Political Appointments, Civil Service Systems, and Bureaucratic Competence: Organizational Balancing and Executive Branch Revenue Forecasts in the American States

George A. Krause   University of Pittsburgh
David E. Lewis   Princeton University
James W. Douglas   University of North Carolina at Charlotte

Scholarship on executive politics provides conflicting views about whether staffing administrative agencies through politicized or (politically) autonomous means is the best method for maximizing bureaucratic competence. We offer a theoretical account which maintains that obtaining a proper balance between both types of personnel systems across the supervisory and subordinate levels of an organization will best foster bureaucratic competence. We evaluate our organizational balancing thesis using data on executive branch general revenue fund forecasts in the American states from 1987 to 2002. States with a combination of politically appointed agency executives and merit-selected subordinates generally provide more accurate revenue forecasts than states that possess uniformly politicized personnel selection systems. Conversely, states with a combination of department head–appointed executives and subordinates chosen from an at-will system (i.e., nonmerit) produce more accurate forecasts than states with uniformly autonomous personnel selection systems. Our statistical findings underscore the positive consequences associated with balancing politicized and autonomous means of selecting personnel within hierarchies of political organizations.

Governance in democratic society is premised on the simple notion that the citizenry can effectively control their government. Responsiveness to the broader polity, however, requires that government exhibit competence in the tasks it is delegated to perform. While elected officials seek responsiveness from the bureaucracy, a government that is ineffective at executing policy cannot be responsive to the broader polity. This tension between political responsiveness and bureaucratic independence governs the selection of unelected officials to fill government positions. The selection methods used for staffing unelected posts provide direct insight into how elected officials weigh their desire of minimizing agency problems arising from delegated authority with their need to provide bureaucracies sufficient slack to effectively perform tasks.

While a separation-of-powers framework is useful for understanding the appointment of U.S. federal agency heads (McCarty 2004) and independent regulatory commissioners (Snyder and Weingast 2000), personnel selection systems are not always politicized. Some agency personnel arrangements do allow personnel to be hired and fired at will by political actors—i.e., an "at-will" system. Other agency personnel systems stipulate that personnel be selected by merit-based civil service procedures or some other politically autonomous means. Agencies

often use one arrangement for the selection of executives and another to fill subordinate positions. For example, in some agencies political actors select the executives but subordinate positions are filled by civil service procedures. In other agencies, executives are solely chosen by other unelected government officials but subordinate positions can be filled without reference to civil service rules. This begs the question: How do different means of selecting personnel systematically influence bureaucratic behavior? This is an important question to pose since the procedural mechanisms used to select agency officials can affect bureaucratic policy decision making. For instance, Krause (1994, 1996) has shown that different appointment mechanisms used to select Federal Open Market Committee members can explain variations in consensual decision making regarding U.S. monetary policy.

We maintain that striking a balance between at-will and (politically) autonomous personnel systems at different levels of an administrative organization is essential to maximizing bureaucratic competence. One important reason this is true is that using information provided by both at-will and autonomous personnel will produce more accurate bureaucratic decisions than when either agent type dominates both levels of an administrative organization. We test our thesis with new panel data on executive branch general fund revenue forecasts in the American states from 1987 to 2002. Our empirical evidence shows that this type of organizational balancing in the selection of agency personnel across executive and subordinate levels yields more accurate general fund revenue forecasts in the American states.

Neutral Competence versus Responsive Competence in the Administrative State

The modern debate on the virtues of a politicized versus an independent bureaucracy within political science can be traced to Heclo (1975) and Moe (1985). It was Heclo (1975), and Kaufman (1956) two decades earlier, who most convincingly argued that American presidents should seek neutral competence by emphasizing the hiring of career professionals over political appointees when making executive staffing decisions. Neutral competence entails the application of bureaucratic expertise in an objective manner to obtain the best outcomes possible (Kaufman 1956; see Rourke 1992, 539). Agency personnel selected outside the political appointment process are more apt to possess specialized policy expertise, meaning-

ful experience, public management skills, and relationships with key stakeholders (Heclo 1975). They also enjoy long-standing relationships with other institutional actors and also serve as "honest brokers" in a world of partisan and ideological divisions.

According to the neutral competence perspective, politicized selection of agency personnel is thought to be inversely related to bureaucratic competence by those inside government (e.g., 2003 National Commission on the Public Service—i.e., the 2nd Volcker Commission Report). Political appointees are more apt to have their professional orientation colored by democratic institutions responsible for their appointment (Moe 1982; Nathan 1983; Wood and Waterman 1994) and politicized personnel systems exhibit greater turnover than civil service or merit systems, thus producing both vague and volatile signals regarding an agency's policy objectives (Ban and Ingraham 1990; Heclo 1977; Mackenzie 1987). Recent empirical research supports this view by demonstrating that an increasing reliance on political appointees at the expense of career civil servants in several countries has hurt performance by reducing morale, shortening tenures among career civil servants, and erecting barriers to effectively recruiting highly qualified individuals into government service (Suleiman 2003). Relatedly, Gilmour and Lewis (2006) show that U.S. federal programs run by career managers get systematically higher management grades than programs administered by political appointees.

Moe (1985), however, provides a more skeptical (and contrasting) view of the professional bureaucracy by claiming that their general lack of both understanding and loyalty necessitates presidents to make the bureaucracy more responsive by "manipulating civil service rules, proposing minor reorganizations, and pressing for modifying legislation . . . to increase the number and location of administrative positions that can be occupied by appointees." (Moe 1985, 245).[1] Political appointment thus "counteracts inertia, ensures an influx of new ideas, and keeps government in touch with a variety of interested groups and constituencies" (Bok 2003, 265). Such individuals are also necessary for creating the requisite risk taking and entrepreneurial climate found in high-performing organizations (Bilmes and Neal 2003).

Politicized appointment systems are thus thought by adherents to a responsive competence perspective to contain certain benefits with respect to improving bureaucratic competence vis-à-vis autonomous appointment

[1]Moe (1985) implies that political appointees can provide enough capacity to execute agency tasks. If valid, then differences between autonomous and politicized personnel systems for performance should either be negligible or favor the politicized system.

systems. Recent empirical evidence has claimed that the performance gap between political appointees and careerists is either decreasing or nonexistent as the result of better education and more prior government experience for political appointees (Aberbach and Rockman 2000, 164; Donahue 2003; Michaels 1997), as well as staying longer in politically appointed positions than commonly perceived (Maranto 1998). Nonetheless, a more educated, experienced, and committed public sector workforce comprised of political appointees does not necessarily guarantee an enhanced level of agency performance. Any performance gains can be offset by a lack of professional objectivity and also a willingness to abdicate their professional judgment in response to both external political pressures and competing information sources that vie for elected officials' attention (Rourke 1992).

Next, we provide a theoretical account that attempts to reconcile these competing explanations of executive politics by demonstrating how the design of agency personnel systems shapes the quality of bureaucratic performance. Our story is predicated on the hierarchical relationship between *supervisory* and *subordinate* levels of a bureaucratic organization. We posit that having bureau personnel at the supervisory and subordinate levels chosen by a different personnel selection process improves administrative performance.

Organizational Balancing Between Politicization and Autonomy within Agencies

While there are many factors that influence bureaucratic performance, including task complexity, budgets, and agency culture, an important and often overlooked determinant of an agency's outputs is its personnel selection system. This is an important dimension of bureaucratic performance because it presumes that the type of individuals who comprise unelected government positions do affect the substantive content and quality of public policy. If the appointing (contractual) principal is an elected official, bureaucratic agents face greater pressures to be politically responsive (Wilson 1989, 197–200). Elected officials possessing formal appointment powers are more capable of shaping agency behavior than those who do not directly select agency executives (Krause 1994, 1996). Organizational pathologies arising from politicized personnel selection mechanisms range from lower human capital, higher personnel turnover, decreasing morale, and greater difficulty with recruitment and retention of high-capacity bureaucrats. When the appointing principal(s) is not an

elected official (e.g., agency head; civil service systems; professional or licensing boards; citizen or public interest groups), bureaucrats will be more politically insulated since the acquisition of or removal from their positions is beyond the purview of elected officials.[2]

Effective coordination within hierarchies is difficult because of conflicting interests across levels of an organization (Miller 1992, 196). These vertical dilemmas can hinder bureaus' ability to respond to both political demands and public pressures. Executive authority residing with supervisors rests directly upon the acceptance or consent of subordinates within the agency since bureaucratic compliance cannot be achieved by fiat (Barnard 1938, 164; Brehm and Gates 1997; Simon 1976). Yet, it remains possible that varying preferences within an organizational hierarchy will have positive consequences if used to improve the amount and quality of policy information at a bureau's disposal for decision-making purposes. Different personnel selection methods can thus mitigate any ingrained biases or inherent weaknesses separately held by political appointees and civil servants resulting from their respective orientations centered on political responsiveness and professional norms. When agency supervisors and subordinates are both selected independent of electoral institutions, however, they are not only better insulated from political pressures, but also tend to be isolated from other sources of useful policy information and expertise located elsewhere within the executive and legislative branches. In other words, both types of agency personnel enjoy a comparative advantage in terms of the skills and orientation that they each possess. Therefore, we assert that balancing politicized and autonomous personnel selection between supervisory and subordinate levels of an organization will maximize the level of bureaucratic competence.[3]

The microlevel underpinnings of our organizational balancing theory are rooted in policy information counterbalancing among supervisory and subordinate levels within an organization. This, in turn, can provide an effective means to mitigate information distortion attributable to hierarchies within organizations and hence result in more accurate decision making (Cyert and March 1963,

[2]This is not to suggest that autonomous personnel systems cannot be influenced by politics. Instead, we claim that political responsiveness differs across personnel selection mechanisms and thus has a variable impact on bureaucratic outputs and outcomes.

