Fiscal Implications of Legislative and Executive Term Limits at the State Level[1*]

Dillon T. Klepetar

1 * All written correspondence can be sent to Dillon Klepetar, School of Public Affairs, American University, 4400 Massachusetts Avenue NW, Washington, D.C. 20016. Data requests and other queries by email to klepetar@american.edu. Special thanks goes to Professor Laura Langbein for her comments on an earlier draft of this manuscript.

Abstract

Proponents of term limit legislation have long maintained that these institutional constraints would end the era of the "career politician". One justification for doing away with senior representatives is the argument against a culture of spending which supposedly indoctrinates the longest serving members in government who seek reelection regularly. Recent work has demonstrated that adoption of term limits had the unintended consequence of increasing spending levels—a finding which would surprise diehard advocates of the reform. As with all scientific findings which are incorporated into our working knowledge, findings must be consistent across empirical studies. Questions which seek to better understand how fiscal appropriation responds to institutional changes are perhaps the most important to theory-building and are therefore worth revisiting. Using a set of economic and political indicators, to include legislative and gubernatorial term limits, captures much of the variance we observe in state expenditures. In contrast to previous research, results here indicate that term limiting legislators does not affect overall state spending at traditional confidence intervals. Moreover, improvements to previous models suggest that term limits have a differential effect on expenditures when imposed on state senators, state house members and on governors. This pattern is not surprising given that member goals are shaped by their unique institutional orientation which responds to rule changes in equally distinctive ways.

Introduction

This article began as a replication of a study that found state legislative term limits had the unintended consequence of increasing expenditure levels. Once I collected the data and began inspecting it, patterns emerged that stood in contrast to the conclusions put forth in the original publication. What follows is a review of the original work and the improvements I have been compelled to make along the way. The term limits story is far more complex and a great deal more nuanced than originally thought.

I begin this article with a short description of the study I began replicating. From here, I give an abridged overview of the relevant literature in order to situate the research question and to justify the merits of the study. Next, I review the methodology of the original article so that the reader understands the task at hand. In this section, I pay close attention to methodological issues I aim to address as I revise the model. Then the method that will be used for the present study is laid out in some detail. This discussion should make clear any adjustments I have made to the original work with respect to data sources, coding variables, and regression analysis. Finally, I demonstrate the results in words and graphics, including several regressions. For purposes of comparison, regressions from the original paper are also reported. Several variables omitted from the earlier analysis are included in this model, and the findings suggest that these revisions improve our understanding greatly. Since the original data was unavailable for replication, a portion of the results section is also dedicated to addressing the comparability of these two studies. I conclude by discussing research outcomes, acknowledge lingering questions on this topic, and providing directions for future research.

One recent study by Abbie H. Erler (2007) on the effects of state legislative term limits suggests that the advent of this "reform" brings higher levels of government spending, a finding that defies conventional wisdom on the subject. Fortunately, the data used can be generated readily using a number of resources including the National Governors' Association, the National Association of State Budget Officers as well as U.S. Census data on economic and social indicators now widely available on the internet. My analysis of *Legislative Term Limits and State Spending* will be as much an effort to replicate the findings as to improve upon the method. The data presented by Erler (2007) has several shortcomings that I hope to address.[2]

Term limits are a hot-button issue in many states as well as in the US Congress. Advocates on both sides have logical reasons for taking their position. Proponents of term limitations have long argued that the reform will oust career politicians along with the loose spending habits they develop.[3] Opponents claim that senior politicians provide the fiscal discipline that comes with experience in the legislature. Still others contend that representatives in the state house and senate often have many years to execute their political aims, efforts which invariably cost taxpayers. By reducing the number of terms members can serve, the personal

2 The merits of this study will not be fully entertained here since they have been justified in the original paper. I should be clear at the outset what this study is and what it is not. The goal of this research is simply to observe and measure the relationship between term limits and spending, not to make causal arguments about their association.
3 Robert W. Reed et al., "The Relationship between Congressional Spending and Tenure with an Application to Term Limits," Public Choice 94 (1998): 98. For a discussion of incumbent pork spending as a reelection strategy see Barry Weingast (1994) on universalism and Robert M. Stein and Kenneth N. Bickers (1995) on particularistic constituency benefits.

goal of making good public policy must be achieved hurriedly before a legislator is term limited.[4] Some champions of term limits see reelection as a perpetual incentive to dole out constituency benefits in exchange for votes, a pattern that puts fiscal burdens on the state tax pool.[5] Another empirical study on the issue will certainly not put an end to debate; rather, this investigation has higher aims. In fact, the following study may provide insights into the institutional incentives, or more appropriately, the institutional constraints, that affect fiscal policy outcomes at the state level.

