach is that it results in loss of the information contained in the rest of cross-sectional dimension. Note in particular that under strong sustainability $\beta = 1$ for all states, which means that the information regarding β can be pooled. The state-by-state approach does not make use of this information and is, therefore, inefficient. Moreover, we cannot apply pooled OLS, as this requires that the regression error, ε_{it} , to be cross-sectionally independent. Given the high degree of intra-U.S. fiscal collaboration and synchronization and the presence of federal policy shocks that affect the fiscal positions of all states, the assumption of cross-sectional independence seems implausible.

In order to assess the significance of the cross-state correlation problem, we compute the pair-wise correlation coefficients of the state level surplus, $R_{it} - G_{it}$. The simple average of these correlation coefficients across all the 1,081 state pairs, together with the associated cross-sectional dependence (CD) test statistic discussed in Pesaran *et al.* (2008), are given in Table 2 for each of the balances B1-B4. The average correlation coefficient ranges between 0.21 and

0.84, and the CD statistic is highly significant for all four balances, which is suggestive of strong cross-state dependence.

To formally illustrate the implications of this dependence, suppose that the regression error in (4) has the following common factor structure:

$$\varepsilon_{it} = \lambda_i' \mathbf{f}_t + \nu_{it} \quad (6)$$

where \mathbf{f}_t is an r -dimensional vector containing the unobserved factors, which could represent regulations, fiscal shocks, or any other feature affecting revenues that is common for all states. The disturbance ν_{it} is assumed to be mean zero and uncorrelated across states, but potentially correlated over time. The factors in (6), which are allowed to be also serially correlated, are introduced in order to model potential cross-state dependence in ε_{it} . The extent of this dependence is determined by λ_i , which is a vector of loading parameters that measure the effect of the common factors. By inserting (6) into (5) we obtain

$$G_{it} - R_{it} = (1 - \beta)G_{it} - \alpha_i - \lambda_i' \mathbf{f}_t - \nu_{it}. \quad (7)$$

This means that now $G_{it} - R_{it}$ is allowed to be dependent across states, which is also consistent with the estimated cross-state correlations reported in Table 1. As an indication of the reasonability of the assumed factor model, suppose for simplicity that $\beta = 1$ and that there is only one factor, in which case \mathbf{f}_t can be well-approximated by the average surplus for each t , see Pesaran (2007). By simply taking deviations from this average, the highest correlation in Table 2 drops from 0.837 to 0.157. Hence, one factor is enough to capture most of the correlation.

The above results suggest that the regression errors are correlated across states, and therefore, the pooled OLS method cannot be applied. But even if the errors are cross-sectionally independent, OLS will still be inefficient if the regressors are correlated with the idiosyncratic error term. A common approach to alleviate this problem is to use fully modified (FM) estimation techniques. In our case, there is not only this correlation, but also the correlation between the

regressor and the common factors. The FM estimator of Bai and Kao (2005) accounts for both of these correlations and is, therefore, appropriate for our purposes. The estimator, which can be seen as a factor augmented version of the more conventional FM estimator of Kao and Chiang (2000), is implemented in two steps. The first step is to estimate the unknown factors, which can be done by using the method of principal components. In the second step, β is estimated by FM techniques conditional upon the first-step factor estimates.

5. Empirical Results

5.1 Unit Root Tests

Since the data appear to be cross-sectionally correlated, we cannot use the conventional approach of just combining individual Dickey-Fuller unit root tests as if they were independent. For our purpose, we employ the $\overline{\max}$ bootstrap test of Smith *et al.* (2004), which uses a sieve sampling scheme to account for error dependence across both the time series and cross-section dimensions of the panel.⁸ The test, which can be seen as a bootstrap version of the well-known Im *et al.* (2003) test, is constructed with a common 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. On the other hand, if the null is accepted, we conclude that the panel is $I(1)$ as a whole, which is a necessary condition for cointegration.

The order of the sieve is permitted to increase with T at the rate $4(T/100)^{2/9}$ and so is the lag length of the individual Dickey-Fuller test regressions. As for the deterministic component, since the data are clearly trending, the test regressions are fitted with a constant and a linear time trend. The bootstrap distributions are based on 1,000 replications.

