NEIL MALHOTRA
Stanford University

Disentangling the Relationship between Legislative Professionalism and Government Spending

Recent movements to deprofessionalize American state legislatures have been driven partly by the notion that professional legislators spend more than their citizen counterparts. This article explores the relationship between legislative professionalism and government spending, a connection complicated by the possibility that legislators in high-spending states may choose professional institutions to handle their responsibilities more effectively. I employed propensity score matching, an increasingly used technique of causal inference, to disentangle the relationship. Contrary to previous academic work and popular notions, I found that professional legislatures do not spend significantly more than part-time bodies do, if one accounts for the fact that legislatures in high-spending states have a greater need to be professionalized and therefore select those structural frameworks. These findings have important implications for the study of the effects of legislative institutions on public policies more generally and attest to the utility of recently developed techniques of causal inference to disentangle these relationships.

One of the most important institutional developments in American state legislatures is the professionalization revolution. Salaries, session lengths, and staff resources all substantially increased in the 1960s and 1970s. A major, yet inadequately answered question is, What has been the effect of these structural changes on state public finance? Some scholars have argued that professional legislators, who are career politicians, engage in greater and more-inefficient spending than do their counterparts in citizen legislatures. Professional state legislators are presumed to be subject to the same incentive structures that facilitate logrolling and particularistic spending in the highly professional U.S. Congress. This conception may explain the professionalization plateau of the 1980s (Squire and Hamm 2005) and the popular movements to deprofessionalize legislatures via term limits (Kousser 2005) and the expansion of direct democracy (Matsusaka 2004) in the 1990s.

There has been much popular speculation about the effect of professionalism on budget sizes and legislative efficiency. After California state lawmakers frustrated the governor's efforts to pass workers' compensation reform in 2004, Governor Arnold Schwarzenegger proposed making the legislature, the most professional in the country, into a part-time body.[1] In congressional testimony, Stephen Moore of the Cato Institute argued that the U.S. Congress should become less professional, citing anecdotal evidence that the highly professional Massachusetts legislature imposes a higher tax burden on citizens than the nonprofessional New Hampshire legislature does (U.S. Congress 1996). Editorial writers have also speculated on the subject: "Do citizens suffer when politicians only spend part of their time making laws? In Virginia [a hybrid legislature poised between professional and citizen status], it's probably not a coincidence that we have low taxes and very little regulation."[2]

Scholarly evidence to support or rebut these claims is sparse. The main obstacle to empirical analysis of the effect of professionalism on spending is the possibility that high-spending states may *choose* to be professional. Some evidence suggests that states construct professional legislatures to cope with the increasing demands of a rapidly growing public sector (Malhotra 2006). In other words, states may not be randomly assigned to receive the "treatment" of professionalism. In this article, I attempt to clarify the relationship between legislative professionalism and government spending in the American states using techniques of causal inference. The results have important implications not only for state politics but also for our broader understanding of the effects of institutions on policy outputs.

Professionalism has been used as an independent variable to explain a host of dependent variables, including divided government and partisan composition (Fiorina 1994; Squire 1997; Stonecash and Agathangelou 1997), interest group activity (Berkman 2001), membership diversity (Squire 1992), policy responsiveness (Maestas 2000), gubernatorial effectiveness (Dilger, Krause, and Moffett 1995), incumbent reelection (Berry, Berkman, and Schneiderman 2000), congressional candidacies (Berkman 1993, 1994; Berkman and Eisenstein 1999), education policy (Sabloff 1995), and membership stability (Squire 1988). Surprisingly, there has been limited attention to professionalism's influence on fiscal policy. In an early study, Carmines (1974) found that professionalism mediated the effect of interparty competition on welfare spending. Parties with comparable electoral strength within a state must compete for votes among lower socioeconomic groups, and this translation of interparty conflict to public policy is presumably stronger in states where legislators and parties are more organized within the chamber. In the most recent analysis of the topic, Owings and Borck (2000) found a positive association between professionalism and spending, and claimed that citizen legislators spend less. Their results are mainly correlational, however, and do not delve closely into concerns about nonrandom assignment.[3] In a related study, Carey, Niemi, and Powell (2000) found that professional legislators enjoy greater incumbency advantage, perhaps because of their ability and incentives to secure more spending on behalf of their districts.

Several studies have not considered professionalism explicitly but have examined the power of entrenchment for fiscal outcomes. For instance, a model by Herron and Shotts (2006) delineated conditions under which term limits prevent voters from selecting the representatives who are best at delivering particularistic benefits. Similarly, Carey, Niemi, and Powell (1998) and Carey et al. (2006) found that legislators who are term limited spend less time worrying about delivering pork to their districts. Deprofessionalizing legislatures via term limits may therefore reduce spending. Similarly, direct democracy tools—such as initiative and referenda, which have produced fiscal reforms (for example, tax and expenditure limits)—presumably move legislative and budgetary processes away from professional politicians and toward the citizenry. Matsusaka (1995, 2004) observes that states where citizens can initiate and approve laws by popular vote spend less than do states where only legislators have that ability. Finally, restraints on legislators' fiscal capacities have been found to hinder spending, as well (Primo 2006). Many observational studies have produced results suggesting a link between professionalism and spending.

For this article, I used propensity score matching to assess the effect, if any, that professionalism has on spending beyond the selection effects already described (Malhotra 2006). I compared legislatures that were equally likely to be professional, some of which received the "treatment" of professionalism while others did not. This approach limited my examination to cases on "common support," excluding those that could not be matched according to a set of observable covariates and were therefore not true counterfactuals.

Using this technique, I found evidence contradicting existing scholarly work and popular conceptions of the effects of legislative professionalism. Once I accounted for the fact that legislatures in high-spending states have a greater need to be professionalized (and thus they select into those institutional frameworks), I found that professional legislators spend no more than their counterparts in unprofessionalized bodies. I estimated treatment effects with King's (2000) and Gilligan and Matsusaka's (1995) pooled dataset of four legislative sessions across ten-year intervals: 1963–64, 1973–74, 1983–84, and 1993–94. These are the same data used by Owings and Borck (2000), allowing direct comparability. My analysis not only addresses the question of whether or not professional legislatures are more spendthrift, but also presents a method with the broad application of disentangling the relationships between other institutional structures and policy outcomes, which are often prone to self-selection into treatment and control.

The article is organized as follows. Section 1 presents a brief theoretical overview of competing arguments as to the effect of professionalism on government spending. In Section 2, I describe the statistical methods and the components of the empirical models. In Section 3, I present the results of these tests and provide analysis. I conclude with a discussion of the implications of this research and possible extensions.

1. Overview

In their article positing that citizen legislators spend less money than professional members spend, Owings and Borck (2000) cite several studies from the political science and economics literatures in support of their hypothesis. I will only briefly summarize those studies here. They argue that professional legislators have greater means and incentive to engage in pork-barrel spending. Reed and Schansberg (1996) have suggested that since logrolling is a repeated game, it is more feasible in professional legislatures, where there is less turnover and members meet for longer periods of time. Moreover, logrolling requires political skill (Coates 1999), so professional legislators have greater abilities to engineer deals that result in pork-barrel spending. And because seats in professional bodies are more valuable than seats in citizen legislatures, professional members have a greater reelection-incentive and, consequently, greater motivation to secure pork-barrel projects to gain electoral support from constituents (in the form of votes) and interest groups (in the form of campaign contributions) in their districts (Owings and Borck 2000). Finally, Fiorina (1994) has reported that professional legislatures are more likely to attract Democrats and, hence, would be filled with members who ideologically prefer greater public spending.

These arguments also imply that spending growth induced by professionalism should be accompanied by offsetting tax increases.