Method: Erler (2007)

To measure the effects that term limits have on state spending, Erler "analyzed fiscal data from 47 continental US states from 1977-2001."[6] The unit of analysis is the state-year during the panel. Her dependent variable is "general spending" which includes investments in "education, highways, welfare and interest on the debt," however general spending is not used in the final analysis.[7] Data sources for the original research were drawn from the "Statistical Abstract of the United States, State Government Tax Collections, (sic) the Fiscal Survey of the States and the Book of States".[8] Although the author provides information about data collection, she does not reveal which resources were used to extract variables. For instance, state expenditure data is available both in the Statistical Abstract of the United States (SAUS) and in the Fiscal Survey of the States (FSS) publications.

The theoretically important variable in Erler (2007) is the implementation of term limits in the state legislature, which she translates into a dichotomous variable. Within the panel, the intuitive procedure would be to code states with 1 when term limits were in place and 0 otherwise in a particular year. Erler (2007) instead believed that this process would not allow for the anticipation legislators would feel two years prior to being term limited. This logic informed her decision to essentially lag term limits by one year in states who adopted the legislation. Handling term limits in this way had major implications for way the data was coded. Missouri began limiting its legislators in 2002, however Missouri is considered a "term limits state" in the original sample because of the coding system outlined above. Within the group of states with term limits, there are also several which had not fully implemented term limits when the panel expires in 2001. Some states had term limited their house members prior to 2001 but had not yet had constrained state senators due to the longer terms in that chamber. Evidence presented later suggests these factors could have heavily distorted the results.

4 Richard Fenno (1978) demonstrates that legislators have three major goals: reelection; power or influence within the chamber; and so-called good public policy. Term limits have obvious ramifications for the goal of reelection and potentially less explicit effects on the latter two. Influence within the chamber is traditionally tied to seniority, a norm which would be undermined with tenure limitations. Potential impacts on the pursuit of good public policy are discussed in the text.

5 Doug Bandow, "Real Term Limits: Now more than Ever," Cato Institute's Policy Analysis 221 (1995) and George F. Will, Restoration: Congress, Term Limits, and the Recovery of Deliberative Democracy (New York: Free Press, 1992).

6 Erler (2007) justifies dropping Nebraska from the sample since that state has nonpartisan election; Hawaii because public school expenditures come from the state instead of local municipalities; and Alaska because the majority of their revenue stream comes from oil and not taxes.

7 Abbie H. Erler, "Legislative Term Limits and State Spending," *Public Choice* 133 (2007): 482. The dependent variable in the original regression tables include spending per capita and spending as a percent of overall income.

8 Erler, "Legislative Term Limits and State Spending," 482.

One other consideration that Erler (2007) recognizes is that adoption of term limitations is probably not entirely an exogenous institution. This pattern is undeniably clear when one looks at which states have adopted the constraint and which states have not. At one time or another between 1977 and 2001, the following states began limiting the number of terms their state representatives could serve: Arizona; Arkansas; California; Colorado; Florida; Maine; Michigan; Montana; Ohio; and South Dakota. Any student of state politics recognizes these states as among those who have the public initiative option, a form of direct democracy. State representatives with potentially endless careers are not likely to self-impose these constraints. Therefore, the ballot option allows voters to impose term limit laws on their elected representatives in these states. Despite this clear pattern, Erler (2007) considers term limits to be "an exogenous institution."[9] We might also believe that among states with referendum or initiative voting, higher levels of spending drive support of term limits legislation within the electorate.

Erler (2007) distinguishes between strong and weak term limit laws, a variable that taps qualitative differences in the language of term limit legislation. The variety of term limit legislation includes those placing a "cap" on the number of terms served, while others require members to take time off (after being term limited) before running for office again. Several term limit laws apply to the cumulative number of terms a member has served in the house and senate. Still other states prevent members from serving anywhere in government after their tenure has expired. The objective severity of term limit legislation may have implications for state spending but in further dividing a small treatment group Erler (2007) sacrifices the generalizability of her results.

Previous literature also guided Erler's decision to include several control variables that are known to affect spending on the state level. Alt and Lowery (1994) demonstrate the importance of personal income, the state-wide unemployment rate and federal grant revenue, as these all affect expenditures. Reed (2006) argues that the party of the governor and the population density also explain some of the variance we observe in spending levels across states over time. As a result, these control variables are included in the original article. Erler (2007) used the quadratic form of income, density and unemployment instead of the raw values. Squaring income is conventionally done to normalize the distribution of observations, however density and unemployment are seldom seen in this form. In Erler's (2007) coding scheme, the governor's party is a dummy variable which takes on a value of 1 if the governor is a Democrat and 0 if she is a Republican. Again, the original article does not reveal the source of political variable data so replication data may not be identical.[10]

To analyze these data and test the relationship between term limits and state spending, Erler (2007) uses "linear cross-sectional time series models…estimated using OLS with panel-corrected standard errors" which were "corrected for first-order correlation".[11] The regression analysis also employed fixed effects for state and year to control for variation that was specific

9 Erler, "Legislative Term Limits and State Spending," 485.
10 The descriptive statistics in the Appendix reveal that data generated by the author is roughly the same as that used by Erler.
11 Erler, "Legislative Term Limits and State Spending," 485.

to the time period and states in the sample. The results from the original study can be found in Table 3 in the column labeled "Erler (2007)".