The bootstrap p -values reported in Table 3 suggest that the evidence against the unit root null is fairly weak. At the 5-percent level, there are only 14 rejections of the null, and at the 1-

⁸ Smith *et al.* (2004) also propose four other tests, denoted as \overline{t} , \overline{LM} , $\overline{\min}$, and \overline{WS} . However, the $\overline{\min}$ test generally performed best in their simulations, and we will therefore only consider this test.

percent level, there are no rejections at all. Note that the number of rejections does not systematically vary when we consider subsamples defined based on the criteria C1-C4. The same is true when considering the results from across the four balances B1-B4. Thus we proceed as if both revenue and expenditure variables are $I(1)$ for both the full sample and the subsamples.

5.2 Estimation

The first step in obtaining the Bai and Kao (2005) FM estimator is to apply principal components to extract the factors. The number of factors to estimate is determined using the IC_1 information criterion of Bai and Ng (2002) with the maximum number of factors set to five.

In the second step the first-step factor estimates are used to correct the pooled OLS estimator for the correlation between the regressor and the regression error, which requires estimation of the associated nuisance parameters. For this purpose, we follow Bai and Kao (2005) and use the Newey and West (1994) procedure. For better estimation accuracy, Bai and Kao (2005) suggest iterating between steps one and two, which can be continued until convergence. However, since the results were basically identical, we only report the two-step estimates.

The FM t -test has an asymptotic standard normal distribution under the null hypothesis, suggesting that in large samples normal p -values can be used. However, as Westerlund (2007) demonstrates, these p -values can sometimes provide a poor approximation to the empirical distribution. Therefore, we follow his proposal and compute bootstrapped p -values.⁹ The estimation results are reported in Table 4. The p -values for the test of the hypothesis of $\beta = 0$ versus $\beta > 0$ were all close to zero, and are thus not reported. The p -values in the table correspond to the hypothesis of $\beta = 1$ versus $\beta \neq 1$.

Moving across the balances in Table 4 first, we see that the slope estimates corresponding to the broadest balance B1 (total revenues less total expenditures) are generally larger than the

⁹In contrast to the sieve approach of Smith *et al.* (2004), in which the serial correlation is modeled, the bootstrap of Westerlund (2007) is based on a nonparametric block sampling scheme, where the length of each block is allowed to increase with T at the rate $T^{1/5}$. The number of bootstrap replications is set to 1,000.

hypothesized value of unity, although not statistically so. The slope estimates corresponding to the B2 balance (general revenues less general expenditures) are somewhat smaller in size, but still not statistically different from one. This suggests that the strong version of sustainability holds for these broadly defined state balances. Note that these results hold regardless of the sample composition suggesting that broadly defined revenues are (more than) sufficient to support corresponding expenditures independent of the absence or presence of corresponding BBRs. For the narrower balance B3 (general revenues less current expenditures) similar results are obtained, although in a few cases the null of $\beta=1$ is rejected in the right tail of the hypothesized distribution under the null hypothesis, suggesting that $\beta > 1$. The results corresponding to the narrowest of the four balances B4 (own general revenues less current expenditures), however, are strikingly different. In all cases, the null is strongly rejected in the left tail, suggesting that $\beta < 1$. This is consistent with the weak version of sustainability. Again, the results are independent of the sample composition suggesting own-source general revenues are not adequate to intertemporally match current expenditures irrespective of the degree of fiscal stringency.

Moving down along the sample split criteria, we note that the estimated β s are almost uniformly larger in the first subsample than those in the second subsample regardless of the fiscal stringency criterion employed. In particular, the differences are relatively large corresponding to criteria C2-C3 under B3 and criteria C1-C2 under B4 balance definitions. The differences, however, are not always statistically significant. More importantly, in no case the coefficient estimates allow us to conclude that strong (weak) fiscal sustainability is only associated with states characterized by a relatively high (low) degree of fiscal stringency. Note that the full-sample results are consistent with the subsample results in that they support strong sustainability in all but the narrowest of the four balances.

5.3 Cointegration Tests

The cointegration test that we consider is taken from Westerlund and Edgerton (2007), who develop a sieve bootstrap test for the null hypothesis of cointegration versus the alternative that there is at least one state for which cointegration does not hold. The test is constructed with cointegration under the *null hypothesis*, which is more appropriate than having it under the alternative. In particular, with no cointegration under the null, then the alternative is that there is at least one state where cointegration holds. This means that we cannot really say in which states G_{it} and R_{it} are cointegrated in case of a rejection.

