

The Consequences of Presidential Patronage for Federal Agency Performance

Nick Gallo, David E. Lewis

Vanderbilt University

ABSTRACT

In this article, we examine the relationship between presidential patronage and federal agency performance. Using Program Assessment Rating Tool (PART) management scores for 1,016 federal programs during the Bush Administration, we compare the performance of federal programs administered by appointees from the campaign or party against programs run by other appointees or career professionals. We introduce new means of overcoming the shortcomings of PART scores in order to make reliable inferences from this measure of federal program performance. We find that federal programs administered by appointees from the campaign or party earn lower PART scores than programs run by other appointees or by career executives. We conclude that although appointing persons from the campaign or party provides presidents an important source of political capital and arguably improves accountability, it also has costs for agency performance.

When presidents reward campaign staff or political supporters with federal jobs does this hurt management performance? From very early in President Obama's tenure, critics have charged that appointments to some key positions as close as the White House and as far away as ambassadorships were made more for political or campaign support than demonstrated competence (Nakamura 2009; Weisman and Hayashi 2009). President Obama, like presidents before him, appears to be using some plum jobs to reward supporters and the politically connected. President Bush was similarly criticized for his appointments of persons like Michael Brown (Federal Emergency Management Agency [FEMA]), Julie Myers (Immigration and Customs Enforcement), and David Safavian (Office of Federal Procurement Policy) to key jobs in the administration despite scant credentials (Eggen and Hsu 2005; Smith and Schmidt 2005; Tumulty, Thompson, and Allen 2005). Yet, presidents generally defend these appointments as appropriate, arguing that they conform to historical norms, have a general (though not specific) competence, and that their proximity to the

A previous version of this article was presented at the 2010 annual meeting of the Southern Political Science Association, Atlanta, GA, January 7–10. We thank Camille Burge, Josh Clinton, Gabe Horton, and Karen Hult for valuable assistance and feedback on this project. Address correspondence to the author at david.e.lewis@vanderbilt.edu.

doi:10.1093/jopart/mur010

Advance Access publication on May 16, 2011

© The Author 2011. Published by Oxford University Press on behalf of the Journal of Public Administration Research and Theory, Inc. All rights reserved. For permissions, please e-mail: journals.permissions@oup.com

president can compensate for a lack of credentials.¹ For example, the White House recently defended the appointment of John Roos as Ambassador to Japan, arguing that President Obama had always said some ambassadorial picks would be political, that Roos would work closely with the President, and that he was very accomplished. The White House did not make any reference to expertise in Japanese history, culture, or politics or knowledge of the important foreign policy questions facing the region.

Cases like these raise the important question of what impact campaign or political supporters have on performance when they receive political appointments. Congress and the press severely criticized Brown for his role in FEMA's flawed response to Hurricane Katrina and provided further evidence that the skills necessary to win campaigns are distinct from those required to govern. Yet, presidents frequently maintain that what some appointees lack in demonstrated credentials, they make up for in general competence and a close connection to the president. Indeed, the connections and experience that come with work for the campaign or party may provide executives precisely what is needed for them to do their job well. An important component of agency leadership is *political* work such as strategic planning, policy making, and building relationships with key stakeholders (Maranto 2008). It is also possible that appointees with few credentials are appointed to precisely those jobs where very little specific expertise is required (Lewis 2009). In ambassadorships, for example, it is often career deputies who do most of the serious foreign policy work.

In this article, we examine the relationship between appointees chosen for campaign or political experience and federal management performance. Using 1,016 Program Assessment Rating Tool (PART) scores—numerical measures of federal program performance used during the Bush Administration—we compare the performance of federal programs administered by appointees from the campaign or party against programs run by other appointed or career executives. We introduce new ways of accounting for shortcomings in the PART scores themselves in order to make reliable inferences from this measure of federal program performance. We find that federal programs administered by appointees from the campaign earn the lowest PART scores, followed by programs directed by other appointees, and programs administered by career professionals. We conclude that although choosing appointees from the campaign or party has important political benefits for parties and presidents, its persistence in American democracy also has deleterious consequences for federal management performance.

PATRONAGE, PRESIDENTIAL APPOINTMENTS, AND MANAGEMENT PERFORMANCE

A number of excellent works describe the different factors that explain presidential appointment decisions such as loyalty, competence, and patronage but also sociodemographic characteristics, geography, relations with Congress, and interest group lobbying (see, e.g., Heclo 1977; Light 1995; Mackenzie 1981). Recent work has particularly emphasized loyalty and, to a lesser extent, competence as the key factors in presidential selection (Edwards 2001; Moe 1985; Weko 1995). This increased presidential control of the process

1 In fact, some argue that wealth can improve ambassadorial performance since many embassies are underfunded (Maranto 2005, 110-11).

and focus on loyalty and competence has been aided by a professionalization of the White House personnel operation (Pfiffner 1996; Patterson and Pfiffner 2001; Weko 1995). Generally, however, for most appointments, a host of factors are considered and weighed together prior to selection. Where appointments can be distinguished is in the weight presidents assign different factors. For a class of appointments, much greater weight is placed on political factors than experience or expertise and these appointees are the focus of attention here.