Design and Analysis

The present study is technically an attempt to replicate Erler's work, although some important changes have been made. Improvements to the original model are addressed throughout this section. To test the relationship between spending and term limits, I rely on a multiple interrupted time series design with comparison groups. The main advantage of this design is that it allows the researcher to isolate the effect of the treatment (term limits) by comparing outcome variables for the treatment and control groups both before and after the treatment is applied. This type of study is technically a quasi-experimental design because the treatment is administered exogenously however Erler (2007) and I both agree that term limits are not purely exogenous.

Variables for the replication study are substantively indistinguishable from those employed in the original paper although there remain questions with respect to Erler's (2007) data sources.[12] The main theoretical variable is whether or not a state had term limits in place; information that is widely available on the internet. My data on term limits came directly from a term limit bulletin released by National Conference of State Legislatures.[13] As mentioned, Erler (2007) lagged this variable, meaning that states who implemented term limits would be coded 1 two years before any members were to be term limited. There is some justification for altering the data in such a manner, but doing so compromises our ability to isolate the effect of term limits given the possibility of multiple treatment interference. Dummy variables that pick up more than one moving part are likely to capture effects that are beyond the scope of the theoretically important variable. If we conceptualize term limits as an institutional constraint that begins to affect spending habits before it is in place, we assumed the very relationship we are trying to test. Given the tautological risk of assuming when members begin to react to term limit laws, it is better to begin with a narrow definition of term limits than one that is too encompassing.

For states who were early adopters of term limits, legislators understand that they will be subject to term limitations from the day they are elected, yet there is no way to apply term limits only to a subset of members and track their spending habits individually. I agree that members may change their behavior in anticipation of being term limited; however time series analysis is employed for this very reason—the effect of term limits may not be noticeable right away and if treatment effects are strong enough they will be observed as the pool of legislators facing tenure constraints grows in size. Nonetheless, we are most interested in the overall effect of this institutional change without guessing how far in advance members begin reacting to their tenure limits. Coding the data intuitively (1 for the first year members are term limited) still measures the year preceding the exit of term limited legislators.[14] This cod-

12 I contacted Erler in order to clarify the sources of her data, however she did not respond. A request for the data used in her article was also refused.

13 Data were obtained from the following website: http://www.ncsl.org/Default.aspx?TabId=14842

14 This modified coding scheme technically eliminates Missouri from having any term limit observations within the panel seeing as its term limit legislation took effect in 2002. For purposes here, it is considered a non-term state.

ing scheme is also more consistent with the argument against term limits which claims that representatives will "squeeze in" pet projects just before their time is up.

The main dependent variable, state expenditures per capita, was generated using figures published directly by the U.S. Census Bureau. State level data on what the census calls "general expenditures" was divided by yearly population figures thereby creating expenditures per capita.[15] Data on unemployment rates are sourced from the Bureau of Economic Affairs, a branch of the Census. The figures represent the average state unemployment in a given year and were squared to match the original data. Total personal income within a state (in real U.S. dollars), the amount of revenue from federal grants and land area in square miles (for calculating density) were all taken from the U.S. Statistical Abstract, grants and income were subsequently squared as in Erler (2007). All Census data used here was made available upon request in a excel spreadsheet.[16] The governor's party is also a dummy variable coded 1 when the governor is a Democrat and 0 otherwise. This information was provided by State Politics and Policy Quarterly as was data on unified government. Unified government, a dummy variable, is coded 1 when the governor and both chambers are of the same party and 0 otherwise.[17] In extremely rare cases, the governor is replaced in the middle of his term by someone of the opposite party. Instead of dropping observations, I coded the variable with the fraction of time that the Democrat was in power. For instance, 9 months of Democratic control and 3 months of Republican control would be entered as .75.[18] I should point out that because data sources in the original work remain uncertain, we cannot be sure that the data used for replication came from the same sources. Descriptive statistics in the appendix indicate that the original article used a different indicator of income and perhaps grants.[19]

The reader can compare the descriptive statistics from this study to those of Erler (2007) in Appendix, Table 4. The most striking differences are in the "income^2" column and in the "term limits" column. It is possible that income was drawn from a different source or operationalized differently than the measure of income used here. The Census is the most authoritative source on state income levels, so we can be confident that our results are as reliable, if not more reliable, than those from the original article.

The other major difference is in the mean scores of "term limits", the variable we are most interested in. The result of lagging this variable as in Erler (2007) would mean the inclusion of 22 more observations for states with term limits. It appears that her unique coding scheme accounts for this disparity.

15 "General Expenditures comprises all expenditure except that classified as liquor store, utility, or insurance trust expenditure. As noted above, it includes all such payments regardless of the source of revenue from which they were financed." See also: http://www.census.gov/govs/www/class_ch8.html#S8.2

16 I am grateful to Russell Pustejovsky, Statistician, State Finance and Tax Statistics Branch, Governments Division, U.S. Census Bureau for sending me state-level data from the early 1940's in manageable form.