[3]Horn (1995) contends that a mixture of politicized and autonomous personnel selection systems will result in greater transaction costs than a unitary personnel selection system. We maintain, however, that transaction costs will be more than offset by the higher caliber of agency performance produced by a mixed personnel system. This rival hypothesis is empirically testable within the context of our statistical analysis.

67–82; Downs 1967, 121–22).[4] Organizational balancing can thus foster increased bureaucratic competence via the "mutual policy learning" that often occurs when politicized and autonomously chosen agency personnel question each others' underlying perspectives or assumptions (e.g., Ascher 1978; Klay 1985). Politicized selection of bureaucrats can improve bureaucratic performance by advocating greater flexibility and innovation in public agencies, being more attuned to the polity's will through elected representatives, and bringing policy information derived from outside sources to bear on administration. On the flip side, a politicized agency has a tendency to commit myopic decision-making errors when it is preoccupied with ideological, partisan, or electoral demands at the expense of technical concerns and professional norms. Bureaucrats selected by an autonomous personnel system are afforded organizational stability and memory, as well as greater discretion to utilize their technical expertise. This selection method, however, also engenders its own pathologies arising from professional norms, inertia arising from an agency's culture or history, and a myopic commitment to the agency's policy mission. Complete bureaucratic independence can lead to a narrow common agency perspective divorced from both the larger democratic concerns of the polity, as well as better alternatives for executing administrative and policy tasks.

Ideally, if personnel selection systems can be designed to offset these distinct set of shortcomings between supervisory and subordinate levels of an organization, better overall performance will ensue. This is because organizational balancing improves the quality of bureaucratic competence by exploiting both generalist and specialist skills used to conduct administrative tasks (Aberbach and Rockman 2000, 56). More specifically, we maintain that a mixed personnel selection system relying on a combination of politicized supervisory (subordinate) appointments and autonomously chosen subordinates (supervisor) will yield the highest level of bureaucratic competence since competing sources of policy information and expertise are embedded within the organization.

This leads us to offer a set of straightforward theoretical hypotheses concerning bureaucratic performance under different personnel arrangements at the executive and subordinate levels within agencies. Agencies use different arrangements for the selection of executives (at-will, autonomous) and subordinates (at-will, autonomous), and we can use this variation to evaluate the impact of organizational balancing on the quality of agency performance. For simplicity, in the succeeding discussion we will often interchange the term "politicized" for at-will selection of executives and "merit" for (politically) autonomous processes of selecting subordinates since these are the most common arrangements in federal and state agencies in the United States.

> *H1*: Given an at-will subordinate selection system, agencies whose executives are selected via an autonomous appointment system will outperform bureaus whose executives are selected by a political appointment system.
>
> *H2*: Given a merit subordinate selection system, agencies whose executives are selected via a political appointment system will outperform those bureaus whose executives are selected by an autonomous appointment system.
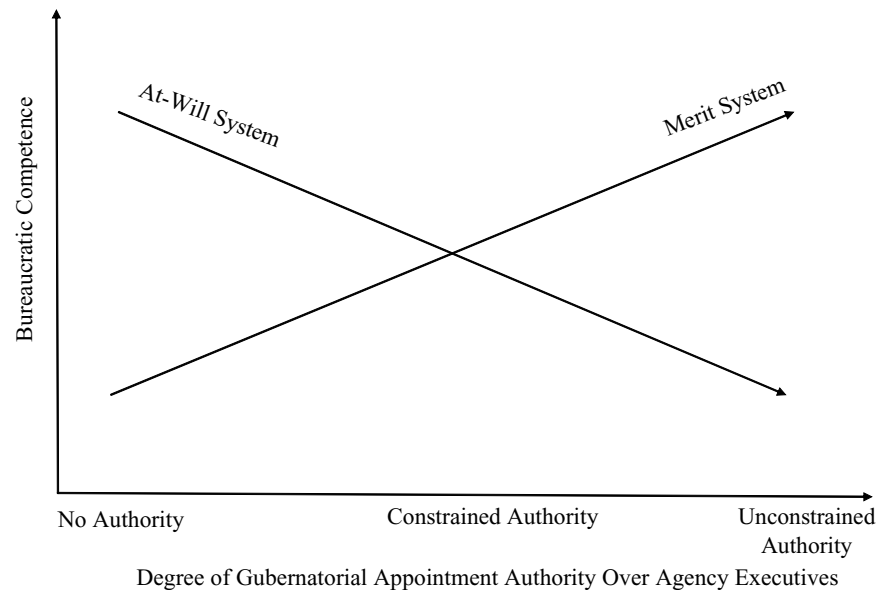>
> *H3*: Given a political executive appointment system, agencies whose subordinates were chosen through a merit system will outperform bureaus using an at-will selection system.
>
> *H4*: Given an autonomous executive appointment system, agencies whose subordinates were chosen through an at-will system will outperform bureaus using a merit selection system.

H1 simply means that under an at-will (i.e., nonmerit) subordinate personnel system, we expect public bureaus whose agency executives are chosen by a (politically) autonomous process to be more competent than when they are chosen through a politicized process.[5] H2 indicates the exact opposite pattern. Under a merit system, we predict

---

[4]Policy information counterbiasing presumes that (1) supervisors must have knowledge of the type of distortion emanating from subordinates—i.e., the former must know whether the latter obtained their position from a politicized or autonomous personnel selection process; and (2) it must also be in the interest of supervisors to offset or reduce such distortion problems (e.g., agency production of either observable outputs or outcomes). Within the context of our organizational balancing theory, this suggests that supervisors will encounter *different types* of distortion from subordinates based upon whether the latter were selected via politicized or autonomous means. Underlying our theory is the notion that policy information counterbiasing is easier to undertake when the supervisor was chosen by a different personnel selection mechanism than their subordinates, ceteris paribus.

[5]We are neither implying that "at-will" subordinates are necessarily inferior to those chosen by a merit system, nor are they identical to politically appointed supervisors insofar that they perform the same roles and tasks within a public organization. Rather, we are claiming that at-will chosen subordinates will be more sensitive to political pressures than merit-system chosen subordinates; and that the former subordinate type faces very similar incentives as politically appointed supervisors since both serve at the pleasure of electoral institutions that possess formal authority regarding both their selection to and removal from office.

FIGURE 1   Theoretical Relationship Between Agency Personnel
Selection System and Bureaucratic Competence

that bureau competence is higher when agency executives are chosen by a politicized process compared to an autonomous process. H3 predicts that given a political executive appointment system, bureaus staffed through the merit system at the subordinate level will outperform bureaus relying on an at-will system. H4 predicts that when agency executives are chosen independent of elected officials, bureaus staffed through the merit system at the subordinate level will be less competent than those relying on an at-will system. Put simply, we hypothesize that personnel selection system heterogeneity (homogeneity) across the supervisory and subordinate levels of an administrative organization yields increases (reductions) in bureaucratic competence (see Figure 1).

Case Selection, Data, Variables, and Methods

One difficulty in the study of both policy formulation and bureaucratic performance is the ability to quantitatively measure the quality of decisions that are comparable both across time and jurisdictions. We test our theoretical hypotheses with new panel data on executive branch general fund revenue forecasts in the 50 American states from 1987 to 2002. Examining revenue forecast accuracy provides an innovative way to evaluate directly bureaucratic competence since it reveals the extent to which

agencies' expectations regarding fiscal conditions actually mirror reality. Policy decisions exhibiting greater (lesser) accuracy are indicative of superior (inferior) bureaucratic competence.

A focus on states' general fund revenues is commonplace in both the revenue forecasting (Cassidy, Kamlet, and Nagin 1989; Rodgers and Joyce 1996) and fiscal political economy literatures (Alt and Lowry 2000; Poterba 1994) for a variety of reasons. First, on a policy level, revenue forecasting is one of the most vital functions of state-level bureaucracies since these estimates play a large role in determining resource allocation for state government programs. Second, estimates of general fund revenues comprise the largest component of total state government revenues. Third, it is hard to analyze forecasts at a different (lower or higher) level of aggregation. General sales, personal income, and corporate income taxes constitute 76% of all state general fund revenues (NASBO 2004, 94), but data is not available from all the states or years for each of these sources of revenue. We cannot analyze estimates of total state revenues since earmarked funds are sometimes not forecasted by executive branch budget offices, but rather the line agencies obtaining such funds (e.g., Franklin and Douglas 2003).[6] Thus, we cannot

---

[6]Furthermore, the political incentives surrounding forecasts of general revenue funds differ from the incentives associated with forecasting earmarked revenue funds since the former (latter) entails a discretionary (nondiscretionary) source of revenue (e.g., Buchanan 1963; Patashnik 1997).

directly assess comparable institutions performing comparable tasks if we analyze disaggregated or more fully aggregated revenue estimates.