Although it is generally acknowledged that presidents reward campaign and party staff with jobs as a reward for past work and to induce future support, no large-N studies attempt to evaluate the influence of this practice on management performance (Bearfield 2009; Mackenzie 1981). A number of recent studies focus on the influence of appointees on management performance relative to career professionals. These works describe the differences between appointees and careerists that matter for performance. Specifically, Collins et al. (2006), Gilmour and Lewis (2006), and Lewis (2008) find that programs and agencies administered by appointees perform systematically worse than those run by career professionals. Lewis (2008) finds that career professionals have more direct agency experience and longer tenures on average and both of these factors lead to better management performance. Appointed managers are more likely to have higher education levels and more nongovernment management experience, but these factors had no influence on program performance. Yet, few direct attempts have been made to determine whether appointees selected primarily for campaign experience or connections actually do adversely influence performance.

One obstacle to this line of research is that it is very difficult for scholars to measure US government performance systematically across different contexts. A number of studies have used revenue forecasts (Krause and Douglas 2005, 2006; Krause, Lewis, and Douglas 2006) or responses to employee surveys to measure agency performance (e.g., Brewer and Selden 2000; Choi and Rainey 2010; Chun and Rainey 2005; Oh and Lewis 2008). Others have used PART scores as a measure of management performance (e.g., Gilmour and Lewis 2006; Jung and Rainey 2008, 2009; Lewis 2008). Each of these measures has limitations. Specifically, it is unclear whether budget forecasting agencies are like other agencies and whether results from those studies are generalizable to other agencies. Surveys of federal employees rely on the impressions of federal employees who may or may not be close enough to agency senior leaders to accurately evaluate performance. PART scores were generated by a presidential agency, applied unevenly across agencies, and there is evidence that some programs received higher scores simply because some executives were better at the PART process (see, e.g., Gilmour 2006; Metzenbaum 2009; Moynihan 2006, 2008; Posner and Fantone 2007; Radin 2005; US Government Accountability Office 2004, 2005, 2008). It was also difficult to identify the precise official responsible for program performance (Moynihan 2009). These difficulties can create problems for inference.

Recently, new data have emerged to mitigate some of the shortcomings of the PART scores. Specifically, Bertelli et al. (2008) fielded a confidential survey of 7,448 career and appointed executives across the executive branch in 2007–08 that included questions about the PART process. The survey asked career executives involved in the PART process the extent to which PART scores picked up real differences among programs in their agency. If career professionals in the agencies who are most familiar with the programs and the PART scores believe that the PART scores measure real differences in

program performance, this should give us confidence in using these measures to evaluate comparative management performance. Parsing out only those programs whose PART scores are considered valid by those most knowledgeable should provide an important robustness check on any analyses using PART scores as a measure of performance.

THE PATRONAGE-PERFORMANCE LINK

Existing research has identified two pathways by which appointees systematically influence agency performance relative to career managers. The first pathway is through differences in background and training between these two classes of federal executives. Appointed and career managers have different backgrounds and qualifications (see, e.g., Aberbach and Rockman 2000; Lewis 2008; Maranto 2005). Appointees average fewer years of public management experience. They are less likely to have worked in the agency they manage before being selected to lead it, and they have fewer years of federal government experience overall. The informational asymmetries between appointed managers and subordinates are greater than for careerist managers and their subordinates. These factors influence the ability of appointees to monitor the programs they direct and implement policies they prefer.

Appointees do have the advantage of political connections. They have significantly more political experience and often have a closer connection to political stakeholders than career managers. Their political connections may be parlayed into greater resources for the agency. Appointed executives also average slightly more education and have significantly more private or not-for-profit sector management experience although it is unclear how easily higher levels of education or general management experience translate into improved agency performance due to important differences between managing in the private and public sector.

The second pathway by which appointee management influences performance is through systematic effects on the agency personnel system. Even if the population of appointee managers and career managers is of comparable experience and background, appointee management can have deleterious consequences for agency performance (Lewis 2008). Appointees stay for shorter tenures than career managers (Hecl 1977). This creates greater management turnover and vacancies in programs and agencies run by appointees. Regular turnover makes it difficult for the agency to communicate agency goals, credibly commit to reform, monitor agency activity, and generally poorer performance (Boylan 2004; Hecl 1977; Lewis 2008; O'Connell 2009). Agencies with appointee managers also have a hard time recruiting and retaining top quality career professionals (Gailmard and Patty 2007). When careerists do not have access to the highest paying or most influential positions in agencies, they have less of an incentive to stay or invest the time and effort they would otherwise expend in pursuit of those positions. Careerists see less qualified but politically connected persons taking the top jobs in their agency, and this is harmful for morale. Appointee management influences not only the qualifications and performance of program managers but also the qualifications and motivation of career professionals that work below these program managers.

Appointees from the Campaign or Party versus Other Appointees

The discussion of differences between programs administered by appointees versus those run by careerists does not directly speak to differences among *types* of appointees, but its implications are clear. The sources of the appointee-careerist gap in performance are even more pronounced for appointees selected from the campaign or party. Since these

appointees are selected importantly due to campaign or political experience, the chance that they are selected for demonstrated competence is lower and the skills necessary for managing a campaign differ from those required for governing. This is not to say that all executives from the campaign or party will perform worse than their other appointee colleagues. Indeed, many executives with campaign and party experience are skilled managers. Rather, the chance that patronage appointees have demonstrated expertise or credentials is lower. To the extent that demonstrated credentials or previous experience or expertise are related to agency performance, programs run by patronage appointees should perform worse than programs run by other appointees or careerists.