17 From Stefanie A. Lindquist (2005), "Predictability and the Rule of Law: Overruling Decisions in State Supreme Courts." Available at: http://academic.udayton.edu/SPPQ-TPR/tpr_data_sets.html

18 This method has demonstrated its validity in several setting. For a discussion of this measurement schema see Mark A. Smith, "The Nature of Party Governance: Connecting Conceptualization and Measurement," American Journal of Political Science 41, no. 3 (1997): 1042-1056.

19 Erler (2007) did not provide descriptive statistics for her "grants" variable.

In addition to the variables included in the original analysis it seems appropriate to improve the model where important factors were overlooked. Looking further into term limitation legislation I notice that state house and state senate term limit laws do not always go into effect at the same time. Given the staggered nature of representatives' terms, the law could not affect both chambers at exactly the same time. To accommodate for this nuance I broke out term limits into two separate variables in the later analyses.

Although density is included in the original analysis, it does not necessarily capture the way in which state demographics affect economies of scale. Money spent on public goods within each state is efficient where populations are concentrated however overall population also plays an independent role. If high density helps states spend money efficiently per unit of area, population helps boost expenditure efficiency per person. The state will inevitably provide some basic services that cost taxpayers money but because the initial investment in infrastructure, education and the like will be the most costly, population increase will distribute the cost (to a certain point) while lifting individual tax burdens.

Take highways for example: state governments have partially shouldered the burden for building a modern highway network. The cost is determined independently of the population in the state, however the more people with access to the good reduces its cost per capita. Notice that this dimension is different than the efficiencies from having a densely populated state (which most *geographically* small states have). If economies of scale affect expenditures per capita, both population and population density will help to explain more of the variance in the model.

Erler (2007) incorporates governor term limits into her model but only as a check on robustness. Governor term limits are not a new idea and many states have this institutional constraint embedded in their state Constitutions however their effects are seldom studied. Nonetheless, it is a variable that may be impacting spending and is somewhat related to the expectation that term limits affect the spending habits of government. I include a dummy indicating whether or not state has term limits on their governor in the improved models.

In order to analyze the variables discussed above, I use a regression analysis similar to the one used by Erler (2007) estimated by the following model already described above.[20]

$$\text{Yit} = \alpha + \beta T\,\text{it} + \omega C\,\text{it} + \partial K\,\text{it} + \varphi D\,\text{it} + \varepsilon\,\text{it}$$

Where Y represents state expenditures per capita, α is the intercept term, T is vector for the term limits dummy variables indicating whether or not a state had term limits at time t. T is specified three different ways in the regression; first as a unidimensional variable, then

20 Spending at year t-1 would also be a good predictor of spending at time t although the lagged expenditure variable was dropped because it was collinear with the other controls. The spending level at year t-1 could affect the party of the governor in year t as well as some of the other political and economic variables thereby leading to spurious findings. First order autocorrelation controls accommodate for part of outcome variation that is simply a function of prior spending within states over time, an acknowledgment of the autoregressive nature of state spending.

in bicameral perspective and finally to indicate constitutionally-mandated gubernatorial term limits. *C* is a vector of control variables used in the original work as well as here. *K* is a vector of variables that improve upon the original method, some of which necessitate that βT be dropped because of collinearity.[21] *D* is a vector of dummy variables that includes 46 state dummies and 24 year dummies excluding their respective reference groups. This vector holds constant any variation that is specific to particular years and particular states—so-called fixed effects. ε is the unobserved error which is assumed to be 0 in the estimating equation.

Erler (2007) claims that term limits are an exogenous institution, yet we suspect that spending and term limit legislation may be endogenous; that is, states with higher spending are more likely to support tenure limits with hopes of reigning in expenditure levels. As I have mentioned, the initiative process is a fair predictor of which states adopt term limits laws. Therefore, it seems a convenient instrumental variable since it affects adoption of term limit laws but it does not necessarily affect levels of spending.[22] Using similar notation as above we arrive at the following two stage equation.

$$T_{it} = a + B_{it} I + \mu_{it}$$

$$Yit = \alpha + \beta \hat{T} it + \omega C it + \partial K it + \varphi D it + \varepsilon it$$

Where $\hat{T}$ is the estimated value of having term limits in place as a function of I, a dummy denoting whether or not a state has the initiative process.

Econometric results which follow were generated using time-series ordinary least squares estimators with panel corrected standard errors. Fixed effects were also included. All significance tests are two-tailed with minimum 95% confidence intervals. Because non-independence of observations is a potentially serious problem for state-level time series data, I used the Durbin-Watson statistical correction for serial autocorrelation which detects correlation in the residuals and normalizes them. Regressions were computed using STATA 10.

In addition to regression analysis, I performed several different types of t-test's in order to test whether states who adopt term limits have higher spending levels than those who do not. I rely on the t-statistic again when I use a difference-in-differences test to compare changes over time in the treatment group with change over time in the group of states who failed to adopt term limit legislation during the sample frame.