Our main dependent variable assessing forecast accuracy is the absolute percentage forecast error (APFE) in the executive branch estimate of state general fund revenues in each state for each year from 1987 to 2002 [(|actual state general fund revenues–projected state general fund revenues|)/actual state general fund revenues] * 100.[7] In these models a positive value implies more error and less accurate forecasts. The minimum mean value of this dependent variable by state is 2.78% (Wisconsin) and the maximum mean value for a given state is 24.34% (Alaska). The overall mean absolute forecast error is 7.11% (SD = 7.65%).[8] While it is true that these executive branch estimates of general fund revenues will not always bind the legislature since the latter can adopt forecasts generated outside the executive branch, the presence of competing forecasts that can be used for such purposes are accounted for in our subsequent multivariate statistical analysis.[9]

Different Personnel Selection Systems within Bureaucratic Hierarchies

States have different means of selecting budget agency personnel at both the executive and subordinate levels. Some states require that budget agency subordinates come from a merit-based civil service system. Others have no such requirement. States also vary in gubernatorial control over selection of the budget agency director. In some states, governors select the budget director without any requirement for consultation. In other states, the budget director is selected by another bureaucrat without the requirement of gubernatorial consultation. In between these two extremes are states that require that the budget director pass legislative confirmation.[10] Our theory

predicts that states possessing a mix of at-will and autonomous selection procedures at the executive and subordinate levels will produce forecasts that yield the highest level of accuracy.

To account for how budget agency directors (supervisors) and subordinate personnel are selected,[11] we include a dummy variable—"G"—for budget agencies where the governor appoints the director without any requirement for approval by the legislature or another outside party (0, 1; 53%;[12] unconstrained gubernatorial authority/politicization). We also create an indicator—"D"—for whether a department head selects the director without any requirement for gubernatorial or legislative approval (0, 1; 6%; autonomous). The base category includes state years where the governor selects the budget director with the legislature's approval (20%), a gubernatorial appointee selects the director with the governor's approval (18%), and South Carolina where a politically appointed panel produces the governor's forecast (base category—constrained gubernatorial authority/politicization).[13] We hypothesize that greater gubernatorial appointment influence should reduce executive branch general fund revenue forecast accuracy. We also include an indicator—"Merit"—for whether budget agency subordinates are selected through civil service procedures (1, 0; 73%) or whether employees are selected at the director's discretion outside the civil service. For the 1987–2002 period, between 34–40 states selected budget agency personnel through a civil service system. We surmise that states with merit system selection of budget agency subordinates will produce more accurate forecasts, independent of how the budget director is chosen.

Our set of organizational balancing hypotheses predict that states with a mix of at-will and autonomous personnel selection at the executive and subordinate levels

---

[7]National Governors Association (NGA) and National Association of State Budget Officers (NASBO). *The Fiscal Survey of the States*, various years. Washington, DC.

[8]Since this variable is skewed towards zero, in auxiliary analysis we also estimated all the models based on a Box-Cox transformation of this variable. The correlation between the Box-Cox transformed absolute percentage forecast error and its untransformed counterpart is 0.84. Auxiliary analyses showed that our results were robust to this transformation of the dependent variable.

[9]We organized our data set so that all of the data are aligned by the appropriate fiscal year. For example, if unified government began in a state in calendar year 2000, the state would be recorded as beginning unified government in FY2001 since the revenue forecast generated early in calendar year 2000 would be for the FY2001 budget.

[10]South Carolina has a separate politically appointed budget board.

[11]Each of these variables was constructed from data given in the *Budget Processes in the States*, 1987–2002 (Washington: The National Association of State Budget Officers). Unfortunately, this data source is not published annually. We contacted officials in states with missing or inconsistent data in the documents to ascertain the accuracy of the data as well as the exact year in which appointment processes changed. A list of the officials contacted in each state along with their phone numbers can be obtained from the authors.

[12]The reported percentages associated with this and subsequent binary dummy variables pertain to the percentage of cases that are coded as being equal to 1.

[13]There do not appear to be any notable regional biases in the selection of agency budget directors. For instance, "G" is the modal category in three out of four geographical regions (South, Pacific & Rocky Mountain, and Midwest) and is an approximately bimodal category for the lone remaining region (New England & Mid-Atlantic). "D" occurs with some nonzero frequency in all but the Midwest region.

will produce more accurate general fund revenue forecasts than those states that rely only on either type of system. This is because those agency personnel selected by political means will be predisposed towards making optimistic revenue projections in order to create favorable economic news, and provide an impetus for tax cuts (Republican governors) or increased government spending (Democratic governors). Conversely, agency personnel chosen through a (politically) autonomous selection process will be more inclined to err on the side of conservative forecasts (Bretschneider and Gorr 1987). Professional norms push forecasters to be more conservative since overestimating revenues can lead to adverse political and fiscal consequences. Conversely, political pressures often counterbalance these distortions by inducing forecasts to be more optimistic since there is rarely a political constituency for unspent cash on hand. Therefore, organizational balancing can improve the accuracy of bureau forecasting decisions by offsetting internal biases within administrative agencies.

At first glance, the average absolute percentage forecast error for states with organizational balancing is substantially lower than in states without balancing. In states with politically appointed executives and merit-system subordinates the average forecast error is 6.31%, significantly less than the average forecast error of states with other personnel selection systems (7.50%, p < 0.05). In states where agency executives are chosen by bureaucratic procedures and subordinate positions can be filled by at-will employees without regard to civil service rules the average forecast error becomes even lower (4.42%) relative to other combinations of personnel selection (7.17%, p < 0.05).

Statistical Controls

Of course, other factors may influence the quality of state revenue forecasts, including aspects of the political environment such as unified government and election cycles. For instance, increased agreement between the executive and the legislature make it more likely that political pressure will influence the budget forecast (Bretschneider et al. 1989). To account for this we include an indicator that is coded 1 if the two branches of government share the same party and 0 otherwise (40%).[14] Election cycles might also lead to worse forecasts as political pressures increase

to alter forecasts. We include an indicator for gubernatorial election year (0, 1; 26%) and expect the coefficient to be positive (more error).[15] We also include controls for the party of the governor. States with Republican governors are coded with a 1 (48%), states with independents a 0 (2%), and Democrats −1 (49%). In terms of forecast accuracy, past research has found that Democratic governors tend to produce less accurate revenue forecasts than Republican counterparts (Bretschneider and Gorr 1987). We include party control measures in each model to account for any partisan differences that might exist.

States vary in both economic conditions and structure of the budget process. To account for this we include a series of controls. We include an indicator for whether the state has a binding balanced budget requirement (0, 1; 52%)—that is, whether a balanced budget must actually be passed by a state's legislature. We choose this form because the final legislative appropriations are the true targets that governors are trying to influence. The presence of a strict requirement creates an incentive for governors to provide more conservative forecasts in order to produce a number that the legislature can realistically stay beneath (Cassidy, Kamlet, and Nagin 1989). This, of course, should lead to lower levels of forecast accuracy because, in balanced budget states, conservative projections are more important than accurate ones.[16] We also include a control for fiscal slack that involves the combined size of the state's rainy day and surplus general funds as a percentage of actual general fund revenues ($\bar{X} = 5.35$, SD $= 11.83$). States with large levels of fiscal slack can afford to produce less accurate revenue estimates.[17]

The relative percentage of general fund revenues from sales tax sources is included to control for the relative volatility of a state's income stream. Because the sales tax is a relatively stable revenue source, states with a greater

---

[14]These data come from *The Book of the States*, 1986–2003. Lexington: The Council of State Governments. In auxiliary analyses, we have also estimated models interacting this unified partisan branch dummy variable with a measure of political appointment of agency executive selection expecting the coefficient to be positive, implying that unified government matters more in states with politicized executive selection. These alternative empirical tests accounted for both unconstrained gubernatorial appointment authority and gubernatorial appointment with legislative confirmation, respectively. We could not reject the null that the coefficient on these interactions was zero and thus fail to detect statistical evidence supporting this pattern of conditional political influence.

[15]Since one might expect politicized budget agencies to be most sensitive to political pressures stemming from the electoral cycle, we have also estimated models including an interaction of the measures of relative politicization with gubernatorial election year. We could not reject the null that the coefficient on this interaction term was zero.

[16]We constructed this variable using the *Budget Processes in the States*, 1987–2002, in conjunction with phone interviews of executive budget agency officials in all 50 states. A list of these individuals and their phone numbers can be obtained from the authors.

[17]This variable is constructed from data in *The Fiscal Survey of States*, 1986–2003.

dependence upon the sales tax should have an easier time predicting their total general fund revenues. Therefore, they should produce more accurate general fund revenue forecasts. To measure relative sales tax dependency, we utilize the proportion of general fund revenues that come from the sales tax for each state in a given year ($\bar{X} = 0.24$, SD $= 0.11$).[18] Similarly, we include the percentage change in the state's real per capita income as a way of measuring economic activity that might influence forecast performance ($\bar{X} = 3.12$, SD $= 3.14$). Large changes in personal income should make revenues more difficult to predict, resulting in less accurate forecasts.[19] We include an indicator for whether states produce more than one revenue forecast (0, 1; 78%) since competing forecasters, such as legislative budget offices and consensus forecasts, serve as a check against executive power, and thus should result in greater forecast accuracy (Bretschneider and Gorr 1987; Bretschneider et al. 1989). We account for whether a state operates under a biennial budgeting cycle (0, 1; 40%). Our expectation is that states with biennial budget cycles will have less accurate forecasts since they confront more uncertainty than under an annual budget cycle scenario. We also control for whether a state issued an official consensus group revenue forecast in a given year (0, 1; 47%). These groups are comprised of partisan and/or nonpartisan members, and the legislature is required to work from these particular revenue forecasts in preparation of fiscal spending plans.[20] The sign of this coefficient is ambiguous. On one hand, states with consensus forecasts will be able to mitigate the particular biases of executive branch general fund revenue forecasts (negative sign). However, these consensus forecasts might provide information that, if utilized, can adversely affect the accuracy of these forecasts in a conservative manner (positive sign).[21] This is because the consensus general fund revenue forecasts might provide an external institutional check on the level of relative optimism reflected in executive branch forecasts.