Patronage appointees are also likely to turn over more quickly than other appointees. Since patronage appointees are drawn from the political world, they are also more likely to make career decisions with a future career in mind. Rather than view a political appointment as a capstone to a long career, patronage appointees tend to be younger and seeking other jobs in the party, a future administration, or with a firm or group with a close connection to the president's party. This leads to shorter tenures on average. Agencies that are regularly on the receiving end of patronage appointments have higher turnover and vacancy rates.

Since patronage appointees are likely to have fewer qualifications than other appointees, the effects of their selection on executive career choices are more pronounced. Some agencies are fortunate enough to be administered by a string of very talented appointed managers. Others like FEMA are less fortunate, receiving third and fourth tier political types consistently (Lewis 2008, 150). For example, in the aftermath of President George W. Bush's initial appointment of Joseph Allbaugh to FEMA, one career manager said, "There are plenty of Republican emergency managers, fire chiefs, or police chiefs around. And they pull this guy who's a campaign manager?" (Klinenberg and Frank 2005). A large number of experienced career professionals left FEMA during the first 2 years of the Bush Administration. Although the increasing penetration of appointees into agency management can have deleterious consequences for efforts to recruit and retain career professionals, this effect is even more dramatic when those appointees selected are noticeably less qualified than career professionals who otherwise would hold those jobs.

In total, appointees who obtain jobs through work for the campaign or party should be less qualified than other appointees and careerists. They should stay for shorter tenures and programs they run should perform systematically worse than other programs.

DATA, VARIABLES, AND METHODS

To evaluate the relationship between patronage appointees and performance, we use PART scores, a performance measurement scheme devised by the Office of Management and Budget (OMB) during the George W. Bush Administration. Under the PART system, budget examiners in the OMB in cooperation with agency officials graded almost all federal programs (98%) one time over a 7-year period between 2002 and 2008. OMB worked with agencies and gave grades from 0 to 100 for different aspects of program performance based upon a series of 25–30 yes/no questions. Answers to the questions provide the basis for numerical scores in four categories of performance (program purpose and design, strategic planning, program management, and program results), a weighted total score, and an overall program evaluation (ineffective, results not demonstrated, adequate, moderately effective, effective). The questions and the grading scheme were devised through the Federal Advisory Commission Act process and with the input of numerous parties including the National Academy of Public Administration, Congress, and other interested parties. In

total, between 2002 and 2008, 1,016 federal programs were graded. The average total grade was 66.92 (18.34 standard deviation [SD]; minimum 10; maximum 100), and the ratings percentages were ineffective (2.56%), results not demonstrated (17.03%), adequate (29.33), moderately effective (32.09%), and effective (19.00%).

To determine whether programs were administered by appointees or careerists, we relied on the PART worksheets.² These worksheets list both the agency and the bureau that house each program. We consider the head of the bureau as the person responsible for the program unless the program itself is an agency or is obviously part of a specific subagency.³ To determine whether the person was an appointee (either Senate-confirmed or in the Senior Executive Service) or a careerist, we relied on information from the *Federal Yellow Book*.⁴ In total, we could find biographical information about managers of 977 of the 1,016 federal programs. Of these 977 federal programs, 8.80% (86) were managed by career executives (Career appointment), 11.26% (110) by politically appointed members of the Senior Executive Service (Noncareer appointment), and 79.43% (776) by Senate-confirmed appointees (Presidential appointment with Senate confirmation). There are also five programs administered by personnel from agencies with their own personnel system, making it difficult to classify them as either career professionals or political appointees.⁵ For many programs, the easiest official to identify was a cabinet secretary, either because a bureau was not listed or a program cut across several bureaus. It is unlikely that the secretary was directly involved in the day-to-day operations of these programs, so we also evaluate the data below without these cases as a robustness check.

There were 358 different managers of the 977 programs.⁶ We collected background information on each of the 358 managers from a variety of sources including data from Lewis (2008), publicly available biographies, and the *Federal Yellow Book*.⁷ Among the data we

2 For an example, see <http://www.whitehouse.gov/omb/expectmore/detail/10002014.2004.html> (accessed August 3, 2010). For the complete data, see <http://www.expectmore.gov>.

3 For example, the Western Area Power Administration (WAPA) is part of the Power Marketing Administration in the Department of Energy. WAPA was considered to be a program for evaluation by the OMB. Rather than consider the head of the Power Marketing Administration as the agency head, we considered the head of WAPA itself as the head. Similarly, if a program involved Native American health but the worksheet only listed the program as being part of the Department of Health and Human Services, we considered the head of the Indian Health Service as the program manager. In some cases, the worksheet lists no bureau. It either leaves that entry blank or it lists the agency name again. In these cases, we list the agency head as the program manager. We have also estimated models excluding cases where no bureau is listed and the results look similar to what is reported here. The results are included in Appendix 1.

4 The *Federal Yellow Book* is published by Leadership Directories, Inc. It lists contact information for agency officials, their appointment authority, and often biographical information. We assume that all persons are careerists unless otherwise indicated by the *Yellow Book*.