Primo (2006) and Bohn and Inman (1996) have found that budget constraints often bind in state legislatures, with taxes and expenditures moving together. Like securing spending projects, passing tax bills requires time and skill. Yet spending can often be targeted to individual districts, whereas the tax burden is generally shared, producing a common-pool problem wherein legislators have an incentive to overspend and overtax (Chen and Malhotra 2007; Weingast, Shepsle, and Johnsen 1981). Professionalism may exacerbate inefficiencies arising from the common pool.

Owings and Borck also cite several studies that imply that citizen legislators should spend *more* than professional ones. Building on Reed et al.'s (1998) study of interest group lobbying in Congress, Owings and Borck speculate that, because citizen legislators are more reliant on groups for information and research, they will be more susceptible to these groups' demands for increased particularistic spending.[4] Further, because professional legislators are more secure in their reelection (Carey, Niemi, and Powell 2000), they have less incentive to engage in logrolling to curry favor back home (Fiorina 1996).

The most troublesome argument against the hypothesis that professionalism increases spending (at least for the purposes of empirical testing) is that legislatures faced with large budgets simply select more full-time responsibilities. Building on Fiorina and Noll's (1978) conception of legislators as "ombudsmen to the bureaucracy," I have previously argued (Malhotra 2006) that the demands of reelection and constituency service compelled members to respond to the burgeoning state budgets of the postwar period by professionalizing. Using a longitudinal analysis and instrumental variables regression, I found empirical evidence that increased spending causes professionalism, independent of any differing incentives between professional and citizen legislators. This study builds on my previous results by examining if this nonrandom assignment accounts for the observed correlation between professionalism and spending in the American states.

2. Methods and Measurement

In this section, I discuss the methodology of the statistical tests as well as the measurement of the principle variables of interest (professionalism and spending) and relevant covariates. To provide a baseline for comparison, I replicated Owings and Borck's (2000) constant-elasticity model of state spending, which is estimated with standard least squares regression. I estimated parameters from both models with pooled historical data (Gilligan and Matsusaka 1995; King

TABLE 1
Descriptive Statistics

Variables	Mean	Std. Dev.	Min	Max
Professionalism				
Log Squire Index	–1.65	.54	–3.08	–.31
Kurtz-Squire Dummy	.52	.50	0	1
Session Length Restriction	.68	.47	0	1
Spending				
Log Total Expenditures	7.09	.41	5.97	8.05
Log Social Services Expenditures	6.61	.48	5.29	7.51
Log Non–Social Services Expenditures	6.09	.39	5.25	7.30
Control Variables and Covariates				
Democratic Control	.41	.50	0	1
Divided Control	.27	.44	0	1
Log Revenue from Federal Government	5.92	.47	4.54	6.96
Log Mineral Revenue	5.44	1.61	1.36	9.46
Log Population	14.87	1.01	12.75	17.26
Log Personal Income	9.29	.26	8.48	9.88
Population Growth	.06	.05	–.03	.42
Percentage Metropolitan Population	.61	.24	0	1
Population Heterogeneity	.47	.05	.33	.59
Gubernatorial Power	14.65	3.10	7	20
Opportunities to Advance	.005	.007	0	.053

Note: N = 191. All monetary figures are normalized on a per capita basis. See appendices for measurement details.

2000) from four legislative sessions: 1963–64, 1973–74, 1983–84, and 1993–94. I adjusted all dollar figures for inflation and measured them on a per capita basis. Measurement details and data sources are provided in Appendix 1. Technical details, such as missing data and nonspherical errors, are discussed in Appendix 2. Descriptive statistics appear in Table 1.

Constant-elasticity Model

As the benchmark analysis, I used Owings and Borck's (2000) main empirical specification, derived from a constant-elasticity model of government spending and simplified in this presentation. Using the Cobb-Douglas functional form, Owings and Borck define spending (S_{it}) as a production function of professionalism (P_{it}) and a vector of economic, social, and institutional control variables ($\mathbf{x}_{it}$) for each state i at time t:

$$S_{it} = e^{\alpha} P_{it}^{\beta} \mathbf{x}_{it}^{\gamma} e^{\mu} . \tag{1}$$

Taking the natural logarithm of both sides of equation (1) yields

$$ln\ S_{it} = \alpha + \beta(ln\ P_{it}) + \boldsymbol{\gamma}(ln\ \mathbf{x}_{it}) + \mu_{it}\ , \tag{2}$$

where μ_{it} represents a disturbance term of the form $\mu_{it} = \lambda_t + e_{it}$, and λ_t represents a year fixed-effect, and e_{it} is random error with mean 0.[5] This log-log specification allows one to interpret the β parameter as representing a β% change in spending associated with a 1% increase in professionalism.

Propensity Score Matching

The principal method used to deal with nonrandom assignment is propensity score matching, which allows estimation of the treatment effect of professionalism on spending via comparison of states that were about equally likely to select similar legislative structures. Matching requires conceiving of professionalism as a dichotomous treatment, which I will later discuss in further detail. Following Holland (1986), I define the treatment effect (τ_i) as the difference in government spending (S_{it}) for a state under two conditions, professional (p) and nonprofessional (u):

$$\tau_i = S_{itp} - S_{itu}. \tag{3}$$

Unfortunately, for a given state, we cannot simultaneously observe spending under both conditions (the counterfactual is exactly that—what did not occur). If, however, we could randomly assign a population of states to be either professional or nonprofessional, then we could calculate the average treatment effect (ATE), which is the expected value of the difference between the professional and non-professional conditions:

$$\text{ATE} = E(S_{it} \mid P_{it} = 1) - E(S_{it} \mid P_{it} = 0). \tag{4}$$

Here, P_i represents assignment to the professional condition (1) or the nonprofessional condition (0). In expectation, all states are identical, but some are assigned to the treatment and the others are assigned to the counterfactual control. Obviously, state legislatures are not randomly assigned to varying levels of professionalism, which is the implicit assumption of correlational studies that make causal claims.

Nevertheless, matching attempts to proxy the optimal experimental paradigm of random assignment by comparing samples of states that are (nearly) identical on all observable covariates relevant to spending except for one: professionalism (Rosenbaum and Rubin 1983). In other words, the goal is to match states (professional versus nonprofessional) that were equally likely to receive the treatment of professionalism and compare state spending levels. I calculated the likelihood of receiving the treatment with propensity scores (p_{it}), that is, the predicted probabilities of a logistic regression predicting professionalism, the dichotomous treatment variable, with the set of observable covariates ($\mathbf{x}_{it}$):

$$\Pr(P_{it} = 1 \mid \mathbf{x}_{it}) = \Lambda(\boldsymbol{\gamma}\mathbf{x}_{it}). \tag{5}$$

$\Lambda(\cdot)$ is the logistic cumulative distribution function.[6]

I matched pairs using the "nearest neighbor" method, in which each professional state is matched to a nonprofessional state that has the closest value of p_{it}. Following Dehejia and Wahba (2002), I matched "with replacement," meaning that multiple professional states could be matched to multiple nonprofessional states. All states not on common support (that is, nonprofessional/professional states with propensity scores lower/higher than the lowest/highest propensity score for a professional/nonprofessional state) were dropped from the analysis, so that true counterfactuals were compared. Using the matched sample, I then tested for balance, ensuring that the means of $\mathbf{x}_{it}$ were similar between professional and nonprofessional states and that I was therefore comparing states that were equally likely to receive the treatment. Once the sample was verified to be balanced, I estimated ATE, weighting by the number of times a state was matched. I estimated standard errors of ATE via bootstrapping (Dehejia and Wahba 2002).