Methods

In modeling executive branch forecasting performance in the American states, we face a panel design structure in which our model estimation choices must be guided by both substantive theory and data constraints. Because the number of cross-sectional units exceeds time units by over a factor of 3 (N $= 50$, T $= 16$), we do not adopt statistical methods commonly employed by political scientists that are primarily designed to handle research designs where T $\geq$ N (e.g., Beck and Katz 1995, 644; Stimson 1985, 928–29). Further, accounting for conventional cross-sectional fixed effects (CSFEs) is not a sound practice here due to problems arising from (1) collinearity (see Baltagi 1999, 309); (2) weakly time-invariant nature of our bureaucratic personnel selection system dummies;[22] and (3) the standard rank-condition assumption pertaining to the CSFEs will not be met (Assumption FE.2; Wooldridge 2003, 269).[23]

Therefore, we adopt two distinct strategies for estimating our statistical models to ensure the robustness of our results. The first approach involves modeling cross-sectional heterogeneity as substance, rather than nuisance, by not employing a cross-sectional effects estimation strategy. We model the cross-sectional heterogeneity in our data through our theoretical variables of interest and instead treat timewise heterogeneity as either unobserved (random effects) or observed (fixed effects). This is because each year will bring common circumstances that executive agency personnel in the American states must confront when formulating general fund revenue forecasts across different years and budget cycles since they respond to top-down federal level macroeconomic policy and outcomes via intergovernmental grants, monetary policy, and the like. Further, the use of time dummies to account for timewise fixed effects (or random effects) is appropriate in shorter panels since proper stochastic modeling of the dependent variable is difficult when T is small (Arellano 2003, 60–64). Therefore, it is preferable to allow for time-varying intercepts when one

---

[18]These data come from *The Book of the States*, 1986–2003, and the U.S. Department of Commerce, Bureau of the Census (http://www.census.gov/govs/www/state.html).

[19]Data for this variable come from the U.S. Department of Commerce, Bureau of Economic Analysis (http://www.bea.gov/regional/statelocal.htm).

[20]We classified Texas as a consensus group state. Although the Texas Comptroller produces this official revenue forecast, it acts as a consensus forecast for this state given that it binds the legislature in the resulting fiscal policy process similarly to other states possessing consensus groups.

[21]The Spearman rank rho correlation between the competing and consensus forecast dummy variables is of moderate magnitude (0.41).

[22]By *weakly* time invariant, we mean that a rather modest change in a given variable occurs within a given cross-sectional unit (state). By *strongly* time-invariant, we mean that such variables are temporally fixed for a given cross-sectional unit.

[23]Hsiao echoes this sentiment by contending that "if the explanatory variables contain some time-invariant variables, z_i, their coefficients cannot be estimated by CV (covariance estimation), because the covariance transformation eliminates z_i from" the covariance transformed equation (2003, 35). Thus, modeling the cross-sectional heterogeneity as either a random or deterministic process risks "throwing out the baby with the bathwater" by treating important substantive cross-sectional differences as nuisance.

has a cross-sectional dominant panel (large N relative to T; Wooldridge 2003, 170).[24]

Our second strategy is to employ a *fixed effects variance decomposition* (FEVD) estimation approach that exploits the dominance of *between* (spatial) variance for those exogenous variables that follow either a strongly or weakly time-invariant process (Plümper and Troeger 2004; see also Hsiao 2003, 52).[25] This estimator is a three-stage technique where the conventional CSFE unit effects are estimated in the first stage; the next stage involves decomposing these unit effects by partitioning them into separate (strongly or weakly) time-invariant and residual components; and the third stage reestimates the original model via pooled OLS containing both the weakly time-invariant variables and the second-stage residuals noted above. The temporal dynamics for the FEVD estimator are accounted for by a Prais-Winsten AR(1) serial correlation correction. Because this technique accounts for cross-sectional heterogeneity while simultaneously "purging" common variance between nuisance and substantive effects, this estimator's finite sample properties—based on Monte Carlo evidence—has the desirable dual properties of (1) increasing the efficiency of CSFE estimates, and (2) reducing the likelihood of omitted variable bias associated with ignoring cross-sectional heterogeneity that is orthogonal to both strongly or weakly time-invariant exogenous variables.[26]

Statistical Findings

The statistical results for the model specifications appear in Tables 1–4 and proceed as follows.[27] Table 1 includes

estimates from baseline models that do not account for organizational balancing. Table 2 includes models that account for organizational balancing by interacting variables accounting for personnel selection mechanisms at the executive and subordinate levels. Table 3 includes Wald tests associated with the models in Table 2.[28] Finally, in Table 4 we include models of forecast *bias* rather than accuracy to illuminate why some personnel selection arrangements may be more accurate than others.

In Table 1 Models 1–4 assess the additive effect of agency executive and subordinate personnel selection systems on states' absolute percentage forecast error (APFE). Models 1 and 2 distinguish between unconstrained politicized ("G"), constrained politicized (base category: "constant"), and autonomous ("D") executive appointment systems, while Models 3 and 4 do not make a distinction between the two types of politicized executive appointment systems. Since our primary empirical results are robust across both timewise heterogeneity and FEVD estimation strategies, our subsequent interpretations are limited to the former estimates of these forecast accuracy models (Models 1 and 3).[29]

---

[24]We remain agnostic on this particular issue since differences between timewise fixed and random effects potentially become larger when N > T, thus requiring the application of the Hausman test to determine which modeling choice is more appropriate (Hsiao 2003, 51).

[25]We thank Thomas Plümper and Vera Troeger for making their STATA program code available to us.

[26]Plümper and Troeger's (2004) Monte Carlo evidence determines that the appropriate *between* (spatial) variance to *within* (temporal) variance ratio is approximately equal or greater than the 2.5. In our empirical analysis, the following exogenous variables can be defined as being at least weakly time invariant: "D" (4.36), "G" (3.81), Merit (3.19), "G" × Merit (2.85), "D" × Merit (3.62), Fiscal Slack (4.49), Relative Reliance on Sales Tax Revenues (6.56), Biennial Budget Cycle (3.35), Competing Forecasts (2.44), and Consensus Forecasts (2.31).

[27]The relatively low R^2 statistics associated with these statistical models are largely due to the fact that the dependent variables are expressed as (absolute) percentage forecast errors and thus are analogous to undergoing a first differencing transformation. Also, such forecast error variables are not prone to the stochastic trends typically found when separately analyzing many aggregate macroeconomic variables.

[28]In auxiliary analysis, we also estimated models that separately excluded Alaska and Texas because each contained outlying general revenue forecast accuracy observations to ensure that our statistical results were not the potential product of one state's overly influential extreme observations. The estimates from these alternative models are largely consistent with those produced by our reported results. A few relatively minor exceptions occur that are directly relevant for testing of our theory of organizational balancing. First, the deletion of Texas from the data set increases the statistical significance associated with Test 2D–Model 7: "D" + "D" × Merit = 0 from p = 0.11 to 0.06. Conversely, deletion of Texas cases results in a decline in statistical significance for Test 1A—Model 5: D = G = 0 from 0.05 to 0.19. The deletion of Alaska from the data set decreases the statistical significance associated with Test 2D—Model 7: "D" + "D" × Merit = 0 from p = 0.11 to 0.17, and also for Test 3A—Model 6: "G" + "G" × Merit = 0 drops from p = 0.00 to 0.052. However, dropping Alaska observations enhances the precision (and corresponding statistical significance) when estimating revenue forecast bias in Model 9 with respect to the individual coefficients for "G" and "G" × Merit, and "D" and "G" × Merit in Model 10, respectively.

[29]One difficulty in this analysis is that the different state institutions could be endogenous to a state's inherent forecasting difficulty. That is, states where forecasting is difficult may have one set of institutions and states where forecasting is easy may have another. If this is the case, it is hard to disentangle the distinct influence of the institutions themselves on forecast accuracy. Since our data do not contain the time period when most states chose their budget institutions, the endogeneity is difficult to model directly. Instead, we focus on those states that changed their institutions during the time period of our study. A *necessary condition* for the existence of a potential endogeneity problem is that states that either change their executive or staff personnel selection method should have faced a relatively more difficult forecasting period preceding this institutional change relative to the other 49 states covering this same time period. Out of the nine possible cases where this system change occurred in our data, we found no such pattern when

Our statistical results reveal that the institutional mechanism used for selecting the agency budget director has no discernible effect on state revenue forecast accuracy. This conclusion is also confirmed by the failure to reject the null that "D" is equal to "G" (Model 1: $\chi^2 \sim (1) = 0.00$, p = 0.99) for the Wald test restriction results appearing near the bottom of Table 1. States that select agency subordinates through merit systems typically produce more accurate revenue forecasts than states operating under an at-will staff system. The relationship between budget agency merit systems and executive branch general fund revenue forecast accuracy possesses the correct hypothesized sign and is statistically significant at p < 0.001. Specifically, merit-system regimes are estimated to have about 4% lower absolute percentage forecast error than at-will system regimes in Models 1–4. This substantive finding is consistent across all models in Table 1 but also in subsequent analyses conducted in this study.