Results

A simple t-test reveals that states that adopted term limits and states who did not, have statistically identical expenditures per capita, both in the time period before any state had implemented term limits and in the short time afterwards. Cut points for separating the time period before and after term limits became popular are identical to those used in the original

21 When I separate term limits into house term limits and senate term limits, the simple term limits variable is eliminated.

22 Jeffrey S. Zax, "Initiatives and Government Expenditures," Public Choice 63 (1989).

paper. The complete results from the t-test as well as a visual representation of expenditure levels can be seen below. These findings quickly put to rest my suspicion that term limit adoption and spending levels are an endogenous institution. Figure 1, convincingly illustrates that term states (mainly the initiative-referendum states) have roughly equal expenditures per capita as those states who failed to adopt term limits in the sample frame.

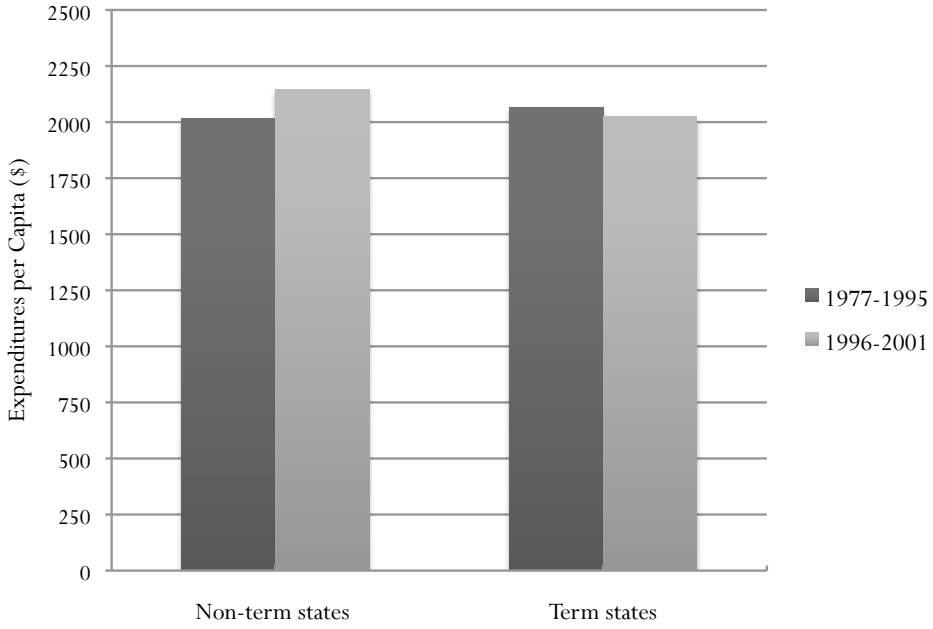
Table 1, Student's T-test, Expenditures per Capita for Treatment and Comparison Groups, 1977-1995

	N	Avg.	Std. Error	
Non-term states	703	2016.70	35.95	
Term states	190	2062.68	70.17	
Combined	893	2027.46	32.05	
Difference		-45.98	75.73	
				t ≈ -0.61
				df = 891

Table 2, Student's T-test, Expenditures per Capita for Treatment and Comparison Groups, 1996-2001

	N	Avg.	Std. Error	
Non-term states	222	2144.21	71.04	
Term states	60	2022.90	109.36	
Combined	282	2115.82	60.12	
Difference		121.31	142.07	
				t ≈ 0.85
				df = 280

Figure 1, Expenditures per Capita for Term Limit States and Non-term Limit States



In addition to t-tests in each time period, a difference in difference test was employed to establish whether the change in spending per capita was significantly different for the term limit states compared to those states who failed to implement term limit legislation. The t-statistic for the difference in difference model (implied in Figure 1) is -1.09, indicating that the change in expenditure levels between the treatment and control groups were not significantly different in each of the time periods. Therefore, we fail to reject the null hypothesis that changes in per capita expenditures between the two groups are equal.

Although t-tests are useful statistical measures, they are blunt. From t-tests we cannot rule out that changes in spending levels in each of the groups were from term limit laws. Observed changes in spending might have been more or less dramatic if other intervening variables were not in the picture. To hold other variables constant over the entire time period, multivariate regression analysis is necessary. The full results can be found in Table 3. The first column is simply copied from the original study for purposes of comparison; Column 1 is the column that I hope to replicate. The figures in Model 1 (column 2) represent my replication of Erler (2007) using the same variables and a nearly identical method of analysis. The only change made to the original method is the coding of term limits which has already been discussed. In Model 2, I drop the term limits variable in order to measure separate indicators of term limits in the state house and state senate. In model 3, I introduce two variables; a dummy variable for whether or not there is gubernatorial term limits and total state population. Column 4 is an attempt at using an instrumental variable approach. In this case, the term limits variable is actually the fitted values for term limits which were predicted using a dummy variable for the initiative process.