Propensity score matching offers two main advantages over standard regression techniques in dealing with the problems of causal inference described earlier. First, matching increases the comparability between professional and nonprofessional states by aligning their distributions of observables. Least squares regression does not have this balancing feature but "instead controls for observable differences between the treatment and control groups by assuming the conditional mean function is correctly specified by the regression equation" (Behrman, Cheng, and Todd 2004, 115). Thus, unlike matching, regression relies on strict functional-form assumptions. Second, matching only compares groups in the area of common support, which restricts attention to the observations that can serve as true counterfactuals. In contrast, "regression estimators typically use all the observations in estimation . . . and extrapolate over any regions of $\mathbf{x}_{it}$ where the supports do not overlap" (Behrman, Cheng, and Todd 2004, 115).

One potential problem with propensity score matching is that selection into the treatment may affect the outcome variable beyond what can be predicted from the observables. In other words, there may be unobserved variables that influence assignment. In this case, selection effects may remain and there may be little difference between matching and standard least squares regression (Arcenaux, Gerber, and Green 2006). If, however, we show that matching on *some* observable characteristics eliminates the correlation between spending and professionalism, then we reveal the fragility of the relationship, which is likely to be biased.

Measuring Professionalism (P_{it})

Measuring a legislature's level of professionalism is difficult because there is no obvious or direct statistic that can quantify this qualitative property. Within the expansive literature, legislative professionalism has been measured in myriad ways. Owings and Borck (2000) used the Squire index (Squire 1992), which averages three generally accepted proxies of professionalism: legislator compensation, the amount of time the legislature spent in session, and the amount of resources available to the member (staff, operating budgets, and so on). To standardize these scores over time, scholars calculate the professionalism of each legislature as a percentage of Congress's level of professionalism. Although the highest possible value of the index is, technically, infinity, the index generally lies between 0 (a completely unprofessionalized legislature) to 1 (a legislature as professionalized as Congress). The Squire index is the most commonly used measure of professionalism in the literature (Squire and Hamm 2005); consequently, using it in my current analysis allows for direct comparability to existing research. Finally, the Squire index is easily and intuitively interpreted; it measures how professional a legislature is compared to the U.S. Congress, the most professionalized body in the world.[7]

A limitation of the Squire index when used for propensity score matching is that the index is continuous, whereas matching requires a dichotomized treatment.[8] Kurtz (1992) divided state legislatures into three categories (professional, hybrid, citizen), taking into account legislator salaries, session lengths, staff resources, and interviews with the legislators themselves. Unfortunately, the Kurtz index only exists for the 1993–94 session. One straightforward way of bifurcating states from previous periods is to estimate a logistic regression predicting professional or hybrid status with the Squire index (P_{it}) for the 1993–94 session:

$$\Pr(K_{it} = 1 \mid P_{it}) = \Lambda(\beta P_{it}), \tag{6}$$

where K_{it} is coded as 1 if legislature i is categorized by Kurtz as professional or hybrid and as 0 if the legislature is categorized as citizen.[9] The regression estimates allow one to predict legislatures' probabilities of being professional/hybrid for the entire historical sample.[10] If the probability was greater than 50%, then I classified the legislature as professional.[11] In subsequent analyses, I refer to this variable as the "Kurtz-Squire Dummy."

To assess the robustness of the results, I used another measure of professionalism that is truly dichotomous: the presence of a session length restriction. Previous studies (King 2000; Mooney 1995) have observed that legislatures with constitutional or statutory restrictions on session length are more likely to be nonprofessional. We can intuitively see that, when the legislature can technically meet for an unrestricted amount of time, session lengths will increase, salaries will rise commensurately, and staffs will be expanded to deal with the heavier workload.[12] Indeed, the polychoric correlation between the Squire index and the session-length-restriction dummy in these data is –.64, indicating that states with restrictions are less likely to have high salaries, long sessions, and noteworthy staff resources. Consequently, I also measured professionalism with a dichotomous measure of whether or not the state has a restriction on session length, which shares a tetrachoric correlation with the Kurtz-Squire dummy of –.71. To compare the matching results to the other methods, I also estimated equation (2) with the Kurtz-Squire dummy and an indicator variable representing a session length restriction. Thus, a session length restriction can be conceived of as a true treatment: How would spending increase if a state lifted its constitutional or statutory restriction on session length?

Measuring Spending (S_{it})

I measured government spending straightforwardly as the expenditure level of the state that is produced by a legislative session, adjusted for inflation and normalized on a per capita basis. For example, the

1964 budget is produced by the 1963–64 legislative session. Much of the existing theory posits that professionalism increases capital expenditures in the form of pork, not general public-welfare programs. Therefore, I divided state budgets into two categories: social services expenditures and non–social services expenditures. Social services expenditures are the sum of the state's spending on education, public welfare, hospitals, health, and employment security services. Non–social services expenditures comprise budgetary line items not directly and specifically related to social services (corrections, police, general financial administration, highways, natural resources).

Covariates of Spending and Professionalism (x_{it})

The control variables for the models predicting spending (and the covariates used to calculate propensity scores) come from Owings and Borck (2000), who use the specifications of Gilligan and Matsusaka (1995). The theoretical and empirical bases of these variables can be found in Rubinfeld's (1987) theory of public finance, which suggests that there are three main determinants of state government spending: socioeconomic characteristics, sources of revenue, and political variables. Detailed justifications for the use of these controls can be found in the original papers, but I will briefly summarize the rationales here. I adjusted all dollar figures for inflation and normalized them on a per capita basis.

Four socioeconomic variables are included in the regression models: Population, Population Growth, Percentage Metropolitan Population, and Per Capita Personal Income. Higher-population states may require greater spending because they are very complex to govern. There also may exist economies of scale in government spending that produce a negative effect of Population. States that have experienced significant, recent population growth may face a short-run demand for public services. Moreover, states that are more metropolitan may face different needs than more-suburban or rural states, since urban constituents often require more social services and capital projects. Finally, wealthier states may be able to spend more because of greater tax revenues and looser budget constraints.

States may also have varying sources of revenue, which may affect spending levels. Revenue from the federal government in the form of grants generates wealth effects that may allow more spending. Mineral revenue must also be considered because states with large mineral deposits can levy severance taxes on nonresidents, thereby generating funds without placing stress on the tax base. Finally, I included dummy variables for Democratic-controlled and divided governments because Republican governments (the baseline category) are presumed to be more fiscally conservative.

To improve balance when performing propensity score matching, I also included variables known to influence professionalism in equation (5). I borrowed two sets of variables from Mooney's (1995) and King's (2000) models of state legislative professionalism: socio-economic factors and structural characteristics. First, as the citizenry becomes more diverse and heterogeneous, the legislature must serve many different groups, necessitating professionalization. Therefore, I included the Sullivan index of diversity, which measures the probability that two randomly selected individuals from a state will differ along various demographic characteristics. Two structural features of state government could also have an effect on professionalization: gubernatorial power and opportunities to advance. An expansion of gubernatorial power may impel the legislature to become more professional in order to provide an effective check against the executive branch. Power can be expanded in a variety of areas, including appointments, vetoes, budget making, and length of tenure. Accordingly, I used the Schlesinger/Beyle index of gubernatorial power, which is constructed from these characteristics. Additionally, professionalized legislatures are expected to exist in states where there are many opportunities for legislators to advance to higher office. These state houses serve as training grounds for career politicians hoping to move up the occupational ladder. I used Maestas's (2000) measure of advancement opportunities, which is a function of the number of House seats in a state, how often those seats turn over, the percentage of those seats held by former state legislators, and the total number of state legislators (see Appendix 1). An increase in opportunities for advancement should be associated with higher levels of professionalism. Lastly, following Bailey (2006), I included fixed-effects for year in the propensity score equation.