The results also shed light on how important control variables influence state revenue forecast accuracy. For instance, neither unified government nor election year pressures have any meaningful bearing on forecast accuracy. We obtain some evidence in Models 2 and 4 that Republican governors produce less accurate forecasts than their Democratic colleagues contrary to the findings of Bretschneider and Gorr (1987). States with formidable balanced budget restrictions produce significantly less accurate general fund revenue forecasts than those states that either have weak or no such limitations. This makes sense given that the states facing tough balanced budget restrictions may have greater incentives to alter forecasts for political expediency. States that place a comparatively greater reliance on the stable source of sales tax revenues are not only more apt to produce executive branch forecasts that are more accurate, but also more willing to commit forecasting errors overstating the actual amount of general fund revenues coming into state government coffers (see Table 4 below). Higher state economic growth translates into greater executive branch general fund revenue forecast accuracy. Finally, states whose executive branch agencies experience bureaucratic competition

from other state-level government sources do not produce more accurate general fund revenue forecasts compared to those states where the executive branch budget agency has a monopoly over these tasks. As expected, these forecasts are less accurate in states with biennial budgets. States with a consensus forecast produce significantly less accurate general fund revenue forecasts than other states largely because their projections are too conservative (see Table 4 below).

The results from the additive models in Table 1 lead us to conclude that only the subordinate personnel selection mechanism matters for bureaucratic performance. States whose budget agencies are staffed by merit-based civil servants at the subordinate levels typically produce more accurate forecasts. One difficulty with these models is that they fail to consider how the hierarchical interplay between executive and subordinate levels of an organization affects the quality of bureau output.

We address this issue by estimating models (Table 2) that include interactions between executive and subordinate personnel selection variables consistent with our theoretical hypotheses. The corresponding Wald coefficient restriction tests that enable us to test theoretical relationships involving the equality or sum of coefficients appear in Table 3. Notable differences in forecast accuracy emerge among the various agency executive and subordinate selection system configurations when we account for organizational balancing. We look first at cases where the state has subordinates selected outside the merit system (H1). Models 5–8 indicate that having department heads select the budget director without the requirement of gubernatorial or legislative consultation significantly improves forecasts by 3.06–6.11% compared to the base category. States where governors are allowed to select the budget director without consultation with the legislature produce forecasts that are indistinguishable from states where they must get confirmation from the legislature (base category). The tests for differences between at-will and autonomous executive appointment systems do, however, uncover significant differences consistent with H1 based upon the individual coefficient (t) tests appearing in Table 2. Such differences are supported by the Wald coefficient restriction tests at the 0.05 level (Table 3: Test 1A).[30] In states possessing at-will subordinate personnel selection systems, having the department head

analyzing lagged volatility in actual general fund revenues from the three previous years (t − 2, t − 3, and t − 4), and only three instances of this pattern (Arizona: 1990, Colorado: 1988, and Idaho: 1997) with respect to revenue shocks representing the unanticipated component of general fund revenues (e.g., Crain 2003, 74–75). In these latter sets of exceptional cases, it must be stated that both the executive and staff selection methods changed simultaneously in one state (Arizona) for the only time in our data set, and another state only yielded a single observation from one year to base our inferences upon (Colorado). The paucity of such evidence from only a small subset of our broader data set is evidence in favor of the reliability of our estimates.

[30]We also find that states where subordinate positions are filled through the merit system typically produce lower average forecast errors than "at-will" states with a constrained politicized executive appointment system (base category). In other words, in states where the governor nominates and the legislature confirms the budget director, at-will states are estimated to have about 4.17% more absolute forecast error compared to merit states.

TABLE 1 Additive Impact of Agency Personnel Selection Systems on States' Executive General Fund Revenue Forecast Accuracy

Variable	(1) Absolute Forecast Error (Accuracy)		(2) FEVD – AR1 Abs. Forecast Error (Accuracy)		(3) Absolute Forecast Error (Accuracy)		(4) FEVD – AR1 Abs. Forecast Error (Accuracy)	
Executive and Subordinate Personnel Selection Systems								
Governor (G)	0.77	(0.59)	1.05*	(0.63)	—		—	
Department Head (D)	0.76	(1.18)	−0.94	(1.26)	0.29	(1.12)	−1.60	(1.19)
Merit (0, 1)	−3.94**	(0.65)	−4.21**	(0.69)	−4.10**	(0.63)	−4.52**	(0.67)
Statistical Controls								
Unified government (0, 1)	−0.53	(0.55)	−0.61	(0.56)	−0.69	(0.55)	−0.60	(0.56)
Gubernatorial election year (0, 1)	−0.53	(0.75)	−0.81	(0.52)	−0.44	(0.67)	−0.81	(0.52)
Party of governor (−1, 0, 1)	0.32	(0.28)	0.60**	(0.29)	0.41	(0.28)	0.60**	(0.29)
Balanced budget requirement (0, 1)	1.80**	(0.55)	1.97**	(0.59)	1.67**	(0.54)	1.77**	(0.58)
Fiscal slack (rainy day & surplus general funds)	0.01	(0.02)	−0.12**	(0.03)	0.01	(0.02)	−0.12**	(0.02)
Relative reliance on sales tax revenues	−18.14**	(2.60)	−17.95**	(2.78)	−17.21**	(2.58)	−17.57**	(2.77)
Change in real personal income	0.07	(0.09)	0.20**	(0.10)	0.005	(0.09)	−0.20**	(0.10)
Competing forecasts (0, 1)	0.46	(0.71)	0.26	(0.76)	0.39	(0.71)	0.26	(0.76)
Biennial budget (0, 1)	0.95*	(0.55)	1.07*	(0.59)	0.82	(0.55)	0.96	(0.59)
Consensus forecast (0, 1)	2.11**	(0.61)	2.40**	(0.65)	1.84**	(0.60)	2.29**	(0.64)
η	—		0.90**	(0.09)	—		0.90**	(0.09)
Constant	11.26**	(1.32)	12.80**	(1.37)	12.11**	(1.23)	13.72**	(1.25)
H_0: G = D (X^2, 1df)	0.00	[0.99]	2.66	[0.10]	—		—	
ρ	—		0.18		—		0.18	
F (13, 14, 13 df); Wald X^2 (12 df)	7.74**	[0.00]	13.95**	[0.00]	95.41**	[0.00]	15.05**	[0.00]
R^2	0.10		0.21		0.10		0.21	

Note: Dependent variable is [(|actual state general fund revenues – projected state executive general fund revenues|)/actual state general fund revenues] ∗ 100. N = 790, 735, 790, 735. Standard errors are inside parentheses. Probability levels inside brackets. Estimates for Model 1 are based upon timewise fixed effects regression and Model 3 estimates are based on timewise random effects regression. We could not reject the null in Model 3 that the coefficients in the timewise random and fixed effects models were the same (p < 0.70). *Significant at the 0.10 level; **Significant at the 0.05 level in two-tailed tests.

select the budget director produces more accurate executive branch general fund revenue forecasts compared to when the governor possess such appointment authority. Our statistical evidence unequivocally lends strong credence to H1.

If we change our focus to states where subordinates are selected through the merit system, we predict that having the governor appoint the director would be better for performance than having a department head do it (H2). While the hypothesized pattern is evident in both the numerical values and statistical significance associated with these individual coefficient estimates, these differences become more ambiguous in inferential terms when we assess the linear combination of these coefficients. For example, Model 5 estimates suggest that in merit states different configurations of executive appointments pro-

duce the following average absolute forecast errors: unconstrained gubernatorial authority: 5.91%, constrained gubernatorial authority: 5.34%, autonomous means via a department head: 7.65%.[31] The Wald tests in Table 3, however, show that we cannot easily reject the null of no difference among the different means of selecting the budget director for the timewise heterogeneity models at

[31]We calculate these numbers directly from the estimates. We set control variables at their means or reasonable values, take the product of the coefficients and values, and then sum up these quantities (−2.23%). We then add this value to the values we get for different configurations of the key variables. For example, for the results for unconstrained politicized executive appointment system we calculate from Model 5 the following: constant + Merit + G + Merit∗G or 11.75 − 4.17 + 0.87 − 0.30 = 8.14%. The simulated average percentage forecast error is 5.91% (8.14% − 2.23%). All similar analyses are calculated in this manner.