Just as in the Smith *et al.* (2004) test, the order of the sieve approximation is allowed to increase with T , and number of bootstrap replications is set to 1,000. All long-run variances are estimated using the Newey and West (1994) procedure. Since the bootstrap p -values for both the full sample and the subsamples are all equal to one, at least down to the third decimal, we do not report them. The results suggest that there are no violations of the cointegration condition, implying that for B1 and B2, and also for B3 at the 1-percent significance level, the strong version of the sustainability hypothesis cannot be rejected.

5.4 Robustness Checks

The robustness of our results has been tested in several ways. As a way of checking the robustness of the finding that for B1-B3 the deficit is strongly sustainable, we applied the Smith *et al.* (2004) test to the deficit, which should be $I(0)$ since revenues and expenditures are cointegrated and one-to-one. As expected, the bootstrap p -value is zero to the third decimal, meaning that the unit root null must be rejected at all conventional significance levels.

Moreover, since there is no consensus about the measures of revenues and expenditures in the empirical literature, we repeated our analysis replacing the level variables with the same

variables scaled by state personal income.¹⁰ The results indicate that our original findings with respect to the unit root and cointegration conditions seem to be very robust to this change. However, there are some differences in the slope coefficient estimation results. In particular, the slope estimates based on the scaled variables reported in Table 5 are generally smaller than before. This implies that state expenditures were more elastic than state revenues with respect to changes in the size of the state economy over time.

A noteworthy aspect of the results in Table 5 is that in a number of cases associated with B1-B3, we cannot reject the null of $\beta = 1$ for the first subsample, but can (marginally) reject it for the second subsample. This result is particularly observed in relation to B2, which is a balance that includes only general expenditure and revenues, and the debt limit provision (BBR5). Thus, for scaled variables the evidence is consistent with the conjecture that BBRs matter in relation to strong fiscal sustainability. Note, however, that for the most narrowly defined of the four balances (B4) we have consistent and strong evidence of weak sustainability as in Table 4. Finally, note that in sharp contrast to Table 4 results, the full-sample results in Table 5 indicate that the strong version of sustainability is rejected in all but the broadest of the four balances.

6. Summary and Concluding Remarks

In this paper, we expanded the empirical literature on state government fiscal behavior by examining whether (a) the fiscal sustainability hypothesis was supported in the American states and (b) the evidence supporting the hypothesis was more pronounced in states characterized by a higher degree of fiscal stringency. We found statistical evidence that a sufficient condition for strong (weak) version of sustainability was consistently satisfied for the full panel of 47 states *and* most subsamples of different degrees of fiscal stringency when the budget balance was the most broadly (narrowly) defined of our balances. These results, which held regardless of whether

¹⁰ *Bureau of Economic Analysis* does not provide state gross product (GSP) data on a consistent basis for the entire sample period.

the fiscal variables were expressed in levels or ratios, implied that the degree of fiscal stringency did not make much difference in relation to the sustainability of these extreme balances. For balances lying between the two extremes, strong sustainability was also supported for the ratio variables; albeit less consistently. Moreover, when the ratio variables were used, evidence of strong (weak) sustainability was observed in subsamples that were characterized by relatively high (low) degrees of fiscal stringency.

According to our results, fiscal sustainability seems to be sensitive to the definition of the budget balance employed and the way the fiscal variables are measured. However, considering that in growing economies scaled variables are more suitable than level variables and that a portion of the broadest balance is actually subject to balance budget requirements (BBRs), we have good reasons to give more weight to part of the results associated with the ratio variables and the middle balances. Based on these results, we may then tentatively conclude that BBRs in the form of constitutional no-deficit-carry-over provision, anti-deficit measures in the implementation phase of the budgetary process, reliance on own revenue, and debt limits are associated with fiscal sustainability.