3. Results

Whereas the traditional regression technique finds a positive and significant relationship between professionalism and spending, a method explicitly designed for causal inference detects no significant association. These results suggest that previous studies may have failed to account for the fact that states nonrandomly select professional legislative structures.

Constant-elasticity Model

First, I replicated Owing and Borck's (2000) constant-elasticity model to provide a baseline with which to compare future analyses.[13] In Table 2, I present parameter estimates for equation (2) using three dependent variables (total expenditures, social services expenditures, and non–social services expenditures) and three measures of professionalism (the continuous Squire index, the Kurtz-Squire dummy, and session length restriction).

As seen in specification (1) of Table 2, professionalism is strongly related to total expenditures (b = .10, $p < .001$), with a 1% increase in professionalism associated with a .1% increase in spending. Similar findings emerge when one dichotomizes professionalism, as in specifications (2) and (3). Being a professional state, as classified by the Kurtz-Squire dummy, corresponds with a 5% increase in per capita spending, as compared to a state with a citizen legislature. The presence of a session length restriction, a sign of a nonprofessional legislature, decreases spending by approximately 7%. When one examines budgetary line items with respect to whether or not they encompass social services, as in specifications (4)–(9), one finds the Squire index again strongly associated with spending. The dichotomized measures only have a significant effect on non–social services expenditures, however, a result consistent with the idea that professionalism increases targetable spending typically associated with logrolling. These results demonstrate a strong correlation between professionalism and spending, even when one controls for socioeconomic factors known to affect budget sizes. Still, as explained previously, these findings may be the consequence of high-spending states selecting professional legislative institutions in order to handle more effectively the demands of a large public sector.

With respect to the control variables, one sees significant relationships in the hypothesized directions. Compared to Republican-controlled governments, Democratic governments spend 4% to 6% more on total spending and 5% more on social services spending. Interestingly, the non–social services spending typically associated with pork appears to be nonpartisan. Moreover, states that receive more revenue from the federal government also have larger per capita budgets, partly because of the availability of matching grants. A 1% increase in federal aid is associated with an approximately .54% increase in total expenditures. State wealth, as measured by personal income, affects non–social services spending but not social services spending, a result perhaps indicating that richer states have greater tax

TABLE 2
Constant-elasticity Models Predicting Government Spending with Professionalism

Dependent Variable	Log Total Expenditures (1)	(2)	(3)	Log Social Expenditures (4)	(5)	(6)	Log Non-Social Expenditures (7)	(8)	(9)
Log Squire Index	.10*** (.02)	—	—	.10* (.03)	—	—	.09** (.04)	—	—
Kurtz-Squire Dummy	—	.05* (.02)	—	—	–.01 (.03)	—	—	.06+ (.03)	—
Session Length Restriction	—	—	–.07** (.02)	—	—	–.03 (.03)	—	—	–.09** (.03)
Democratic Control	.05* (.02)	.04+ (.02)	.06** (.02)	.05** (.02)	.05** (.02)	.05* (.02)	.00 (.03)	.00 (.03)	.02 (.03)
Divided Control	.01 (.03)	.00 (.03)	.01 (.03)	–.01 (.02)	–.01 (.03)	–.01 (.02)	.03 (.03)	.02 (.03)	.03 (.03)
Log Federal Revenue	.53*** (.04)	.54*** (.04)	.54*** (.04)	.46*** (.05)	.46*** (.06)	.46*** (.06)	.53*** (.06)	.53*** (.06)	.54*** (.06)
Log Mineral Revenue	.00 (.01)	.00 (.01)	.00 (.01)	.01 (.01)	.01 (.01)	.01 (.01)	–.01 (.01)	–.01 (.01)	–.01 (.01)
Log Population	–.08** (.02)	–.06*** (.01)	–.07*** (.02)	–.03 (.02)	.00 (.02)	–.01 (.02)	–.16*** (.02)	–.14*** (.02)	–.15*** (.02)
Log Personal Income	.36*** (.09)	.38*** (.09)	.40*** (.09)	.07 (.07)	.15 (.10)	.14 (.10)	.73*** (.13)	.76*** (.13)	.77*** (.13)
Population Growth	–.10 (.20)	–.11 (.21)	–.11 (.21)	–.39+ (.22)	–.38+ (.21)	–.37+ (.21)	.19 (.30)	.19 (.31)	.21 (.30)
Percentage Metro. Pop.	.12+ (.07)	.13+ (.07)	.15* (.07)	.27*** (.05)	.25*** (.06)	.26*** (.06)	–.07 (.11)	–.07 (.11)	–.03 (.11)
Constant	1.82* (.87)	1.03 (.87)	.99 (.85)	3.30*** (.55)	1.89** (.71)	2.18** (.69)	–1.05 (1.36)	–1.73 (1.32)	–1.65 (1.30)
R^2	.92	.92	.92	.83	.93	.93	.92	.92	.93

***$p < .001$; **$p < .01$; *$p < .05$; $^+p < .10$ (two-tailed).
Note: Standard errors in parentheses. N = 191 for all regressions. Panel-corrected standard errors (PSCEs) assuming correlated panels for equations (4), (5), and (6). PSCEs assuming independent panels for equations (7), (8), and (9). Year fixed-effects included but not reported.

bases but also exhibit less demand for social services. Further, population is negatively associated with non-social services spending, suggesting economies of scale in capital projects in that larger states require less per capita spending. Conversely, population is unrelated to per capita spending on social services because benefits often increase linearly with the number of recipients. Finally, states with higher metropolitan areas exhibit higher social services spending but are no more likely to spend more on non–social services items, perhaps because urban constituencies have greater demand for government programs.[14]

Propensity Score Matching

One way to account for nonrandom assignment is to match states that were equally likely to select various levels of professionalism and compare expenditure levels for legislatures that were professional to those that were not. To do so, I constructed a measure of the likelihood (that is, propensity) of being a professional legislature, or the predicted probabilities from logistic regressions predicting the Kurtz-Squire dummy and session length restriction with a set of covariates of professionalism and spending. The results of these regressions appear in Table 3.[15]

Using the algorithm described earlier, I then eliminated state-years not on common support, and I matched professional and nonprofessional states that had the closest likelihoods of receiving the treatment, allowing replacement. With the Kurtz-Squire dummy used as a measure of professionalism, 26 treated and 10 untreated observations are off common support, leaving 155 matched observations. For the session-length-restriction measure, 24 treated and 19 untreated observations are off common support, leaving 148 matched observations.

Before estimating the treatment effects, one should determine if the matched sample has achieved balance. Table 4 presents the means of the covariates for professional and nonprofessional states for both the unmatched and matched samples. Prior to matching, there was clearly extreme imbalance pointing to nonrandom assignment. For both treatment variables (Kurtz-Squire dummy and session length restriction), significant differences exist between the treatment groups along almost all of the covariates. For example, professional states are more likely to have higher populations, richer and more-diverse populations, more aid from the federal government, more-powerful governors, and more advancement opportunities available to legislators. Hence, professionalism is clearly not randomly assigned, and it would be

TABLE 3
Logistic Regressions Predicting Professionalism

Variables	Kurtz-Squire Dummy	Session Length Restriction
Democratic Control	–.57	2.26***
	(.49)	(.57)
Divided Control	.09	.14
	(.56)	(.60)
Log Federal Revenue	2.51*	–1.37
	(1.08)	(1.19)
Log Mineral Revenue	–.38*	.61**
	(.15)	(.19)
Log Population	1.51***	–2.42***
	(.41)	(.59)
Log Personal Income	2.44	5.59+
	(2.33)	(3.02)
Population Growth	1.21	–3.68
	(4.52)	(6.50)
Percentage Metropolitan Population	–.76	4.74*
	(1.54)	(2.08)
Gubernatorial Power	–.03	–.02
	(.08)	(.09)
Population Heterogeneity	8.09	–23.10**
	(6.33)	(7.70)
Opportunities to Advance	29.65	18.94
	(59.41)	(47.02)
Constant	–59.48*	–2.35
	(23.48)	(28.10)
Pseudo R^2	.34	.37
Log Likelihood	–87.69	–74.92

***$p < .001$; **$p < .01$; *$p < .05$; $^{+}p < .10$ (two-tailed).
Note: Standard errors in parentheses. N = 191 for all regressions. Year fixed-effects included but not reported.

incorrect to interpret results from a standard regression as if it were. After matching, however, there are very few significant differences between treatment and control, save for assignment to condition, a scenario approximating the experimental ideal.