Table 3, Effects of Term Limits on State Expenditures per Capita, 1977-2001

Independent Variables	**Erler (2007)**	**1**	**2**	**3**	**4**
Constant	__	4618.8*** (69.9)	4629.1*** (70.3)	4243.8*** (39.6)	4.61*** (.05)
Term Limits	59.8** (20.9)	-25.8 (35.5)	__	__	-11.4 (8.2)
Unified Government	-42.6** (13.4)	-46.3*** (11.3)	-44.8*** (11.4)	-60.4*** (11.3)	-.04*** (.01)
Governor Party (Dem=1)	32.8** (12.7)	24.9* (12.0)	24.7* (12.0)	23.7** (10.7)	.02 (.01)
Grants	.830** (.115)	.00001* (.000005)	.000008 (000006)	.00001* (6.92×10^{-6})	1.05×10^{-8} (6.6×10^{-9})
Income²	.0000001** (0000001)	-2.5×10^{-22} (1.6×10^{-22})	-1.5×10^{-22} (1.9×10^{-22})	1.0×10^{-21}*** (3.0×10^{-22})	-2.4×10^{-25} (2.1×10^{-25})
Population Density²	.0008* (.0004)	.0017*** (.0005)	.0017*** (.0005)	.0016*** (.0004)	1.7×10^{-6}*** (3.3×10^{-7})
Unemployment Rate²	.276 (.392)	.23 (.36)	.28 (.36)	.60 (.42)	-.002 (.003)
House Limits	__	__	159.7** (52.7)	174.7** (64.4)	__
Senate Limits	__	__	-212.2** (67.1)	-218.3** (68.8)	__
Governor Term Limits	__	__	__	487.1*** (52.3)	__
Population	__	__	__	-.0001*** (.00001)	__
Observations	1175	1175	1175	1175	1175
Number of States	47	47	47	47	47
Number of Years	25	25	25	25	25
R^2	.94	.97	.97	.97	.96

Note: Cell entries are unstandardized regression coefficients with panel corrected standard errors in parenthesis. Erler column represents the data presented in Abbie H. Erler, "Legislative Term Limits and State Spending," *Public Choice* 133 (2007): 485. The Erler (2007) column corrects for first order autocorrelation, the remaining models use Durbin-Watson autocorrelation controls. All models include fixed effect dummy variables for state and year (not reported).

* = p<.05; ** = p<.01; *** = p<.001

The regression results provide several statistically and substantively significant findings. In Column 1, we can see that the term limits variable does not reach conventional levels of significance. This suggests that conclusions from the original study may have been biased either because of the coding of the term limits variable or other data problems inconsistencies. With the exception of income^2, results from Erler (2007) are noticeably similar to those produced here. The direction, magnitude and significance of most coefficients are comparable in Erler (2007) and my basic regression found in Column 1. These include unified government, the party of the governor grants, population density and the unemployment rate. Unified governments spend about $46 per capita less than states with divided government and Democratic governors appear to outspend Republicans by roughly $25 per capita. An additional 100,000 dollars in federal grant money corresponds to an average increase in spending by about 1 dollar per capita. Given that state population levels are all over 100,000, we might infer that an additional dollar from the federal government results in one more dollar of state spending. These results are intuitive and consistent with those of Erler (2007). Because the findings are comparable on every variable excepting only term limits, we might suspect that the coding scheme proposed be Erler can account for her divergent conclusion. The reader will notice that the r^2 jumps from .94 to .97 between Erler (2007) and Column 1. This provides, at least superficially, evidence that coding the term limits variable intuitively fits the data better than if it is lagged.

Perhaps the most considerable contribution to the study of term limits is featured in column 2. Here we see that dividing the term limits concept into house limits and senate limits has a profound effect on the regression. Both of these variables reach high levels of significance and surprisingly move in opposite directions. The implementation of house term limits increases spending by about $174 per capita. Conversely, senate term limits are associated with spending levels that are $218 per capita less, when other variables are held constant.

The findings presented in column 3 are equally stunning. The coefficients and significance levels are not drastically altered for house and senate term limits while on average, we see that gubernatorial term limitations increase spending by about $487 per capita, a finding that is statistically significant at the p=.001 level. The overall population of a state also reduces spending levels, thereby supporting the economy of scale argument. An additional 10,000 people living in a state corresponds with an average decrease is spending per capita by about $10. Although the coefficient does not make a strong statement, this finding is substantively significant considering that state populations vary considerably and it is not uncommon for state populations to fluctuate greatly (even by 100,000) in a given year. We should also note that income^2 becomes significant in this model although the coefficient is difficult to interpret and as I've already mentioned this variable is primarily a control.

Column 4 was an attempt at using an instrumental variable approach to deal with the potential endogeneity between state spending and term limit adoption however this model performs poorly in comparison to the others. The values for unemployment, grants, income and population density are noticeably different than we might expect and are difficult to interpret because they are so miniscule. The only variable that is significant in this model is

population density although it is unclear why its coefficient is radically different than in the other equations.