5 There were two managers overseeing five programs. We exclude these two managers and five cases from the main analyses. In general, programs administered by these two managers from agencies with their own personnel system performed poorly. Their total PART score was 52.27. We have also estimated models simply controlling for these cases and the estimates are virtually identical to what is reported here. The coefficient on these cases is large and negative, as expected. We exclude them here because we do not have enough cases to make reliable inferences about programs in agencies with their own personnel systems. The results are available from the authors upon request.

6 The SD for total PART scores overall is 18.34. The average SD for PART scores by manager is 6.47. This suggests that, although there is variance in program PART scores within manager, it is smaller than the overall SD. This implies that managers that direct programs with high scores are less likely to direct programs with very low scores. Managers that direct programs with very low scores are unlikely to also direct programs with very high scores. Since each program in the data set has been evaluated once, there are no cases where a program has been managed both by an appointee and careerist.

7 Lewis data available at <http://people.vanderbilt.edu/~david.lewis/data.htm>.

collected was information about the campaign experience of the manager, including whether or not the manager worked for the campaign (5.08%) or previously worked for the party (3.49%).⁸ The *Yellow Book* lists 8 managers of 57 programs as having worked on the campaign staff and 16 managers overseeing 84 programs as having worked either for the campaign or the national party prior to their appointment. As the descriptive statistics suggest, the number of managers whose publicly available biographical information mentioned work on the campaign is limited. However, a significantly larger number of managers in the administration were associated in some way with the campaign. As a robustness check below, we also evaluate the data by subdividing the data into appointees with *any* public connection to the campaign and appointees with no public connection to the campaign (18.50%).⁹

As a starting point, Table 1 lists differences in background characteristics among appointees from the campaign, appointees not from the campaign, and careerists. Indicators are included to identify statistically distinguishable differences between (1) appointees who did not work on the campaign and careerists and (2) appointees from the campaign and appointees not from the campaign. Important differences are apparent among the three groups. Appointees from the campaign have systematically less education and more political experience than either career professionals or other appointees. Like other appointees, they have less agency experience, shorter tenures, and more private sector management experience than career professionals. They manage fewer employees but a greater number of programs than other types of managers which suggests they manage a greater number of smaller programs.

The differences between appointees from the campaign, other appointees, and career executives appear to matter for performance. Figure 1 includes histograms of total PART scores by type of program manager. The figure illustrates what is suggested in the bottom rows of table 1. Programs administered by appointees from the campaign have systematically lower part scores (55) than other appointees (69) and career professionals (73). This first cut at the PART score data suggests that the naming of appointees from the campaign to administer federal programs and agencies hurts program performance. The skills these appointees exercised working on the campaign do not directly translate into program management.

Controls for Agency and Program Characteristics

Of course, these observed differences in PART scores may be due to factors other than the quality of the managers. To account for this, we estimate a series of models that control for a host of agency, program, and year-specific factors that may influence a program's PART score and be correlated with whether or not the program manager is a career professional or appointee of some type. One difficulty with comparing program performance across a wide variety of programs is that programs differ dramatically in what they do and their size.

⁸ Many other members of the paid campaign staff were rewarded with jobs in the administration but few were directly placed into jobs as agency heads or program managers. We include an indicator for campaign experience (0,1) in some models and an indicator for campaign or party experience (0,1) in what follows. We prefer the latter measure since it includes a greater number of managers.

⁹ To construct this measure, we searched for any reported connection to the campaign in newspapers across the country. We conducted a Lexis-Nexis search of the manager's name within a limited number of words to "Bush" and "campaign," "rally," "introduced," "advisor," or "adviser" and read through all relevant articles. In all cases where an article indicated a connection between the person and the campaign we coded them with a 1. All other managers were coded with a 0.

Table 1
Background Characteristic by Manager Type

Variable	Careerist	Appointee from Campaign	Other Appointee
<i>Manager background</i>			
Education (0–3)	2.07	1.67#	2.16
Previous bureau experience (0,1)	0.85	0.50	0.50*
Experience in another federal department (0,1)	0.45	0.50	0.42
Public management experience (0,1)	0.98	1.00	0.92#
Private sector management experience (0,1)	0.26	0.50	0.56*
Tenure as bureau chief (Years)	3.65	2.03	2.36*
Worked in Congress (0,1)	0.00	0.20	0.15*
Worked in the White House (0,1)	0.00	0.30*	0.06*
<i>Management Environment</i>			
Programs graded (1–3)	1.17	3.65*	1.73*
Average budget of program graded (millions)	\$520	\$959	\$4899
Bureau employment (10–222,715)	7,191	856	13,788
Number of Appointees in Agency	2.4	5.1	10.2*
<i>PART Score</i>			
Program Purpose and Design	88.19	79.53*	87.59
Strategic Planning	79.61	62.40*	78.89
Program Management	88.80	79.02	84.09*
Program Results	59.50	34.73*	52.92*
Average Total PART grade (0–100)	73.07	55.24*	68.63*

Note: N = 351. Education levels (1–4) are bachelors or lower, masters, MD or DDM, and doctorate. There are only 8 managers who came from the campaign and they manage 57 programs. Previous bureau experience indicates work for the bureau prior to becoming its head. For difference of means tests, the “Other Appointee” is column compared to “Careerist” column and the “Appointee from the Campaign” column compared to Other Appointee column.

*Difference in means significant at the .05 level in one-tailed tests; #Difference in means significant at the .10 level in one-tailed tests.