A few concluding remarks are in order. Fiscal sustainability by definition is a long-term condition and evidence in favor it is not inconsistent with presence of (large) fiscal imbalances in the short term. Accordingly, growing state deficits in much of post-sample period (that is, from 2007 on) due to significant shocks to state economies are not necessarily a cause for alarm; unless further data indicate that the shocks permanently shifted the long-term paths of state revenues and expenditures and put state fiscal deficit on an unsustainable trajectory. We tend to think that the robustness of the finding that the most broadly defined state fiscal balance satisfies the condition of strong sustainability should be somewhat reassuring in view of the fact that the sample period includes several periods of state fiscal distress.

Our results are also not inconsistent with individual state-level deviations from long-run sustainability. For example, while our results seem very robust with respect to the sample used in

the estimation, there might still be some states for which the slope coefficients do not satisfy the (strong or weak) sustainability condition, but that these deviations get averaged out in our pooled estimation methodology.

Finally, our statistical evidence regarding whether strong (weak) sustainability is due to the presence (absence) of BBRs is not conclusive. Doubts regarding the effectiveness of BBRs have been expressed by, among others, Briffault (1996) who noted that states are able to operate under BBRs largely because they are neither stringent nor binding. However, based the totality of the statistical evidence presented in this paper, we tend to concur with Snell's (1996, p.7) more optimistic assessment that "State balanced budget requirements matter. States that have the most rigorous requirements for balanced budget are the most likely to balance their budgets. States with less rigorous budgeting requirements make use of them. These outcomes tend to occur even though most states lack strong enforcement mechanisms." Accordingly, more rigorous requirements in some states and improved enforcement mechanism in all states may be crucial in ensuring state fiscal sustainability.

References

- Advisory Commission on Intergovernmental Relations (1985). "Fiscal Discipline in the Federal System: Experience of the States," Washington, D.C.
- Ahmed, S. and J. Rogers, (1995). "Government Budget Deficits and Trade Deficits. Are Present Value Constraints Satisfied in Long-term Data?", *Journal of Monetary Economics*, 36 (2), 351-374.
- Alesina, A. and T. Bayoumi (1996). "The Costs and Benefits of Fiscal Rules: Evidence from U.S. States," National Bureau of Economic Research, Working paper #5614.
- Alt, J. and R.C. Lowry (1994). "Divided Government and Budget Deficits: Evidence from States," *American Political Science Review*, 88, 811-28.
- Bai, J., and C. Kao (2005). "On the Estimation and Inference of a Panel Cointegration Model with Cross-Sectional Dependence," In Baltagi, B. (Ed.), *Contributions to Economic Analysis*. Elsevier, Amsterdam.
- Bai, J. and S. Ng, (2002). "Determining the Number of Factors in Approximate Factor Models," *Econometrica*, 70(1), 191-221.
- Bayoumi, T. and B. Eichengreen (1995). "Restraining Yourself: The Implications of fiscal Rules for Economic Stabilization," *IMF Staff Papers*, 32-48.
- Bohn, H. (1998). "The behaviour of U.S. public debt and deficits," *Quarterly Journal of Economics*, 113 (3), 949-963.
- Bohn, H. (2007). "Are stationarity and cointegration restrictions really necessary for intertemporal budget constraint," *Journal of Monetary Economics*, 57, 1837-47.
- Bohn, H and R.P. Inman (1996). "Balanced-budget Rules and Public Deficits: Evidence from the States," Carnegie-Rochester Conference Series on Public Policy.
- Briffault, Richard (1996). "Balancing Acts: The Reality Behind State Balanced Budget Requirements," Twentieth Century Fund, New York.
- Endersby, J. W. and M. J. Towle (1997). "Effects of constitutional and political controls on state expenditures," *Publius*, 27 (1):83-98.
- Fatás, A. and Mihov, I. (2006). "The Macroeconomics Effects of Fiscal Rules in the US States," *Journal of Public Economics*, 90, 101-117.
- Goyal, Rajan, J. K. Khundrakpam, Partha Ray (2004). "Is India's public finance unsustainable? Or, are the claims exaggerated?" *Journal of Policy Modeling* Vol. 26, Iss. 3; 401-420.
- Hakkio, G. and Rush, M. (1991). "Is the Budget Deficit 'Too Large?'" *Economic Inquiry*, 29 (3), 429-445.
- Haug, A. (1991). "Cointegration and Government Borrowing Constraints: Evidence for the United States," *Journal of Business & Economic Statistics*, 9 (1), 97-101.
- Hou, Yilin and Daniel L. Smith (2006). "A Framework for Understanding State Balanced Budget Systems: Reexamining Distinctive Features and Operational Definitions," *Public Budgeting and Finance*, 26(3), 22-45.
- Hou, Yilin and Daniel L. Smith (2009). "Do State Balanced Budget Requirements Matter?: Testing Two Explanatory Frameworks," Available at <http://ssrn.co/abstract=913842>.
- Im, K. S., Peseran, M. H., Shin, Y. (2003). "Testing for unit roots in heterogeneous panels," *Journal of Econometrics*, 115, 53-74.
- Kao, C., and M. H. Chiang (2000). "On the Estimation and Inference of a Cointegrated Regression in Panel Data," *Advances in Econometrics* 15, 179-222.
- Kia, Amir (2008). "Fiscal sustainability in emerging countries: Evidence from Iran and Turkey" *Journal of Policy Modeling*, 30, 957-972.
- Krol, R. and S. Svorny (2007). "Budget Rules and State Business Cycles," *Public Finance Review*, 35, 530-544.
- Levinson, A. (1998). "Balanced Budgets and Business Cycles: Evidence from the States," *National Tax Journal*, 51, 4, 715-732.