Although this model considerably weakens the strength of the term limits variable, this may indicate the relative exogeneity of term limit adoption with respect to the variables included here. Results from the first stage equation (not reported) suggest that initiative voting is not a strong predictor of term limit implementation. In fact, the initiative dummy variable was not even significant when it was regressed on term limits, calling into question the reliability of the predicted values. Our initial reasoning for pursuing this approach was influenced by our suspicion that higher spending levels drove term limit adoption. Figure 1 and corresponding t-tests indicate that this was not the case.

Taken together, the results presented here are markedly different than those generated by Erler (2007). Differences between the data used here and the data used in the original study might account for this discrepancy. Census data is regarded as the only reliable source for unemployment, population density, and the other control variables used in the model. Descriptive statistics in the Appendix suggest the next obvious place to look for data problems is in the outcome variable, where mean expenditures per capita are nearly 800 dollars different in each of the studies. Spending data on the state level is provided by the U.S. Census, however a variety of other institutions distribute these data publicly. Probably the best known resource for expenditure figures is the National Association of State Budget Officers who annually (occasionally biannually) publish a booklet called the Fiscal Survey of the States. The Fiscal Survey of the States (FSS) is available online from every year in the sample, excepting 1977, and includes multiple measures of spending, revenue on-hand, and other useful information on each U.S. state. Again, I should stress that I am unsure about original data used by Erler (2007). If she was able to find FSS data for the entire sample, it could potentially explain why we came to such divergent conclusions.

I collected the FSS data since my original objective was to use multiple indicators for the dependent variable.[23] However, because the first year of the data was missing, using the FSS data would have created problems in the regression and introduce questionable assumptions into the model. Instead, I decided to assess whether the Census and FSS were measuring the same underlying concept of spending. If the Census and FSS are comparable (or interchangeable) indicators of spending, we can rule out this explanation for the different conclusions reached in the original and replication study. Confirmatory factor analysis is the conventional method to discern how many underlying measures indicators are tapping, so it is seems like the appropriate tool to use in this case.

Results from the matrix (not shown) suggest that among these two variables, measured from 1978 to 2001, only one principal factor emerges. We can safely assume that this factor is state-year expenditures. The eigenvalues are 1.93 for the first factor and -.02 for the second factor. The factor loadings, which are analogous to an r term, are roughly .98 for both the

23 Multiple indicators is always an effective way to improve robustness and augment internal validity. If the empirical story is consistent with multiple specifications, we can begin to evaluate the validity of the causal claim.

Census and FSS data. Therefore we reject the null hypothesis that the outcome variables are measuring two distinct concepts. In addition, these two indicators have a correlation of .97, further evidence that FSS and Census measures approach parity. If Erler (2007) used FSS data instead of the Census measures of expenditure employed here, it would not account for the contradicting conclusions.

Discussion

The results presented here make a strong case for revising our understanding of the relationship between term limits and state spending. Replication indicates that Erler's (2007) conclusions regarding the relationship between legislative term limits and expenditures was based, at least partially, on a coding scheme that accounted for term limits when they didn't exist. Improvements to the original model reveal several important nuances that should be incorporated into our knowledge of term limiting institutions. First, we learned that house term limits and senate house limits have significant effects on statewide spending although in different ways. Term limits in the state house increase spending, on average, while the same institutional constraint in the senate has the opposite effect. This finding is surprising at first glance however the fact that each chamber has different election cycles and varies greatly in membership size may indirectly explain why tenure limitation had unique effects on each state house. The short story is that the two chambers really *are* unique (and intentionally so) which conditions how each of its members behave in electoral and legislative settings.[24]

Most states who implemented term limits set a ceiling on years instead of terms. In most states members are allowed 8 years total, regardless of whether a member is from the senate serving 4 year terms or a house member with a two year term. Because state senator term lengths are at least twice as long, members elected to this house have fewer elections before they are term limited. It is likely that bicameral design interacts with tenure limits to have the impact we observe. Although more research might help uncover why the house and senate respond differently to term limits, this findings is a major improvement upon earlier work.

Another considerable conclusion we can draw from this study is that governors who are term limited spend considerably more money, on average, than governors who are not. The reason for this result is unclear but it likely has to do with the executive nature of the governor's role in state government. Although the governor is technically not responsible for the creation of "good public policy" during her tenure, it is likely that the legacy of governors is far more important than state representatives given that it is a much more prominent post. The gubernatorial term limit may cause governors to adopt an aggressive policy agenda so that they will build a favorable reputation when they leave office. This would be especially important for governors that are term limited because they will inevitably have to find a career

24 Unfortunately, there is a dearth of literature on state bicameralism. At the state level, chamber differences are much less understood than those of the US Congress, and a great deal less developed theoretically. Because of this gap, possible reasons for the house-senate disparity observed here (post term limit) is speculative at best. This present study brings us closer to acknowledging that state house and state senate differences are substantive while the mechanisms behind these differences remain tenuous.