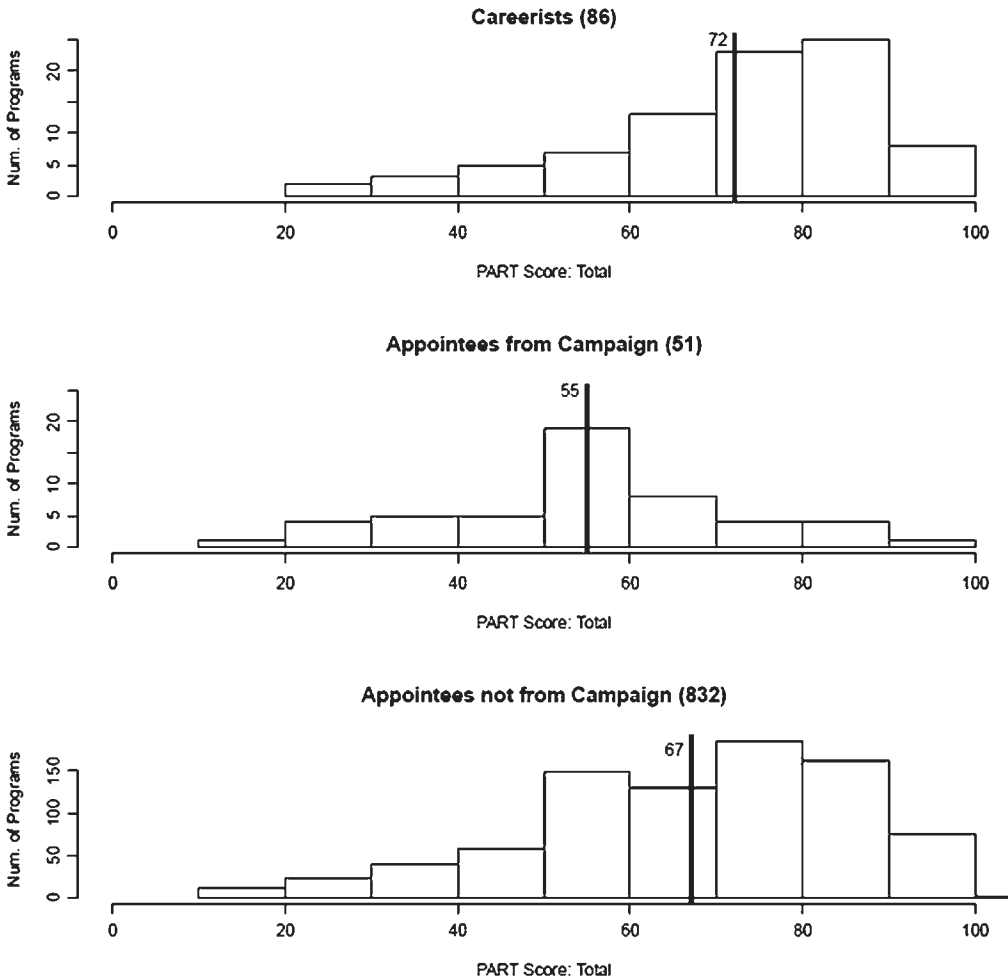
Some programs regulate (e.g., food safety, pest and disease management), others provide credit (e.g., small business development, veterans home loans), and others deliver direct services to clients (e.g., Head Start, Smithsonian operations). To account for these differences and compare different types of programs against each other, we include indicators for each program type—block/formula grant (15.9%), capital assets and service acquisition (8.4%), competitive grant (17.4%), credit (3.8%), direct federal (36.1%), research and development (11.1%), and regulatory (7.3%) programs.¹⁰ The models also control for the natural log of program budget since larger programs may be harder (or easier) to manage and the size of programs differs by type of manager.

Important contextual differences exist among agencies that implement federal programs that may also influence PART scores, including structural features such as the presence of fixed terms (4.8%) and whether an agency is a commission (7.4%) or administration.¹¹

10 These program categories are the primary classifications for each program. Some programs are listed as more than type. There was one program whose type was listed as “mixed.” This program was excluded.

11 We have also estimated models that include indicators for the cabinet departments. In these models, the coefficient on work for the campaign or party is negative but smaller and only marginally significant in two-tailed tests ($p < .11$). The coefficients are still substantively notable (–4 to –6 points on the total PART score). The results suggest that cabinet departments that house program managers that worked on the campaign also house programs that are managed poorly. Is it the department culture that contributes to poor program performance or the patronage appointees that contribute to this culture that lead to poor program performance? The model estimates suggest both. These results are included in Appendix 1.

Figure 1
Total PART Score by Manager Type



Agencies with these features are designed to be insulated from presidential control and this can influence the difficulty or ease of management. Some bureaus manage a greater number of programs than others. This can influence program performance and varies by type of manager. To account for this, we include a control for the number of programs an agency or bureau has evaluated in the year the program is being evaluated (mean 3.5; SD 3.47; minimum 1; maximum 31).