With balance established in the matched sample, I can address the average treatment effect, or the average partial effect of professionalism on a state. As Table 5 shows, for all classes of spending, the average treatment effect is statistically insignificant. Professional legislatures spend no more money than nonprofessional legislatures, if we account for the fact that high-spending states are more likely to

TABLE 4
Evaluating Balance between Treated and Control for Covariate Means

	Kurtz-Squire Dummy						*Session Length Restriction*					
	Unmatched Sample			Matched Sample			Unmatched Sample			Matched Sample		
Variables	Prof.	Unprof.	Dif.	Prof.	Unprof.	Dif.	Prof.	Unprof.	Dif.	Prof.	Unprof.	Dif.
Democratic Control	.36	.46	–.11	.41	.38	.03	.52	.17	.34***	.48	.43	.05
Divided Control	.28	.25	.03	.24	.30	–.05	.24	.32	–.08	.25	.42	–.16+
Log Federal Revenue	6.02	5.84	.18**	5.94	5.92	.01	5.96	5.90	.07	5.92	5.93	.00
Log Mineral Revenue	5.13	5.93	–.79**	5.24	5.05	.18	5.76	4.93	.83**	5.58	5.43	.15
Log Population	15.16	14.43	.73***	14.98	15.00	–.02	14.58	15.34	–.76***	14.76	14.74	.02
Log Personal Income	9.41	9.17	.23***	9.35	9.37	–.02	9.28	9.34	–.05	9.26	9.30	–.02
Population Growth	.05	.06	–.01+	.05	.05	.01	.06	.04	.02*	.06	.06	.00
Percentage Metropolitan Pop.	.69	.51	.18***	.65	.66	–.01	.58	.66	–.09*	.57	.55	.02
Gubernatorial Power	15.47	13.78	1.69***	14.82	14.33	.49	14.39	15.32	–.93*	14.18	12.24	1.94***
Population Heterogeneity	.48	.45	.03***	.47	.47	.00	.46	.49	–.03***	.46	.45	.00
Opportunities to Advance	.007	.003	.004***	.004	.004	.000	.004	.008	–.004***	.004	.003	.001

***$p < .001$; **$p < .01$; *$p < .05$; $^{+}p < .10$ (two-tailed).
Note: *t*-tests used to evaluate significance of difference of means. N = 191 for all regressions.

TABLE 5
Average Treatment Effects of Professionalism on Government Spending

Dependent Variables	Kurtz-Squire Dummy	Session Length Restriction
Total Expenditures	80.07	–49.09
	(105.62)	(82.37)
Social Services Expenditures	34.69	–27.78
	(51.73)	(64.37)
Non–Social Services Expenditures	45.38	–21.32
	(35.68)	(40.23)
Treated (N)	74	106
Control (N)	81	42
Treated Off Common Support (N)	26	24
Control Off Common Support (N)	10	19

***$p < .001$; **$p < .01$; *$p < .05$; $^{+}p < .10$ (two-tailed).
Note: Bootstrapped standard errors in parentheses.

select professionalism as their optimal institutional structure. This finding is not simply a result of changes in the sample size made to achieve balance, but rather due to the increased comparability between professional and nonprofessional states.[16]

Moreover, the treatment effects are substantively minor. Professional states, as classified by the Kurtz-Squire dummy, spend $80.07 more per capita than nonprofessional states spend, which represents 10.5% of *one* standard deviation in total spending in these data. Further, the mean per capita spending of a nonprofessional state is $1,173.21; professionalism would have a minor influence even if these states adopted a professional legislative structure.[17] Similarly, a restriction on session length reduces spending by only $49.09 per capita, representing only 6.4% of a standard deviation. As for budgetary line items, professionalism has a greater (but still limited) effect on non–social services expenditures than on social services expenditures. The average treatment effect of professionalism as measured by the Kurtz-Squire dummy is $45.38 per capita, or 10.5% of one standard deviation in non–social services spending (compared to $34.69 and 8.7% for social services spending).

Discussion

The idea that professional legislators spend more than their citizen counterparts spend has pervaded the popular consciousness. Moreover, it is the conclusion of the only existing academic study on the subject and is implied by the extensive literatures on term limits and direct democracy at the state level. This presumption is consequential, as it has led several movements to turn professional legislatures into part-time bodies, changes that may substantially affect public policies. But these conjectures have been based on either speculation or inattention to the fact that professionalism, as a strategic choice, may be a *response to* and not a *cause of* growth in the public sector. My results provide empirical support for this hypothesis. Using propensity score matching, I found no significant effect of professionalism on several classes of expenditures, both targetable and nontargetable. Furthermore, the substantive size of the average treatment effects are also small, representing a minor percentage of both the mean and distribution of the dependent variable. There are theoretical reasons to believe that both professional and citizen structures would contribute to higher spending. Although confirmation is beyond the scope of this study, it is possible that these effects cancel one another out, resulting in little net effect.

A major focus of state politics research, and legislative studies more generally, has been on ascertaining the effect of institutional structures on public policies. This study highlights a classic case of potentially nonrandom assignment into the institutional "treatment" affecting inference. The observed correlation between professionalism and spending is very high, but we see no effect of professionalism when we compare states that are equally likely to receive the treatment. Indeed, combining these findings with those of my previous work (Malhotra 2006) suggests that states with large and rapidly growing public sectors construct professional legislatures so that their reelection-minded members may better handle the demands of their positions, which include providing constituency service and oversight for state bureaucracies and the executive branch. These results also suggest that previous studies analyzing the influence of professionalism should be revisited if assignment is potentially nonrandom.

Future studies can expand upon this topic by assessing the effects of movements to deprofessionalize legislatures via term limits (Kousser 2005). Although it is difficult to convert a full-time legislature into a part-time body by cutting salaries and session lengths, popular initiatives in several states have successfully prevented legislators from staying in office indefinitely. Term limits have been found to prevent members from building the expertise necessary to craft legislation. It is plausible that legislators who know they will be leaving soon may not invest the time and effort to become experts in logrolling and building coalitions, knowing that they must seek other opportunities soon. It would interesting to see if term limits have curtailed spending, as their proponents, who are supporters of small government, intended. Again, because term limits are not randomly assigned, future work must approximate an experimental design. More broadly, the methods used in this analysis can be exported to any investigation of the effect of institutional structures on policy outputs, particularly in cases where those structures were self-imposed.

A common thread throughout various subfields in political science is that institutions matter. A logical consequence of this fact is that political elites are aware of institutions' importance and make rational, strategic choices when choosing structural frameworks. This study highlights another instance where analyzing the effects of institutions on public policies requires we take into account how those institutions are selected in the first place.