in the private sector, a reality they are aware of on day one.[25]

Our empirical finding that gubernatorial term limits are associated with higher levels of spending is perhaps the most reliable finding given that governor term limits are built into state constitutions (or adopted shortly thereafter). In other words, gubernatorial term limits are as exogenous as statehood; that is, the reasons that some states have gubernatorial term limits while some do not has absolutely nothing to do with state spending. With respect to the adoption of *legislative* term limits, we still face an endogeneity question which should caution the inferences drawn these other term limit variables.[26]

Lastly, and not surprisingly, these results indicate that state governments operate under the "economies of scale" principle whereby population growth increases the efficiency of expenditures. Interestingly enough, increasing population density *increases* spending while increasing overall population decreases spending. Taken together, these results probably indicate that there exists an optimal economy of scale whereby population and area are proportional to the public goods that the state chooses to invest in. Although the population variable was not central to this analysis, it is an important improvement to the original model of Erler (2007).

This study underscores the price that political science can pay when work goes unchecked or non-replicated. In this case, I found no bases for the central finding in the original paper presented by Erler (2007). Legislative term limits do not significantly increase (or decrease) state expenditures on average, except when the concept is broken out and applied to each chamber of the state legislature independently. Therefore, the effects of term limitation are highly dependent on the government institution that they are imposed upon. From a policy perspective, these results raise questions about the ability of term limits to control spending at the state level given the competing theories of legislator motivation provided here.

25 Of course some governors are already wealthy when they assume office, are retired, or plan on serving in another public office. These probably are the exceptions to governors who must look for work after leaving the governor's mansion.

26 Future research would do well to find instruments that predict the adoption of term limits but are unrelated to spending.

REFERENCES

Alt, James E. and Robert C. Lowry. "Divided Government, Fiscal Institutions, and Budget Deficits: Evidence from the States." *American Political Science Review* 88. 4 (1994): 811–28.

Bandow, Doug. "Real Term Limits: Now more than Ever." Cato Institute's *Policy Analysis* 221 (1995).

Besley, Timothy and Anne Case. "Does Electoral Accountability Affect Economic Policy Choices? Evidence from Gubernatorial Term Limits." *Quarterly Journal of Economics* 110.3 (1995): 769–98.

Carey, John M.; Richard G. Niemi; Lynda W. Powell and Gary F. Moncrief. "The Effects of Term Limits on State Legislatures: A New Survey of the 50 States." *Legislative Studies Quarterly* 31.1 (2006): 105–34.

Erler, Abbie E. "Legislative Term Limits and State Spending." *Public Choice* 133 (2007): 479–94.

Fenno, Richard F. *Home Style: House Members in their Districts*. New York: Pearson Longman, 1978.

King, Gary. "Replication, Replication." *Political Science and Politics* 28.3 (1995): 444–52.

Lopez, Edward J. "Term Limits: Causes and Consequences." *Public Choice* 114 (2003): 1-56.

Reed, Robert W. "Democrats, Republicans, and Taxes: Evidence that Political Parties Matter." *Journal of Public Economics* 90 (2006): 725–50.

Reed, Robert W.; Eric D. Schansberg; James Wilbanks and Zhen Zhu. "The Relationship between Congressional Spending and Tenure with an Application to Term Limits." *Public Choice* 94 (1998): 85–104.

Rogers, Diane Lim and John H. Rogers. "Political Competition and State Government Size: Do Tighter Elections Produce Looser Budgets?" *Public Choice* 105 (2000): 1–21.

Stigler, George J. "The Economies of Scale." *Journal of Law and Economics* 1(1958): 54–71.

Will, George F. *Restoration: Congress, Term Limits, and the Recovery of Deliberative Democracy*. New York: Free Press, 1992.

Zax, Jeffrey S. "Initiatives and Government Expenditures." *Public Choice* 63 (1989): 267–77.

Appendix

Table 4, Summary Statistics

Variable	Author	Mean	St. Dev	Min	Max
Unified Gov't (=1)	Erler Klepetar	.46 .44	.499 .49	0 0	1 1
Expend. per Capita	Erler Klepetar	2784 2049	707.6 970.8	1016 507	5220 5229
Gov. Party (Dem=1)	Erler Klepetar	.547 .537	.498 .494	0 0	1 1
Gov. Limits (=1)	Erler Klepetar	.274 .723	.446 .448	0 0	1 1
Income[2]	Erler Klepetar	6.09×10^{8} 2.64×10^{22}	2.30×10^{8} 9.07×10^{22}	2.31×10^{8} 9.17×10^{18}	1.82×10^{9} 1.29×10^{24}
Density[2]	Erler Klepetar	84057 85290	2117724 221517	2.57 18.09	1272479 1310453
Term Limits (=1)	Erler Klepetar	.053 .029	.224 .167	0 0	1 1
Unemployment Rate[2]	Erler Klepetar	41.1 41.1	30.8 30.9	4.84 4.84	324 324