The policy content of what agencies do can also influence the evaluation of specific programs. Some agencies, by virtue of mission, history, and personnel, are more liberal or conservative than others. If PART scores are politicized, these differences in agency ideology will influence the grades programs receive. To measure the ideology of agencies, we use preference estimates created by Clinton and Lewis (2008). Clinton and Lewis conducted an expert survey about whether agencies tended to be liberal, conservative, or neither consistently and used responses to generate estimates, adjusting for different

definitions of liberal and conservative and the quality of the ratings. Some prominent liberal agencies include the Departments of Education, Housing and Urban Development, and Labor. Prominent conservative agencies include Treasury and Defense and notable moderate agencies include the departments of Agriculture and State. Since PART scores are generated by a Republican Administration, programs in liberal agencies will get systematically lower grades if the grades are politicized.

Finally, we include indicators for year a program was assessed since the average PART score may increase over time as agency officials become familiar with the system and more or fewer patronage appointees may serve as the president's tenure progresses.

Methods

We estimate a series of regressions where the unit of analysis is a federal program,¹² the dependent variable is the total PART score,¹³ the key independent variable is the type of manager (careerist, appointee from campaign, other appointee), and the models include a series of agency and program-specific controls. Since bureau chiefs often manage multiple programs and bureaus house multiple programs, we report robust standard errors.¹⁴

One criticism of the PART scores is that they have very little connection to real performance. To help address these concerns, we use data from a recent survey of federal executives that asked the executives themselves about the validity of the PART scores for programs in their agency. In 2007–08, the Survey on the Future of Government Service asked career professionals involved in the PART process the following question,

To what extent did the PART pick up real differences in program performance among programs in your agency? [Almost always reflected real differences (2.62%), generally reflected real differences (14.94%), sometimes reflected real differences (26.58%), rarely reflected real differences (22.70%), PART scores have no connection to real performance (14.18%), don't know (18.99%)].

What is notable about the response to this question is the low percentage overall that believed PART scores measured real differences across programs in their agencies. This gives credence to concerns articulated about these scores and their use in the budgetary process (Gilmour 2006; Metzenbaum 2009; Moynihan 2006, 2008; Posner and Fantone 2007; Radin 2005; US Government Accountability Office 2004, 2005, 2008).

Although the PART scores were less discriminating in some agencies, they were adjudged to pick up real differences in other agencies. This provides a useful means of

¹² We have also estimated models where the manager is the unit of analysis and discuss these below. Since programs are embedded in agencies accounting for the hierarchical structure in the data would be natural. The difficulty with estimating such models in this case is that it is very demanding on the data to estimate an intercept and slope for each agency. Given that there are only 16 managers who were affiliated with the campaign, there are only 1 or 2 such appointees at most in any given agency. Estimating coefficients on the patronage variable by agency with 1 or 2 cases is infeasible. We have, however, estimated hierarchical linear models to test for the influence of appointees generally on program performance (excluding the patronage variable). The average estimated effect of appointees on PART score is sizeable and significant ($B = -4.73$ and $SE = 2.07$) and confirms what is reported in the text.

¹³ We have also estimated models using the overall categorical grade for each program. These models look similar to those in table 2 and are included in Appendix 1.

¹⁴ We have also estimated models where SEs are adjusted for clustering on bureau or manager. These models confirm what is reported here. There are a few outlier cases that may be influential according to leverage plots (`lvr2plot` in STATA 10.0). We have also estimated models excluding these cases and the results confirm what is reported here.

gleaning only the best information from the PART scores. The tables below include columns for models estimated on all PART scores and models estimated on a subset of models using only the most reliable PART scores according to career professionals. Specifically, we estimate models on the subset of agencies where more than half report that PART scores, almost always, generally, or sometimes reflect real differences among programs in their agencies.¹⁵ If the relationship between manager type and performance remain even in the subset of most reliable PART scores, this provides additional confidence that the relationship is not an artifact of some component of the PART scores themselves.

There are two other notable difficulties in model estimation. First, whether or not a program is administered by a careerist, an appointee, or an appointee from the campaign may be endogenous. For example, it could be that appointees from the campaign are selected to run the most difficult and politically sensitive programs. If this is the case, the low PART scores we see in figure 1 for appointee-run programs may be due to the difficulty running the program and not the management ability of the appointee. Of course, the evidence in table 1 suggests that appointees from the campaign or party have fewer credentials than other appointees so it would be somewhat surprising if such persons were selected to run the most challenging programs. All models try to control for relevant aspects of the management environment such as program size and type of program. We have also estimated models on a subset of the data where the appointment authority of the manager is estimated along with the total PART score via two-stage least squares. In these models, we could not reject the null that the ordinary least-squares (OLS) coefficients were the same as those in the instrumental variables regressions ($p < .82$). We focus on the OLS estimates here and include the two-stage least-squares estimates in Appendix 1.¹⁶

A second potential difficulty would emerge if presidents wanted certain programs to fail, particularly programs that the president opposed. If this were the case, a correlation might exist between some appointees and low PART scores not because of poor management but because presidents appointed these managers to hurt these programs. We have estimated models on liberal and conservative agencies separately to see whether the relationship between manager type and PART score remains. If the Republican president wanted programs to fail, this should be most apparent for programs in liberal agencies (e.g., regulatory and social welfare programs). Model estimates look generally similar for liberal and conservative agencies, however, suggesting that the difference in PART scores among manager types is not influenced by systematic presidential attempts to make programs fail.¹⁷

RESULTS

Table 2 includes estimates from four models of PART scores. Models 1 and 2 are estimated on all PART scores, and Models 3 and 4 are estimated using only on the most reliable PART scores according to career professionals. Model 5 includes estimates from a model where the

15 We have also used different cutoffs (i.e., agencies where two-third or three-fourth of careerists have these responses) and the bivariate patterns remain ($p < .00$ for two-third cutoff; $p < .08$ for three-fourth cutoff). More fully specified models are difficult to estimate since reducing the number of cases also reduces the number of managers with previous campaign or political experience.