- Martin, G. M. (2000). "US Deficit Sustainability: A New Approach Based on Multiple Endogenous Breaks," *Journal of Applied Econometrics*, 5(1), 83-105.
- MacDonald, R. (1992). "Some Tests of the Government's Intertemporal Budget Constraint using US Data," *Applied Economics*, 24 (12), 1287-1292.
- Newey, W. and K. West (1994). "Autocovariance Lag Selection in Covariance Matrix Estimation," *Review of Economic Studies*, 61, 631-653.
- Pesaran, H. M., A. Ullah, and Y. Yamagata (2008). "A Bias-Adjusted LM Test of Error Cross Section Independence," Forthcoming in *Econometrics Journal*.
- Poterba, J. (1994). "State Responses to "Fiscal Crises: The Effects of Budgetary Institutions and Politics," *Journal of Political Economy*, 102, 799-821.
- Poterba, J. (1996). "Budget Institutions and "Fiscal Policy in the U.S. States," AEA Papers and Proceedings, 86, 2, 395-400.
- Quintos, C. (1995). "Sustainability of the Deficit Process with Structural Shifts," *Journal of Business and Economic Statistics*, 13, 409-417.
- Rose, S. (2006). "Do fiscal rules dampen the political business cycle? *Public Choice*, 129, 407-431.
- 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.
- Smith, Daniel L. and Yilin Hou (2008). "State Spending Behavior under Heterogeneous Balanced Budget Requirement Systems: A Long-Panel Study," Available at <http://ssrn.co/abstract=1024387>.
- Snell, Ronald K. (1996; updated 2004). "State Balanced Budget Requirements: Provisions and Practice," National Conference of State Legislators.
- Sorensen, Bent E., Lisa Wu and Oved Yosha (2001). "Output fluctuations and fiscal policy: U.S. state and local governments 1978-1994," *European Economic Review*, 45, 1271-1310
- Tanner, E. and Liu, P. (1994). "Is the Budget Deficit 'Too Large?': Some Further Evidence," *Economic Inquiry*, 32, 511-518.
- Trehan, B., Walsh, C. (1991). "Testing Intertemporal Budget Constraints: Theory and Applications to U.S. Federal Budget and Current Account Deficits," *Journal of Money, Credit and Banking*, 3, 210-223.
- von Hagen, J. (1998). "Fiscal Policy and Intranational Risk-sharing," CEPR Discussion Paper, 126.
- Westerlund, J. (2007). "Unbiased Estimation of Cointegrated Panel Regressions with Cross Section Dependence," *Journal of Financial Econometrics*, 5, 491-522.
- Westerlund, J. and D. L. Edgerton (2007). "A Panel Bootstrap Cointegration Test," *Economics Letters*, 97, 185-190.