Neil Malhotra <neilm@stanford.edu> *is Assistant Professor of Political Economy, Stanford Graduate School of Business, Stanford University, 518 Memorial Way, Stanford, CA 94305-5015.*

APPENDIX 1
Measurement of Variables and Sources of Data

Professionalism

The *Squire Index* is the average of legislator compensation, days in legislative session, and legislative resources (all expressed as proportions of corresponding traits of Congress). *Compensation* is the mean annual compensation in salary and living expenses for legislators during the session. *Days in Session* is the number of legislative days, averaged across the session. *Legislative Resources* is the mean annual expenditure for the legislative branch per member (excluding legislator compensation). State compensation and days in session figures came from various editions of *The Book of the States* (Council of State Governments, various years). State legislative resources figures are from various editions of *State Government Finances* (U.S. Census Bureau, various years). I took congressional compensation figures from *Congressional Quarterly's Guide to Congress* (Congressional Quarterly 1991). Congressional days in session and legislative resources figures came from Ornstein, Mann, and Malbin 2002. Data were obtained from King 2000, which provides a more-specific description of measurement techniques and sources (338–39). The *Kurtz* classification scheme is taken from Kurtz 1992, in which the 50 legislatures are divided into three categories:

professional, hybrid, and citizen. *Session Length Restriction* indicates any constitutional or statutory restriction on the number of days the legislature can meet and was obtained from *The Book of the States*.

Spending

Per Capita Total Expenditures is the total amount of money budgeted, divided by state population. *Per Capita Social Services Expenditures* is the amount of money budgeted for health, hospital, education, public welfare, and employment security services, divided by state population. *Per Capita Non–Social Services Expenditures* is the amount of money budgeted for highways, natural resources, corrections, police, and financial administration, divided by state population. I adjusted all spending figures (as well as all variables measured in dollars) for inflation using the Consumer Price Index (CPI) and expressed figures in 1994 dollars. All data are from *The Book of the States* for each cross section: 1964, 1974, 1984, and 1994.

Control Variables

Democratic Control is a dummy variable indicating that the governorship is held by a Democrat and both legislative chambers are controlled by Democrats. *Divided Control* is a dummy variable indicating that neither political party controls the governorship and both legislative chambers. Data on partisan control came from *The Book of the States*. *Population Level* is state population for 1964, 1974, 1984, and 1994. *Population Growth* is the percentage increase in population over the four years prior to year *t*. *Percentage Metropolitan Population* is the percentage of the state population residing in a metropolitan statistical area, as defined by the U.S. Census Bureau. *Per Capita Income* is the total income received by residents of a state, divided by state population. I obtained the population and income variables from the *Statistical Abstract of the United States* (U.S. Census Bureau, various years). *Per Capita Revenue from the Federal Government* is the amount received by each state from the federal government, divided by state population. *Per Capita Mineral Revenue* is the total value of production of petroleum, natural gas, coal, and nonfuel minerals, divided by state population. Data for both revenue variables came from *Government Finances* (U.S. Census Bureau, various years).

Additional Covariates

Population Heterogeneity is the Sullivan index of diversity. I gathered Index scores from 1960 and 1980 from Morgan and Wilson 1990. Scores from 1970 and 1990 were generously provided by James D. King. *Gubernatorial Power* is the Schlesinger index, which considers appointive, budget, and veto powers as well as tenure potential. Data for 1964 came from Schlesinger 1965; data for 1974 came from Schlesinger 1971; for 1984, from Beyle 1983; and for 1994, Beyle 1999. Following Maestas (2000), I defined *Opportunities to Advance* as the average number of House seats in a state that turned over in the three elections prior to the session, multiplied by the percentage of seats held by former state legislators, and then divided by the total number of state legislative seats. I gathered data on seat turnover and House member experience from the *Biographical Directory of the United States Congress*. Data on the number of state legislators came from *The Book of the States*.

APPENDIX 2
Technical Issues

Missing Data

Following previous researchers who have examined state spending (Gilligan and Matsusaka 1995, 2001; Owings and Borck 2000; Primo 2006), I decided to exclude Alaska from the analysis because its immense oil and gas deposits make it an extreme outlier; it spends nearly twice as much per capita as the second-highest-spending state. All four Nebraska sessions and the 1963–64 Minnesota session were also excluded because those legislatures were nonpartisan.

Nonspherical Errors

Because I used panel data, the disturbances may be nonspherical, which is a matter of analytical concern. In the regression analyses, I detected heteroskedasticity and autocorrelation with Breusch-Pagan (Breusch and Pagan 1979) and Wooldridge (Wooldridge 2002) tests, respectively. I used panel-corrected standard errors (Beck and Katz 1995) when I detected nonspherical disturbances. In cases of autocorrelation without heteroskedasticity, I assumed panels to be independent.

NOTES

I gratefully acknowledge Barry Weingast, Alberto Diaz-Cayeros, Jonathan Wand, Morris Fiorina, Simon Jackman, Stephen Jessee, John Bullock, Jowei Chen, Connor Raso, and participants of the Stanford University Workshop in Statistical Methods for valuable comments and suggestions. I also thank James D. King and John G. Matsusaka for generously providing data.

1. John Fund, "Less Is More," *Wall Street Journal*, 19 April 2004, editorial.

2. James Freeman, "Congress, Go Home and Stay Home," *USA Today*, 11 August 1999, editorial.

3. In one paragraph, Owings and Borck (2000) discuss the results of a two-stage least squares regression, but they do not explain or justify their instruments, nor do they test the instruments' validity statistically.

4. Berkman (2001) has countered with a finding that interest group density is greater in highly professionalized states as a result of greater legislator demand.

5. Dummy variables and variables measured in proportions, such as Democratic/Divided Control, Population Growth, and Percentage Metropolitan Population, are not logged.

6. Spending is not part of the propensity score equation because its inclusion would result in the selection of matches based on the dependent variable. One strength of matching is that the outcome variable is completely ignored and observations are selected solely according to the criterion of balance, as in a randomized experiment. As Ho et al. (2007) explain, "To ensure that selection during preprocessing depends only on X_i (to prevent inducing bias), the outcome variable Y_i should not be examined during the preprocessing stage. As long as Y_i is not consulted and is not part of the rule by which one drops observations, preprocessing cannot result in stacking the deck one

way or another" (216). I have found, however, that increases in state expenditure are an important predictor of professionalization (Malhotra 2006). Assessing both directions simultaneously in a matching analysis is not possible. Therefore, exploring the effect of spending on professionalism must be done separately, as I have done in previous work.

7. One complication of using the Squire index in across-time analyses is that the baseline of congressional professionalism changes. Squire (2007) has found, however, that the index correlates extremely highly ($r > .99$) with a revised version that uses a fixed baseline, and it is therefore appropriate for dynamic analyses. Further, Mooney (1994) has shown that professionalism scores without baselines are not comparable across time and that the Squire index is the best measure for historical analyses.

8. Although recent techniques have been developed to analyze continuous treatments (Imai and van Dyk 2004; Imbens 2000; Joffe and Rosenbaum 1999), these procedures require the stratification of observations and, consequently, large sample sizes. Because the dataset analyzed here has only 200 observations, it cannot be used to construct reasonably large strata.

9. "Hybrid" states are pooled with "professional" states because of the small sample size of states in the latter category. The treatment effect can therefore be interpreted as the spending reductions caused by the state having a citizen legislature. I reanalyzed the data, pooling hybrid and citizen states, and found similar results: professionalism does not cause more spending.

10. The equation I used to create the dichotomous treatment variable has high predictive power, suggesting that Kurtz's qualitative categorization is a good representation of the Squire index. The percentage of observations correctly predicted by equation (6) is 78%, the expected percentage correctly predicted (Herron 1999) is 72.3%, and the percentage reduction in error (Hagle and Mitchell 1992) is 38.9%.