16 Details of the estimation are included in Appendix 2.

17 These estimates are included in Appendix 1.

manager is the unit of analysis and the manager's average PART score is the dependent variable. These models are estimated to determine whether the relationship between manager type (careerist, appointee from the campaign, other appointee) and PART score so visible in figure 1 is real or the result of other intervening variables. In general, the estimates confirm the basic result. Federal programs administered by appointees from the campaign earn the lowest PART scores. Programs directed by other appointees earn the next lowest evaluations and programs administered by career professionals earn the highest PART scores. This general pattern is robust even when the analysis is limited to the PART scores that career professionals have identified as reliably differentiating programs on true performance.

Substantively, programs administered by appointees are estimated to earn PART scores from 4.69 (Model 2) to 6.74 (Model 3) points *lower* than other programs. The size of this effect is consistent with what has been reported in other work (Lewis 2008) and provides additional evidence that appointee management has costs as well as benefits. Although appointee management may be useful for improving accountability and advancing the president's agenda, these results suggest appointee management also has systematic harmful effects on program performance on average. The differences in performance may stem from differences in the qualifications of appointees versus careerists or the long-term effects appointee penetration has on executive turnover, incentives to invest in expertise, and efforts to recruit and retain career professionals.

When appointees have backgrounds that include work for the campaign or for the national party, the difference in PART scores is even more dramatic. Federal programs managed by campaign or party appointees are estimated to earn PART scores between 9.10 and 13.47 points lower than programs administered by other appointees, depending upon the model and the measure of campaign work included in the specification.¹⁸ Programs administered by career professionals are estimated to receive PART scores close to 1 SD higher than programs administered by appointees with campaign or party experience. On its face, this is important evidence that some of the observed harmful effects of appointee management are due to the types of persons appointed. Appointees selected more for campaign/party experience than competence perform worse than other appointees and career managers.

The differences in PART scores by manager type remain even when controlling for a host of agency and program-specific characteristics. Several controls have notable effects on PART scores. Programs with a larger budget receive higher PART scores. A 1% increase in a program's budget is estimated to increase a program's PART score by about 1.2–1.3 points in the models including all PART scores (Models 1 and 2) and 0.65–0.73 in the models estimated on only the most reliable scores (Models 3 and 4). This may be due to the fact that budget increases may allay one cause of program problems—underfunding. Some differences among program types appear to systematically influence PART scores. Block/formula grants are estimated to get systematically lower PART scores while research and development programs higher scores relative to regulatory programs (base category). The former is estimated to earn PART scores that are 5 points lower, whereas the latter is estimated to earn scores 8 points higher in Models 1 and 2. These effects are smaller and estimated less precisely in the models using only the most reliable PART scores (Models 3

18 Of course, each of these coefficients is estimated with error, making the true effect possibly smaller or larger. For the sake of simplicity, however, we will refer to the point estimates.

Table 2
Federal Program PART Score by Appointees from Campaign

	(1)	(2)	(3)	(4)	(5)
<i>Appointment Authority</i>					
Appointee (0,1)	-4.96** (1.99)	-4.69** (1.99)	-6.74** (2.35)	-6.60** (2.36)	-5.59* (2.99)
Appointee from campaign (0,1)	-13.47** (2.61)	—	-12.26** (2.85)	—	—
Appointee from campaign or work with national party (0,1)	—	-9.10** (2.02)	—	-9.18** (2.09)	-12.67** (4.73)
<i>Program Characteristics</i>					
Ln(Program Budget)	1.20** (0.31)	1.27** (0.31)	0.65* (0.39)	0.73* (0.39)	0.37 (0.61)
Block/Formula Grant (0,1)	-5.16* (2.89)	-5.18* (2.92)	-4.51 (3.59)	-4.59 (3.61)	-1.47 (5.44)
Capital Assets and Service Acquisition (0,1)	2.73 (3.07)	3.52 (3.11)	2.95 (4.02)	4.09 (4.07)	3.53 (6.32)
Competitive Grant (0,1)	-1.66 (2.68)	-1.58 (2.70)	-2.99 (3.43)	-3.13 (3.44)	2.46 (5.65)
Credit (0,1)	0.76 (3.51)	1.02 (3.56)	1.00 (4.16)	0.75 (4.15)	2.15 (6.02)
Direct Federal (0,1)	1.20 (2.51)	1.49 (2.53)	3.59 (2.97)	3.71 (2.99)	5.63 (4.31)
Research and Development (0,1)	8.07** (2.67)	8.29** (2.68)	4.45 (3.89)	4.60 (3.91)	10.97* (6.43)
<i>Bureau Characteristics</i>					
Fixed Term (0,1)	5.50* (3.07)	4.50 (3.24)	0.09 (4.54)	-0.90 (4.70)	2.45 (5.21)
Commission (0,1)	7.23** (2.41)	6.80** (2.49)	5.58 (3.94)	4.97 (4.07)	8.40* (4.34)
# Programs Evaluated	0.30* (0.17)	0.22 (0.17)	0.56** (0.24)	0.52** (0.24)	-0.42 (0.73)
<i>Political Characteristics</i>					
Agency Ideology (Liberal- Conservative)	1.95** (0.66)	2.40** (0.66)	3.01** (0.90)	3.62** (0.89)	3.84** (1.27)
Constant	57.87** (3.68)	57.57** (3.69)	59.25** (4.95)	58.89** (4.94)	68.26** (5.23)
Only Most Effective PART Scores?	No	No	Yes	Yes	Yes
N	962	962	591	591	208
F(19, 330)	12.09**	10.84**	7.86**	7.60**	3.42**
R ²	0.17	0.16	0.19	0.18	0.13