Table 1. State Balance Budget Requirement (BBR) Systems

State	BBR1	BBR2	BBR3	BBR4	BBR5	BBR6	BBR7	BBR8	BBR9	ACIRFSI	SUM789
Alabama	1	0	1	1	1	0	0	1	1	10	2
Arizona	1	0	1	1	1	0	1	1	1	10	3
Arkansas	1	0	0	1	0	0	0	1	0	9	1
California	1	0	0	1	0	1	0	1	0	6	1
Colorado	1	1	1	1	1	0	0	1	0	10	1
Connecticut	1	0	1	1	0	0	1	1	0	5	2
Delaware	0	0	1	1	0	0	1	0	0	10	1
Florida	1	0	0	1	0	0	0	1	0	10	1
Georgia	1	1	1	1	1	0	1	0	0	10	1
Idaho	1	1	1	1	0	0	0	1	0	10	1
Illinois	1	0	1	1	1	0	1	0	0	4	1
Indiana	0	0	1	0	0	0	0	0	0	10	0
Iowa	1	0	1	0	1	0	0	0	0	10	0
Kansas	0	0	1	1	1	0	0	1	0	10	1
Kentucky	1	1	1	1	1	0	1	1	0	10	2
Louisiana	1	0	0	1	0	0	0	1	0	4	1
Maine	1	0	1	0	0	0	0	0	0	9	0
Maryland	1	0	0	1	0	0	1	0	0	6	1
Massachusetts	1	0	1	1	0	1	1	1	0	3	2
Michigan	1	0	1	1	1	0	1	1	0	6	2
Minnesota	1	0	1	0	0	0	0	1	0	8	1
Mississippi	1	0	1	1	1	0	0	1	1	9	2
Missouri	1	0	1	0	1	0	0	1	0	10	1
Montana	1	1	0	1	0	0	1	1	1	10	3
Nebraska	1	0	1	0	1	0	0	1	0	10	1
Nevada	1	0	1	1	1	0	1	0	0	4	1
New Hampshire	1	0	0	1	0	0	1	1	0	2	2
New Jersey	1	0	0	1	0	0	1	1	0	10	2
New Mexico	1	1	0	0	1	0	0	0	0	10	0
New York	1	1	1	1	0	0	0	0	0	3	0
North Carolina	1	0	1	1	0	0	0	1	1	10	2
North Dakota	0	0	0	0	0	0	0	0	0	8	0
Ohio	1	0	1	1	1	0	0	1	0	10	1
Oklahoma	1	0	0	1	0	0	1	1	0	10	2
Oregon	1	0	1	1	0	0	0	1	0	8	1
Pennsylvania	1	0	0	1	0	0	0	0	0	6	0
Rhode Island	1	0	1	1	1	0	1	1	0	10	2
South Carolina	1	0	1	1	1	0	0	1	1	10	2
South Dakota	1	1	1	1	0	0	0	1	0	10	1
Tennessee	0	0	1	0	1	0	0	1	0	10	1
Texas	0	1	1	1	1	0	0	1	1	8	2
Utah	1	1	1	1	1	0	0	1	0	10	1
Vermont	0	0	1	0	0	0	0	0	0	0	0
Virginia	1	0	1	0	0	0	0	0	0	8	0
Washington	0	0	1	1	0	0	1	0	0	10	1
West Virginia	1	0	0	0	0	0	0	1	0	8	1
Wisconsin	0	0	1	1	1	0	1	1	1	6	3
Total	38	10	34	35	21	2	17	32	8		

Notes: The budget requirements BBR1–BBR9 are defined in Section 3. A BBR's value equals one if associated provision is in place, and zero otherwise. Values have been assigned by authors based on the information in Hou and Smith (2006, Table 2). ACIRFSI is the ACIR index of fiscal stringency. SUM789 is the sum of BBR7–BBR9.

Table 2: Cross-state Budget Balance Correlations

Test	B1	B2	B3	B4
Average correlation	0.656	0.207	0.277	0.837
CD	146.35	46.08	161.78	86.59
<i>P</i> -value	0.000	0.000	0.000	0.000

Notes:

The cross-sectional dependence (CD) test and the associated p-values are for the null hypothesis of no cross-correlation.

B1–B4 refer to the budget balances defined in Section 4.1.