TABLE 2 Multiplicative Impact of Agency Personnel Selection Systems on States' Executive General Fund Revenue Forecast Accuracy

Variable	(5) Absolute Forecast Error (Accuracy)		(6) FEVD – AR1 Abs. Forecast Error (Accuracy)		(7) Absolute Forecast Error (Accuracy)		(8) FEVD – AR1 Abs. Forecast Error (Accuracy)	
Executive and Subordinate Personnel Selection Systems								
Governor (G)	0.87	(1.49)	1.39	(1.56)	—		—	
Department Head (D)	−3.06	(2.37)	−4.85*	(2.52)	−3.91**	(1.99)	−6.11**	(2.14)
Merit System (0, 1)	−4.17**	(1.44)	−4.23**	(1.50)	−4.56**	(0.65)	−4.98**	(0.70)
G × Merit	−0.30	(1.61)	−0.61	(1.69)	—		—	
D × Merit	5.37**	(2.73)	5.60**	(2.91)	5.99**	(2.37)	6.49**	(2.55)
Statistical Controls								
Unified government (0, 1)	−0.48	(0.55)	−0.64	(0.56)	−0.60	(0.54)	−0.59	(0.56)
Gubernatorial election year (0, 1)	−0.53	(0.75)	−0.81	(0.52)	−0.45	(0.68)	−0.81	(0.52)
Party of governor (−1, 0, 1)	0.35	(0.29)	0.61**	(0.29)	0.42	(0.28)	0.60**	(0.29)
Balanced budget requirement (0, 1)	1.72**	(0.55)	1.87**	(0.59)	1.63**	(0.54)	1.71**	(0.58)
Fiscal slack (rainy day & surplus general funds)	0.01	(0.02)	−0.12**	(0.03)	0.01	(0.02)	−0.12**	(0.03)
Relative reliance on state tax revenues	−18.18**	(2.60)	−17.93**	(2.78)	−17.46**	(2.57)	−17.64**	(2.77)
Change in real personal income	−0.06	(0.09)	−0.20*	(0.10)	−0.01	(0.09)	−0.20*	(0.10)
Competing forecasts (0, 1)	0.18	(0.72)	−0.06	(0.77)	0.11	(0.72)	−0.09	(0.77)
Biennial budget (0, 1)	1.08*	(0.56)	1.21**	(0.59)	0.98*	(0.55)	1.12*	(0.59)
Consensus forecast (0, 1)	2.01**	(0.62)	2.29**	(0.66)	1.77**	(0.60)	2.16**	(0.64)
η	—		0.89**	(0.09)	—		0.89**	(0.09)
Constant	11.75**	(1.78)	13.13**	(1.86)	12.69**	(1.27)	14.38**	(1.28)
ρ	—		0.17		—		0.18	
F Test (15, 16, 14 df); Wald X^2 (13 df)	7.11**	[0.00]	12.26**	[0.00]	103.55**	[0.00]	13.97**	[0.00]
R^2	0.11		0.21		0.11		0.21	

Note: DV [(|actual state general fund revenues – projected state executive general fund revenues|)/actual state general fund revenues] ∗ 100. N = 790. Standard errors are inside parentheses. Probability levels inside brackets. Model 5 estimates are based on timewise fixed effects regression. Model 7 estimates are based on timewise random effects regression. We could not reject the null that the coefficients in the timewise random and fixed effects models were the same for Model 7 (p < 0.91). *Significant at the 0.10 level; **Significant at the 0.05 level in two-tailed tests.

p = 0.20 and 0.09, respectively (Table 3: Tests 2A & 2D). Even though the numerical pattern for the FEVD – AR1 estimates follows our theoretical predictions consistent with H2, once again we cannot reject the null hypothesis of no difference among the various means of selecting the budget director either at or near conventional levels of significance. Our statistical results thus provide decidedly mixed empirical support for H2 with respect to state executive branch general fund revenue forecast accuracy.

The third hypothesis (H3) was that in states where the governor appoints the budget director, having subordinates selected through the merit system would lead to better performance. This is what the estimates indicate. In states where the governor selects the budget director, merit subordinate selection reduces the absolute forecast error by about 4.47% compared to at-will states (Model 5). These at-will–merit subordinate selection system differences are statistically discernible from one another at p ≤ 0.00 (Table 3: Test 3A). When the governor appoints the budget director, the statistical estimates suggest that staffing the budget agency through the merit system improves competence. Our empirical evidence provides strong unambiguous support for H3.

Finally, our fourth theoretical hypothesis (H4) asserts that states where the governor has no direct influence over budget director selection will perform better if the state has an at-will subordinate selection system. In such states, having a merit-based subordinate selection process makes the budget agency too insulated from politics, decreasing responsiveness, and thus leading to less realistic forecasts. The individual coefficient estimates suggest that such a difference exists in both numerical and inferential terms.

TABLE 3   Wald Coefficient Restriction Hypothesis Tests for Multiplicative Models of States' Executive General Fund Revenue Forecast Accuracy

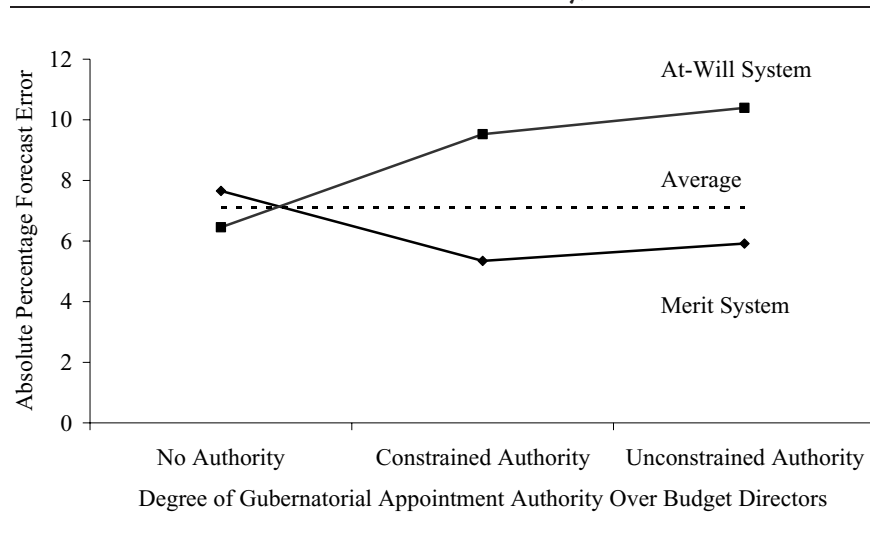| Null Hypothesis	Accuracy (5)	Accuracy FEVD – AR1 (6)	Accuracy (7)	Accuracy FEVD – AR1 (8)
Hypothesis 1: (At-will states) Do autonomous executive appointment systems outperform (unconstrained) politicized appointment systems?				
	Statistic [P-Value]			
Test 1A H_0: D = G	3.82 [0.05]	8.40 [0.00]	—	—
Hypothesis 2: (Merit states) Do (unconstrained) politicized executive appointment systems outperform autonomous executive appointment systems?				
	Statistic [P-Value]			
Test 2A H_0: merit $\times$ D + D = merit $\times$ G + G	1.66 [0.20]	0.00 [0.98]	—	—
Test 2B H_0: merit + merit $\times$ D + D + constant = 0	39.12 [0.00]	32.86 [0.00]	40.09 [0.00]	33.80 [0.00]
Test 2C H_0: merit + merit $\times$ G + G + constant = 0	59.51 [0.00]	76.01 [0.00]	—	—
Test 2D H_0: merit $\times$ D + D = 0	2.88 [0.09]	0.26 [0.61]	2.50 [0.11]	0.07 [0.79]
Hypothesis 3: ([Unconstrained] Politicized executive appointments) Do merit staffing states outperform at-will staffing states?				
	Statistic [P-Value]			
Test 3A H_0: merit $\times$ G + merit = 0	34.34 [0.00]	35.60 [0.00]	—	—
Test 3B H_0: G + merit $\times$ G + merit + constant = 0	59.51 [0.00]	76.01 [0.00]	—	—
Hypothesis 4: (Autonomous executive appointments) Do at-will staffing states outperform merit staffing states?				
	Statistic [P-Value]			
Test 4A H_0: merit + merit $\times$ D = 0	0.28 [0.60]	0.31 [0.58]	0.39 [0.53]	0.38 [0.54]
Test 4B H_0: merit + merit $\times$ D + D + constant = 0	39.12 [0.00]	32.86 [0.00]	40.09 [0.00]	33.80 [0.00]

For example, the estimates from Model 7 suggest that in states where the department head selects the budget director, at-will states produce about 1.20% less absolute forecast error than states where the budget office is staffed through the merit system. The other models produce similar estimates. Statistical evidence in support of H4, however, is more modest than compared to H1 and H3 since the overall magnitude of these forecast accuracy differences is not statistically discernible based upon the Wald coefficient restriction test results (Table 3: Test 4A).

In Figure 2 we show the estimates from Model 5 graphically to refocus attention on the larger theoretical claim about organizational balancing. The graph shows the relationship between different personnel selection systems and average percentage forecast error for executive branch general fund revenues, our empirical measures of bureaucratic competence. Merit system states tend to outperform at-will states since the former states have a lower absolute percentage forecast error than the latter states. The exception to this rule is when the department head appoints the director in merit system states. In such cases, absolute forecast errors increase. In at-will states, these forecast errors decrease when department heads select the budget director. Figure 2 resembles the theoretical expectations appearing in Figure 1 except that the differences between unconstrained versus constrained gubernatorial appointment authority turned out to be less consequential than we expected. This finding is hardly surprising given that elected officials representing different political parties and branches of government share a common incentive for generating optimistically biased general fund revenue forecasts. This common incentive is because Democratic (Republican) politicians tend to prefer to use sanguine general fund revenue estimates to advocate spending increases (tax cuts). In general, however, balancing between politicized and autonomous personnel selection at different levels of an organizational hierarchy improves bureaucratic competence for executive agencies.

Illuminating Forecast Inaccuracy by Analyzing Forecast Biases

In the previous section, we provide empirical evidence that organizational balancing across levels of budgeting agencies results in better performance. Yet, we do not have empirical leverage as to *how* this occurs. We surmise that politically and autonomously appointed agency personnel will share different forecasting preferences. As noted earlier, agency personnel selected by a politicized process will place a greater weight on making sanguine forecast errors, while those selected by an autonomous process will tend to produce more conservative forecasts.