11. To assess robustness, I also cross-tabulated the Squire index with the Kurtz categorization and selected a reasonable cut-point to identify a legislature as professional (Squire index value > .235). I conducted the analyses using this alternative definition and achieved results statistically and substantively similar to those reported in this article.

12. Technically speaking, a legislature that meets biennially has a limit even if the single session can meet for an indefinite amount of time. Following previous research (King 2000; Mooney 1995), I decided not to classify such legislatures as having a session length restriction because they have no constitutional or statutory limit on the session. Nevertheless, I reanalyzed the data, coding all biennial legislatures as having session length restrictions, and the results are statistically and substantively similar to those reported here.

13. Despite my attempts to replicate perfectly Owings and Borck's (2000) regression model, there are still inconsistencies between my results and theirs. Still, the principal relationship between professionalism and spending is similarly strong. The coefficient on professionalism in their main analysis has a t-statistic of 4.12; the coefficient in my analysis has a t-statistic of 4.14.

14. For a simple robustness check on the results of the constant-elasticity model, I regressed current spending against past spending and past professionalism (along with the controls) to determine if professionalization occurred before spending increases, following the logic of the Granger causality test (Granger 1969). Using this modified

approach, I replicated all nine specifications of Table 2, and the professionalism measure never emerged as a statistically significant predictor of spending. Results pointing to the effect of professionalism on spending do not emerge when one accounts for temporal order, a fact that should raise our doubts about the robustness of previous empirical work. The limitation of such an analysis is that if lagged professionalism causes lagged spending, or if professionalism only affects spending contemporaneously, then such a specification may attenuate the effect of lagged professionalism. Accordingly, I chose to focus on the propensity score analyses.

15. I also estimated other specifications of the propensity score equations to obtain matches. The findings are robust to the particular set of observables included in the regression model. The particular specification I chose is appropriate because it includes all available observables and results in balance between the treated and untreated samples. Further, it is important to note that the predictive power of the propensity score equations is irrelevant in assessments of the validity of the matching procedure. The propensity scores serve only as devices to balance the observed distribution of covariates across treated and comparison groups; "the success of the propensity score estimation is assessed by this resultant balance rather than by the fit of the models used to create the estimated propensity score" (D'Agostino and Rubin 2000, 750).

16. After implementing the matching algorithm, I excluded some observations because they did not lie in the area of common support. I reestimated the constant-elasticity models using these restricted samples and found the strong, positive association between professionalism and spending still present. Thus, the difference in the results between the regression and matching analyses is not due solely to the different samples and the removal of a few outliers.

17. A commonly used statistic to assess the power of the treatment on the control units is ATU, the average treatment effect on the untreated. For these data, the ATU is only $79.89 total spending per capita ($p = .58$), again showing that imposing professional institutions would have little effect on the budgets of currently nonprofessional legislatures.

REFERENCES

Arcenaux, Kevin, Alan S. Gerber, and Donald P. Green. 2006. "Comparing Experimental and Matching Methods Using a Large-scale Voter Mobilization Experiment." *Political Analysis* 14: 37–62.

Bailey, Delia. 2006. "Measuring the Effect of Voting Technology on Residual Votes." Presented at the annual meeting of the Midwest Political Science Association, Chicago.

Beck, Nathaniel, and Jonathan N. Katz. 1995. "What To Do (and Not To Do) with Time-series Cross-section Data." *American Political Science Review* 89: 634–47.

Behrman, Jere R., Yingmei Cheng, and Petra E. Todd. 2004. "Evaluating Preschool Programs when Length of Exposure to the Program Varies: A Nonparametric Approach." *Review of Economics and Statistics* 86: 108–32.

Berkman, Michael B. 1993. "Former State Legislators in the U.S. House of Representatives: Institutional and Policy Mastery." *Legislative Studies Quarterly* 18: 77–104.

Berkman, Michael B. 1994. “State Legislators in Congress: Strategic Politicians, Professional Legislatures, and the Party Nexus.” *American Journal of Political Science* 38: 1025–55.

Berkman, Michael B. 2001. “Legislative Professionalism and the Demand for Groups: The Institutional Context of Interest Population Density.” *Legislative Studies Quarterly* 26: 661–79.

Berkman, Michael, and James Eisenstein. 1999. “State Legislators as Congressional Candidates: The Effects of Prior Experience on Legislative Recruitment and Fundraising.” *Political Research Quarterly* 52: 481–98.

Berry, William D., Michael B. Berkman, and Stuart Schneiderman. 2000. “Legislative Professionalism and Incumbent Reelection: The Development of Institutional Boundaries.” *American Political Science Review* 94: 859–74.

Beyle, Thad. 1983. “Governors.” In *Politics in the American States*, 4th ed., ed. Virginia Gray, Herbert Jacob, and Kenneth N. Vines. Boston: Little, Brown.

Beyle, Thad. 1999. “Governors.” In *Politics in the American States*, 7th ed., ed. Virginia Gray, Russell L. Hanson, and Herbert Jacob. Washington, DC: Congressional Quarterly Press.

Biographical Directory of the United States Congress. http://bioguide.congress.gov/biosearch/biosearch.asp [June 1, 2008].

Bohn, Henning, and Robert P. Inman. 1996. “Balanced Budget Rules and Public Deficits: Evidence from the U.S. States.” *Carnegie-Rochester Conference Series on Public Policy* 45: 13–76.

Breusch, Trevor S., and Adrian R. Pagan. 1979. “A Simple Test for Heteroskedasticity and Random Coefficient Variation.” *Econometrica* 47: 1287–94.

Carey, John M., Richard G. Niemi, and Lynda W. Powell. 1998. “The Effects of Term Limits on State Legislatures.” *Legislative Studies Quarterly* 23: 271–300.

Carey, John M., Richard G. Niemi, and Lynda W. Powell. 2000. “Incumbency and the Probability of Reelection in State Legislative Elections.” *Journal of Politics* 62: 671–700.

Carey, John M., Richard G. Niemi, Lynda W. Powell, and Gary F. Moncreif. 2006. “The Effect of Term Limits on State Legislatures: A New Survey of the 50 States.” *Legislative Studies Quarterly* 31: 105–34.

Carmines, Edward G. 1974. “The Mediating Influence of State Legislatures on the Linkage Between Interparty Competition and Welfare Policies.” *American Political Science Review* 68: 1118–24.

Chen, Jowei, and Neil Malhotra. 2007. “The Law of *k*/*n*: The Effect of Chamber Size on Government Spending in Bicameral Legislatures.” *American Political Science Review* 101: 657–76.

Coates, Dennis. 1999. “Efficiency Wages for Politicians: Do Better Paid Elected Officials Produce Better Outcomes?” University of Maryland, Baltimore. Manuscript.

Congressional Quarterly. 1991. *Congressional Quarterly’s Guide to Congress*. 4th ed. Washington, DC: Congressional Quarterly.

Council of State Governments. Various years. *The Book of the States*. Lexington, KY: Council of State Governments.

D’Agostino Jr., Ralph B., and Donald B. Rubin. 2000. “Estimating and Using Propensity Scores with Partially Missing Data.” *Journal of the American Statistical Association* 95: 749–59.

Dehejia, Rajeev H., and Sadek Wahba. 2002. "Propensity Score Matching Methods for Nonexperimental Causal Studies." *Review of Economics and Statistics* 84: 151–61.

Dilger, Robert Jay, George A. Krause, and Randolph R. Moffett. 1995. "State Legislative Professionalism and Gubernatorial Effectiveness, 1978–1991." *Legislative Studies Quarterly* 20: 553–71.