Table 3. Unit Root Test p -values

Criterion	Sample	B1		B2		B3		B4	
		<i>Rit</i>	<i>Git</i>	<i>Rit</i>	<i>Git</i>	<i>Rit</i>	<i>Git</i>	<i>Rit</i>	<i>Git</i>
Subsample results									
C1	ACIRFSI=10	0.015	0.076	0.058	0.064	0.329	0.181	0.086	0.177
	ACIRFSI<10	0.014	0.060	0.055	0.065	0.157	0.024	0.018	0.022
C2	SUM789 \geq 2	0.099	0.117	0.132	0.073	0.115	0.053	0.015	0.041
	SUM789<2	0.087	0.208	0.307	0.188	0.343	0.088	0.075	0.097
C3	BBR2=1	0.089	0.392	0.304	0.250	0.288	0.228	0.253	0.234
	BBR2=0	0.110	0.098	0.225	0.100	0.221	0.041	0.012	0.041
C4	BBR5=0	0.073	0.323	0.486	0.292	0.452	0.161	0.088	0.199
	BBR5=2	0.106	0.062	0.107	0.047	0.105	0.036	0.022	0.030
Full sample results									
		0.077	0.142	0.223	0.090	0.207	0.059	0.207	0.059

Notes:

The p -values correspond to Smith *et al.* (2004) max unit root test. The null hypothesis of a common unit root is tested against heterogeneous autoregressive roots under the alternative.

C1-C4 refer to the sample split criteria; B1-B4 refer to the budget balance definitions (see Section 4.1).

Rit and *Git* are measures of government revenues and expenditures, respectively.

Table 4. Estimated Slope Coefficients (level variables)

Criterion	Sample	B1		B2		B3		B4	
		β	<i>p</i> -value	β	<i>p</i> -value	β	<i>p</i> -value	β	<i>p</i> -value
Subsample results									
C1	ACIRFSI=10	1.101	0.264	1.009	0.789	1.062	0.079	0.738	0.000
	ACIRFSI<10	1.099	0.444	0.993	0.926	1.016	0.763	0.679	0.000
C2	SUM789 \geq 2	1.105	0.306	1.031	0.439	1.087	0.041	0.746	0.000
	SUM789<2	1.099	0.437	0.988	0.856	1.010	0.862	0.677	0.000
C3	BBR2=1	1.103	0.433	1.050	0.491	1.116	0.053	0.640	0.000
	BBR2=0	1.099	0.424	0.980	0.724	0.998	0.984	0.705	0.000
C4	BBR5=1	1.118	0.354	1.015	0.687	1.054	0.199	0.706	0.000
	BBR5=0	1.095	0.444	0.991	0.894	1.016	0.785	0.686	0.000
Full sample results									
		1.100	0.393	0.997	0.958	1.025	0.653	0.690	0.000

Notes:

β refers to the cointegration slope estimated using the Bai and Kao (2004) FM estimator. The regressions are fitted with an unreported intercept.

The *p*-values are for a double-sided test of hypothesis that $\beta=1$.

C1-C4 refers to the sample split criteria; B1-B4 refer to the budget balance definitions (see Section 4.1).

Table 5. Estimated Slope Coefficients (ratio variables)

Criterion	Sample	B1		B2		B3		B4	
		β	<i>p</i> -value	β	<i>p</i> -value	β	<i>p</i> -value	β	<i>p</i> -value
Subsample results									
C1	ACIRFSI=10	0.941	0.623	0.933	0.152	0.890	0.202	0.562	0.000
	ACIRFSI<10	0.859	0.091	0.907	0.029	0.923	0.138	0.571	0.000
C2	SUM789 \geq 2	0.898	0.276	0.910	0.058	0.916	0.181	0.554	0.000
	SUM789<2	0.898	0.300	0.925	0.076	0.901	0.116	0.573	0.000
C3	BBR2=1	0.873	0.287	0.938	0.324	0.824	0.097	0.499	0.000
	BBR2=0	0.903	0.272	0.915	0.030	0.926	0.175	0.582	0.000
C4	BBR5=1	0.941	0.615	0.941	0.233	0.903	0.315	0.529	0.000
	BBR5=0	0.871	0.124	0.906	0.025	0.909	0.064	0.589	0.000
Full sample results									
		0.900	0.270	0.920	0.041	0.906	0.098	0.566	0.000

Notes: See Table 4.