FIGURE 2 **Statistical Relationship Between Alternative Personnel Selection Systems and Executive Branch State General Fund Revenue Forecast Accuracy, 1987–2002**

To address this question, we examine the role organizational balancing plays in affecting the type of *bias* in state level executive branch general fund revenue forecasts. Because bureaucratic competence refers to accurate decision making, this statistical analysis does not provide a test of our theory. What the analysis of forecast bias can shed light on is *how* these variations in forecast accuracy relate to different configurations of administrative personnel systems. We operationalize executive branch general fund revenue forecast bias as [(actual state general fund revenues–projected state general fund revenues)/actual state general fund revenues] * 100. Positive values imply that states are underestimating revenues, while negative values denote the opposite. The overall mean forecast error is 3.82 (SD = 9.73) and state averages range from −1.4 to 10.4. Iowa has the least biased forecasts in our sample ($\bar{X} = 0.06$).

The statistical results for forecast bias multiplicative models appear in Table 4. A summary of the control variable findings indicates that state revenue forecast bias is orthogonal to partisan politics (Party of Governor and Unified Government) and electoral cycles (Gubernatorial election year). States with strong balanced budget restrictions produce more conservative general revenue forecasts than those states that either have weak or no such limitations. Higher levels of fiscal capacity, measured by fiscal slack and relative reliance on sales tax revenues, allow state executive agencies to make more sanguine projection errors.

The models add texture to our claims regarding how organizational balancing affects bureaucratic competence via general fund revenue forecast biases, particularly for merit system states. In merit-system states, executive budget offices produce conservative projections when a department head selects the director. Not surprisingly, giving the governor more control over the selection of the budget director makes the forecasts less conservative (and, presumably more accurate). The conservative bias in forecasts is reduced by 2.44% and 3.20% when the governor appoints the budget director with and without the legislature's approval, respectively. When we compare states with completely autonomous personnel selection to states with completely at-will selection, at-will states produce more optimistic forecasts. Completely autonomous personnel selection states produce 0.68% more of an underestimate than completely at-will states.

The estimates from Table 4 are more erratic for at-will states. This may be due to the fact that states that produce more accurate general fund revenue forecasts will also often have less bias, making the interpretation of bias models difficult. At-will states produce both the most conservative and the most optimistic forecasts depending upon the degree of gubernatorial control in director selection. Having a department head select the director in at-will states produces more accurate general fund forecasts but does so by making them more optimistic.

Discussion

The dilemma of the modern administrative state centers on how to preserve a government's democratic character while also ensuring sufficient capacity to effectively perform delegated tasks on behalf of the polity and their

TABLE 4 Multiplicative Impact of Agency Personnel Selection Systems on States' Executive General Fund Revenue Forecast Bias

Variable	(9) Forecast Error (Bias)		(10) Forecast Error (Bias) FEVD − AR1		(11) Forecast Error (Bias)		(12) Forecast Error (Bias) FEVD − AR1	
Executive and Subordinate Personnel Selection Systems								
Governor (G)	−2.47	(1.83)	−1.19	(2.21)	—		—	
Department Head (D)	−7.58**	(2.92)	−5.72	(3.58)	−5.51**	(2.46)	−4.82	(3.06)
Merit System (0, 1)	−4.89**	(1.77)	−5.46**	(2.13)	−2.43**	(0.81)	−3.71**	(0.99)
G × Merit	3.23	(1.98)	2.62	(2.39)	—		—	
D × Merit	10.78**	(3.36)	8.33**	(4.14)	8.23**	(2.92)	6.65*	(3.64)
Statistical Controls								
Unified government (0, 1)	0.58	(0.68)	−0.25	(0.77)	0.41	(0.67)	−0.17	(0.77)
Gubernatorial election year (0, 1)	−1.13	(0.92)	0.05	(0.65)	−1.14	(0.92)	0.05	(0.65)
Party of governor (−1, 0, 1)	0.15	(0.35)	0.42	(0.40)	0.20	(0.35)	0.42	(0.40)
Balanced budget requirement (0, 1)	1.34**	(0.68)	2.03**	(0.84)	1.23*	(0.67)	1.77**	(0.83)
Fiscal slack (rainy day & surplus general funds)	−0.08**	(0.03)	−0.26**	(0.04)	−0.08**	(0.03)	−0.26**	(0.04)
Relative reliance on state tax revenues	−16.12**	(3.20)	−25.00**	(3.97)	−15.91**	(3.18)	−24.54**	(3.95)
Change in real personal income	0.51**	(0.11)	0.83**	(0.13)	0.51**	(0.11)	0.83**	(0.13)
Competing forecasts (0, 1)	−0.02	(0.88)	−0.07	(1.10)	−0.02	(0.88)	−0.07	(1.10)
Biennial budgeting (0, 1)	0.71	(0.69)	0.40	(0.84)	0.79	(0.68)	0.41	(0.84)
Consensus forecast (0, 1)	2.12**	(0.76)	3.05	(0.93)	2.28**	(0.75)	3.12**	(0.91)
η	—		0.84**	(0.11)	—		0.85**	(0.11)
Constant	8.42**	(2.20)	9.58**	(2.64)	6.31**	(1.50)	8.49**	(1.81)
ρ	—		0.26		—		0.26	
F Test (15, 16, 13, 14 df)	4.29**	[0.00]	9.27**	[0.00]	4.73**	[0.00]	10.59**	[0.00]
R^2	0.09		0.17		0.09		0.17	

Note: Dependent Variable is [(actual state general fund revenues − projected state executive general fund revenues)/actual state general fund revenues] ∗ 100. N = 790, 735, 790, 735. Standard errors are inside parentheses. Probability levels inside brackets. Model estimates for Model 9 and Model 11 are based on timewise fixed effects regression. *Significant at the 0.10 level; **Significant at the 0.05 level in two-tailed tests.

elected representatives. Evidence from state budget forecasting suggests that there are advantages to striking a proper balance between political responsiveness and bureaucratic autonomy (Aberbach and Rockman 2000, 172; Wilson 1989, 202). Our statistical analysis has shown that states with mixed personnel selection systems across executive and subordinate levels produced more accurate executive branch general fund revenue forecasts. Specifically, states with budget agencies staffed by merit-based civil service systems generally outperform states without such staffing practices. These results are contingent, however, on the manner in which the budget director is selected. If states with merit systems have budget directors selected outside the purview of elected officials, revenue forecasts suffer. If these states have budget directors selected by the governor, however, revenue forecasts improve. Similarly,

while states with budget agencies staffed outside the merit system produce worse revenue forecasts on average, they can be improved significantly if the budget director is appointed by a department head rather than a popularly elected governor.

In budget forecasting agencies it is neither desirable for electoral institutions to uniformly control personnel decisions (Moe 1985), nor is it preferable to entirely divorce personnel selection from politics altogether (Heclo 1975; Kaufman 1956). While we think the findings are of general importance, their applicability in other contexts outside of budget forecasting may depend on context-specific factors such as task complexity, unique characteristics pertaining to different forms of organizational hierarchy, or internal agency operations. Since the theoretical account we give is derived from features common to

many bureaucratic organizations, however, our expectations are that these empirical findings are generalizable to a larger class of public organizations within democratic systems.

Recent theoretical work has focused on how elected officials wish to determine the optimal balance involving the tradeoff between political control and bureaucratic capacity (Huber and McCarty 2004; Lewis 2005). We have extended this line of inquiry by analyzing whether the existence of political control or bureaucratic autonomy, through alternative personnel selection mechanisms, has direct consequences for bureaucratic competence. The theoretical account we advance predicts that bureaucratic competence is maximized when agency executives are selected by a different process than agency subordinates. Thus, organizational balancing melds organizational continuity with political responsiveness (Aberbach, Putnam, and Rockman 1981; Heclo 1977, 136) by mixing breadth and depth of expertise (Aberbach and Rockman 2000, 56). Organizational balancing reduces policy information distortions that naturally arise from hierarchies within organizations (Cyert and March 1963, 67–82; Downs 1967, 121–22).

This study has important substantive implications for the study of executive and bureaucratic politics in democratic settings. First, we demonstrate that too much political control, exercised through the design of personnel selection systems, can adversely affect bureaucratic competence. Too little political control can likewise diminish bureaucratic performance. Political insularity can result in conservatism and risk avoidance. Indeed, our statistical evidence shows that the most insular budget agencies do not perform as well "mixed" agencies. Organizational balancing provides politicians with a formal mechanism to engage in "institutional hedging," whereby both political responsiveness and bureaucratic autonomy are not incompatible for maximizing bureaucratic competence.

Second, we uncover that the means of selecting agency executives is of modest importance in explaining variations in bureaucratic competence independent from the personnel selection system used to choose subordinates. In other words, agency leadership only appears to be determinative of bureau output quality conditional on the type of subordinates who are working underneath them. Such a bottom-up perspective highlights the primary role subordinates play in determining the overall performance of a bureaucratic agency (Brehm and Gates 1997). This is a topical issue given the rise of government contracting and the increasing number of agencies with personnel systems outside the traditional civil service system (e.g., Department of Homeland Security, Federal Aviation Administration, Postal Service). Future research analyz-

ing the consequences of appointment politics must focus greater attention on the interaction between supervisory and subordinate levels to obtain a richer understanding of how organizational hierarchies within public bureaucratic agencies influence administrative performance.