Fiorina, Morris P. 1994. "Divided Government in the American States: A Byproduct of Legislative Professionalism?" *American Political Science Review* 88: 304–16.

Fiorina, Morris P. 1996. *Divided Government*. 2d ed. Boston: Allyn and Bacon.

Fiorina, Morris P., and Roger G. Noll. 1978. "Voters, Bureaucrats, and Legislators: A Rational Choice Perspective on the Growth of the Bureaucracy." *Journal of Public Economics* 9: 239–54.

Gilligan, Thomas W., and John G. Matsusaka. 1995. "Deviations from Constituent Interests: The Role of Legislative Structure and Political Parties in the States." *Economic Inquiry* 33: 383–401.

Gilligan, Thomas W., and John G. Matsusaka. 2001. "Fiscal Policy, Legislature Size, and Political Parties: Evidence from the First Half of the Twentieth Century." *National Tax Journal* 54: 57–82.

Granger, Clive W.J. 1969. "Investigating Causal Relations by Econometric Methods and Cross-Spectral Methods." *Econometrica* 34: 424–38.

Hagle, Timothy M., and Glen E. Mitchell II. 1992. "Goodness-of-Fit Measures for Probit and Logit." *American Journal of Political Science* 36: 762–84.

Herron, Michael C. 1999. "Postestimation Uncertainty in Limited Dependent Variable Models." *Political Analysis* 8: 83–98.

Herron, Michael C., and Kenneth W. Shotts. 2006. "Term Limits and Pork." *Legislative Studies Quarterly* 31: 383–403.

Ho, Daniel E., Kosuke Imai, Gary King, and Elizabeth A. Stuart. 2007. "Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference." *Political Analysis* 15: 199–236.

Holland, Paul W. 1986. "Statistics and Causal Inference." *Journal of the American Statistical Association* 81: 945–60.

Imai, Kosuke, and David A. van Dyk. 2004. "Causal Inference with General Treatment Regimes: Generalizing the Propensity Score." *Journal of the American Statistical Association* 99: 854–66.

Imbens, Guido W. 2000. "The Role of the Propensity Score in Estimating Dose-Response Functions." *Biometrika* 87: 706–10.

Joffe, Marshall M., and Paul R. Rosenbaum. 1999. "Propensity Scores." *American Journal of Epidemiology* 150: 327–33.

King, James D. 2000. "Changes in Professionalism in U.S. State Legislatures." *Legislative Studies Quarterly* 25: 327–43.

Kousser, Thad. 2005. *Term Limits and the Dismantling of State Legislative Professionalism*. Cambridge, UK: Cambridge University Press.

Kurtz, Karl T. 1992. "Understanding the Diversity of American State Legislatures." *Legislative Studies Section Newsletter* June: 2–5.

Maestas, Cherie. 2000. "Professional Legislatures and Ambitious Politicians: Policy Responsiveness of State Institutions." *Legislative Studies Quarterly* 25: 663–90.

Malhotra, Neil. 2006. "Government Growth and Professionalism in U.S. State Legislatures." *Legislative Studies Quarterly* 31: 563–84.
Matsusaka, John G. 1995. "Fiscal Effects of the Voter Initiative: Evidence from the Last 30 Years." *Journal of Political Economy* 103: 587–623.
Matsusaka, John G. 2004. *For the Many or the Few: The Initiative, Public Policy, and American Democracy*. Chicago, IL: University of Chicago Press.
Mooney, Christopher Z. 1994. "Measuring U.S. State Legislative Professionalism: An Evaluation of Five Indices." *State and Local Government Review* 26: 70–78.
Mooney, Christopher Z. 1995. "Citizens, Structures, and Sister States: Influences on State Legislative Professionalism." *Legislative Studies Quarterly* 20: 47–67.
Morgan, David R., and Laura Ann Wilson. 1990. "Diversity in the American States: Updating the Sullivan Index." *Publius* 20: 71–81.
Ornstein, Norman J., Thomas E. Mann, and Michael Malbin. 2002. *Vital Statistics on Congress: 2001–2002*. Washington, DC: American Enterprise Institute.
Owings, Stephanie, and Rainald Borck. 2000. "Legislative Professionalism and Government Spending: Do Citizen Legislators Really Spend Less?" *Public Finance Review* 28: 210–25.
Primo, David M. 2006. "Stop Us before We Spend Again: Institutional Constraints on Government Spending." *Economics & Politics* 18: 269–312.
Reed, Robert W., and Eric D. Schansberg. 1996. "Impact of Congressional Tenure Restriction on Spending." In *Legislative Term Limits: Public Choice Perspective,* ed. Bernard Grofman. Boston: Kluwer Academic.
Reed, W. Robert, D. Eric Schansberg, James Wilbanks, and Zhen Zhu. 1998. "The Relationship Between Congressional Spending and Tenure with an Application to Term Limits." *Public Choice* 94: 85–104.
Rosenbaum, Paul R., and Donald B. Rubin. 1983. "The Central Role of the Propensity Score in Observational Studies for Causal Effects." *Biometrika* 70: 41–55.
Rubinfeld, Daniel L. 1987. "The Economics of the Local Public Sector." *Handbook of Public Economics*, ed. A.J. Auerbach and M. Feldstein. Amsterdam: North-Holland.
Sabloff, Paula L.W. 1995. "The Decision Process of a Professionalized Legislature: A Cognitive Anthropology Approach." *Political Behavior* 17: 403–32.
Schlesinger, Joseph A. 1965. "The Politics of the Executive." In *Politics in the American States*, 1st ed., ed. Herbert Jacob and Kenneth N. Vines. Boston: Little, Brown.
Schlesinger, Joseph A. 1971. "The Politics of the Executive." In *Politics in the American States*, 2d ed., ed. Herbert Jacob and Kenneth N. Vines. Boston: Little, Brown.
Squire, Peverill. 1988. "Career Opportunities and Membership Stability in Legislatures." *Legislative Studies Quarterly* 13: 65–82.
Squire, Peverill. 1992. "Legislative Professionalism and Membership Diversity in State Legislatures." *Legislative Studies Quarterly* 17: 69–79.
Squire, Peverill. 1997. "Another Look at Legislative Professionalization and Divided Government in the States." *Legislative Studies Quarterly* 22: 417–32.
Squire, Peverill. 2007. "Measuring State Legislative Professionalism: The Squire Index Revisited." *State Politics and Policy Quarterly* 7: 211–27.
Squire, Peverill, and Keith E. Hamm. 2005. *101 Chambers: Congress, State Legislatures, and the Future of Legislative Studies*. Columbus, OH: Ohio State University Press.

Stonecash, Jeffrey M., and Anna M. Agathangelou. 1997. "Trends in the Partisan Composition of State Legislatures: A Response to Fiorina." *American Political Science Review* 91: 148–55.

U.S. Census Bureau. Various years. *Government Finances*. Washington, DC: Government Printing Office.

U.S. Census Bureau. Various years. *State Government Finances*. Washington, DC: Government Printing Office.

U.S. Census Bureau. Various years. *Statistical Abstract of the United States*. Washington, DC: Government Printing Office.

U.S. Congress. Senate. Committee on Governmental Affairs. 1996. *The Biennial Budget Act of 1996*. 105th Cong., 2d sess., 24 June. Washington, DC: Government Printing Office.

Weingast, Barry R., Kenneth A. Shepsle, and Christopher Johnsen. 1981. "The Political Economy of Benefits and Costs: A Neoclassical Approach to Distributive Politics." *Journal of Political Economy* 89: 642–64.

Wooldridge, Jeffrey M. 2002. *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT Press.