On a more basic theoretical level, this study provides key insights into the seminal literature on hierarchies within political organizations. Scholars have convincingly demonstrated that institutional arrangements involving organizational design are inherently nonneutral since they entail political choices (Hammond and Thomas 1989; McNollGast 1989; Miller 1992; Moe 1989). Furthermore, the lack of neutral hierarchies within such organizations means that government officials must confront tradeoffs among various biases emanating from politicized and autonomous personnel selection systems (cf. Hammond and Thomas 1989, 158). Organizational balancing is an important means of mitigating the biases that naturally occur within political organizations, especially for bureaucratic decision making within a democracy where the normative objective is to maximize the quality of administrative performance.

References

Aberbach, Joel D., Robert D. Putnam, and Bert A. Rockman. 1981. *Bureaucrats and Politicians in Western Democracies.* Cambridge: Harvard University Press.

Aberbach, Joel D., and Bert A. Rockman. 2000. *In the Web of Politics: Three Decades of the U.S. Federal Executive.* Washington: Brookings Institution.

Alt, James E., and Robert C. Lowry. 2000. "A Dynamic Model of State Budget Outcomes Under Divided Partisan Government." *Journal of Politics* 62(November):1035–1069.

Arellano, Manuel. 2003. *Panel Data Econometrics.* New York: Oxford University Press.

Ascher, William. 1978. *Forecasting: An Appraisal for Policy Makers and Planners.* Baltimore: Johns Hopkins Press.

Baltagi, Badi H. 1999. *Econometrics.* New York: Springer-Verlag.

Ban, Carolyn, and Patricia W. Ingraham. 1990. "Short-Timers: Political Appointee Mobility and Its Impact on Political-Career Relations in the Reagan Administration." *Administration and Society* 22(March):106–24.

Barnard, Chester I. 1938. *The Functions of the Executive.* Cambridge: Harvard University Press.

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(September):634–47.

Bilmes, Linda J., and Jeffrey R. Neal. 2003. "The People Factor: Human Resources Reform in Government." In *For the People: Can We Fix Public Service?* ed. John D. Donahue and Joseph S. Nye Jr. Washington: Brookings Institution, pp. 113–133.

Bok, Derek. 2003. "Government Personnel Policy in Comparative Perspective." In *For the People: Can We Fix Public Service?*

ed. John D. Donahue and Joseph S. Nye Jr. Washington: Brookings Institution, pp. 255–272.

Brehm, John, and Scott Gates. 1997. *Working, Shirking, and Sabotage*. Ann Arbor: University of Michigan Press.

Bretschneider, Stuart I., and Wilpen L. Gorr. 1987. "State and Local Government Revenue Forecasting." In *The Handbook of Forecasting*, 2nd ed. Ed. Spyros G. Makridakis and Steven C. Wheelwright. New York: John Wiley and Sons, pp. 118–134.

Bretschneider, Stuart I., Wilpen L. Gorr, Gloria Grizzle, and Earle Klay. 1989. "Political and Organizational Influences on the Accuracy of Forecasting State Government Revenues." *International Journal of Forecasting* 5(June):307–19.

Buchanan, James M. 1963. "The Economics of Earmarked Taxes." *Journal of Political Economy* 71(October):457–69.

Cassidy, Glenn, Mark S. Kamlet, and Daniel S. Nagin. 1989. "An Empirical Examination of Bias in Revenue Forecasts by State Governments." *International Journal of Forecasting* 5(October):321–31.

Crain, W. Mark. 2003. *Volatile States: Institutions, Policy, and the Performance of American State Economies*. Ann Arbor: University of Michigan Press.

Cyert, Richard M., and James G. March. 1963. *A Behavioral Theory of the Firm*. Englewood Cliffs: Prentice-Hall.

Donahue, John D. 2003. "In-and-Outers: Up or Down?" In *For the People: Can We Fix Public Service?* ed. John D. Donahue and Joseph S. Nye Jr. Washington: Brookings Institution, pp. 55–71.

Downs, Anthony. 1967. *Inside Bureaucracy*. New York: HarperCollins.

Franklin, Aimee L., and James W. Douglas. 2003. "Revolving Funds as Budgeting Tools: An Examination of Oklahoma State Agencies." *State and Local Government Review* 35(Spring):90–101.

Gilmour, John B., and David E. Lewis. 2006. "Political Appointees and the Competence of Federal Program Management." *American Politics Research* 34(January):22–50.

Hammond, Thomas H., and Paul A. Thomas. 1989. "The Impossibility of a Neutral Hierarchy." *Journal of Law, Economics, and Organization* 5(Spring):155–84.

Heclo, Hugh. 1975. "OMB and the Presidency—The Problem of 'Neutral Competence.'" *The Public Interest* 38(Winter):80–99.

Heclo, Hugh. 1977. *A Government of Strangers: Executive Politics in Washington*. Washington: Brookings Institution.

Horn, Murray, J. 1995. *The Political Economy of Public Administration: Institutional Choice in the Public Sector*. New York: Cambridge University Press.

Hsiao, Cheng. 2003. *Analysis of Panel Data*. New York: Cambridge University Press.

Huber, John D., and Nolan McCarty. 2004. "Bureaucratic Capacity, Delegation, and Political Reform." *American Political Science Review* 98(September):481–94.

Kaufman, Herbert A. 1956. "Emerging Conflicts in the Doctrines of Public Administration." *Public Administration Review* 50(December):1057–1073.

Klay, William Earle. 1985. "The Organizational Dimension of Budgetary Forecasting: Suggestions from Revenue Forecast-ing in the States." *International Journal of Public Administration* 7(3):241–65.

Krause, George A. 1994. "Federal Reserve Policy Decision Making: Political and Bureaucratic Influences." *American Journal of Political Science* 38(February):124–44.

Krause, George A. 1996. "Agent Heterogeneity and Consensual Decision Making on the Federal Open Market Committee." *Public Choice* 88(July):83–101.

Lewis, David E. 2005. "Presidents and the Politicization of the U.S. Federal Government, 1988–2004." Presented at the annual meeting of the Midwest Political Science Association.

Mackenzie, G. Calvin, ed. 1987. *The In-and-Outers: Presidential Appointees and Transient Government in Washington*. Baltimore: The Johns Hopkins University Press

Maranto, Robert. 1998. "Thinking the Unthinkable in Public Administration: A Case for Spoils in the Federal Bureaucracy." *Administration and Society* 29(6):623–42.

McCarty, Nolan. 2004. "The Appointments Dilemma." *American Journal of Political Science* 48(April):413–28.

[McNollGast] McCubbins, Mathew D., Roger G. Noll, and Barry R. Weingast. 1989. "Structure and Process, Politics and Policy: Administrative Arrangements and the Political Control of Agencies." *Virginia Law Review* 75(March):431–82.

Michaels, Judith E. 1997. *The President's Call: Executive Leadership from FDR to George Bush*. Pittsburgh: University of Pittsburgh Press.

Miller, Gary J. 1992. *Managerial Dilemmas: The Political Economy of Hierarchy*. New York: Cambridge University Press.

Moe, Terry M. 1982. "Regulatory Performance and Presidential Administration." *American Journal of Political Science* 26(April):197–224.

Moe, Terry M. 1985. "The Politicized Presidency." In *The New Direction in American Politics*, ed. John E. Chubb and Paul E. Peterson. Washington: Brookings Institution, pp. 235–271.

Moe, Terry M. 1989. "The Politics of Bureaucratic Structure." In *Can The Government Govern?* ed. John E. Chubb and Paul E. Peterson. Washington: Brookings Institution, pp. 267–329.

Nathan, Richard P. 1983. *The Administrative Presidency*. New York: John Wiley.

National Association of State Budget Officers (NASBO). 2004. *State Expenditure Report: 2003*. Washington: National Association of State Budget Officers.

National Commission on the Public Service. 2003. *Urgent Business for America: Revitalizing the Federal Government for the 21st Century*. Washington: Brookings Institution.

Patashnik, Eric M. 1997. "Unfolding Promises: Trust Funds and the Politics of Precommitment." *Political Science Quarterly* 112(Autumn):431–52.

Plümper, Thomas, and Vera E. Troeger. 2004. "Efficient Estimation of Rarely Changing Variables in Fixed Effects Models." Typescript. University of Konstanz.

Poterba, James M. 1994. "State Responses to Fiscal Crises: The Effects of Budgetary Institutions and Poitics." *Journal of Political Economy* 102(August):799–821.

Rodgers, Robert, and Phillip Joyce. 1996. "The Effect of Underforecasting on the Accuracy of Revenue Forecasts by State Governments." *Public Administration Review* 56(January/February):48–56.

Rourke, Francis E. 1992. "Responsiveness and Neutral Competence in American Bureaucracy." *Public Administration Review* 52(November/December):539–46.

Simon, Herbert A. 1976. *Administrative Behavior*. New York: Free Press.

Snyder, Suzanne K., and Barry R. Weingast. 2000. "The American System of Shared Powers: The President, Congress, and the NLRB." *Journal of Law, Economics, and Organization* 16(October):269–305.

Stimson, James A. 1985. "Regression in Space and Time: A Statistical Essay." *American Journal of Political Science* 29(November):914–47.

Suleiman, Ezra. 2003. *Dismantling Democratic States*. Princeton: Princeton University Press.

Wilson, James Q. 1989. "Bureaucracy*: What Government Agencies Do and Why They Do It.*" New York: Basic Books.

Wood, B. Dan, and Richard W. Waterman. 1994. *Bureaucratic Dynamics*. Boulder: Westview Press.

Wooldridge, Jeffrey M. 2003. *Econometric Analysis of Cross Section and Panel Data*. Cambridge: MIT Press.