Note: Unstandardized regression coefficients and robust SEs reported. Omitted category for appointment authority is career member of the Senior Executive Service. Omitted category for program type is regulatory programs. Year assessed indicator estimates omitted. Model 5 estimated using managers rather than programs as the unit of analysis.

*Significant at the $p < .10$ level; **Significant at the $p < .05$ level in two-tailed test.

and 4). The same pattern holds for programs in commissions. Programs in commissions are estimated to get PART scores that are about 7 points higher than programs in other agencies when all PART scores are examined, but these effects diminish in the models using only the subset of PART scores (Models 1 and 2 vs. 3 and 4). Agencies that have had a greater number of programs evaluated are estimated to do slightly better. This could be a function of the fact that experience with the PART process improves scores (Gilmour 2006).

The most striking finding among the controls is that agency ideology is consistently related to PART scores, with programs in conservative agencies earning significantly higher PART scores than programs in other agencies. A 1 SD increase in the conservatism of an agency is estimated to increase the average PART score by 2–4 points. More concretely, the estimates suggest that programs in conservative agencies such as the Department of the Treasury, the Department of Defense, or the Millennium Challenge Corporation get PART scores that are 2–4 points higher than programs in the moderate Department of State or Department of Agriculture and 4–8 points higher than programs in the liberal Departments of Labor, Education, or Housing and Urban Development. This implies either that these programs are systematically more effective or that PART scores are politicized. Disentangling whether the Bush Administration rated conservative programs highly because conservative programs were more effective or because OMB officials were biased is difficult. This is an issue that could have been revisited in future work had the Obama Administration maintained the PART system. For example, if liberal programs were rated more highly in the Obama Administration, this would lend credence to the view that the PART scores are biased. If conservative programs were still rated more highly in the Obama Administration, this would be evidence that programs in conservative agencies were indeed more effective. Unfortunately, the Obama Administration has chosen not to continue the PART scores and so the analysis can only include data from the Bush Administration. Future research may be able to take advantage of performance measurement efforts at the state level or future administrations to further evaluate whether and when such schemes are politicized.

Robustness Checks

The preceding analysis of the PART scores reveals that programs administered by appointees from the campaign or party received the lowest scores. Yet, the models include programs ostensibly run by a department secretary when these officials are unlikely to direct the day-to-day activities of the program. There are also only a limited number of managers from the campaign or party from whom to make inferences. Both problems threaten the validity of the analysis. To remedy these problems, table 3 includes models that are estimated on a sample that excludes programs previously assumed to be run by department secretaries and including a new measure of campaign connections. The models include an indicator for whether the manager had *any* public connection to the campaign in newspaper reports prior to the election (0,1: 18.50%).

The models generally confirm the estimates in table 2. In models with and without programs ostensibly run by cabinet secretaries programs run by appointees are estimated to earn lower PART scores. The coefficients are negative and significant at the .10 level or better in all four models. Substantively, appointee-run programs are estimated to earn scores 3.63 (Model 3) to 6.74 (Model 2) points lower than programs administered by career professionals. Programs run by appointees with some public connection to the campaign are estimated to perform worse than programs administered by other appointees. The

coefficient estimates are negative in all four models but significant only in those models without programs listed as being administered by cabinet secretaries (Models 3 and 4). Substantively, these models suggest that programs run by appointees affiliated with the campaign earn PART scores 8.14–8.53 points lower than programs run by other appointees. The new models reconfirm that appointees selected partly for campaign work earn lower PART scores when models are estimated on the most reliable data.

Why Patronage Hurts Management

Why do programs run by appointees from the campaign or party get lower PART scores? One possible reason is systematic differences in the background or qualifications of different types of managers. The summary of background characteristics in table 1 revealed that appointees from the campaign serve shorter tenures, have less education, and more political experience (i.e., work in the White House or Congress) than other appointees. Table 4 includes estimates from regressions where PART scores are regressed on the background characteristics of managers along with agency and program-specific characteristics. We code each manager according to whether they have worked in the agency bureau *prior* to becoming its head (0,1), how long they have served in their current position as head (2.4: mean, 0: minimum, 16: maximum), and whether they have had public management experience (0,1). We also include indicators for masters or doctorate-level education. The models include indicators for whether the bureau chief has worked outside the bureau in another federal department (0,1), has private management experience (0,1), and has worked in either Congress (0,1) or the White House (0,1). The models in table 4 include the same controls as in tables 2 and 3. Models are estimated, once again, on the full sample of PART programs and the subset of those deemed most reliable by career professionals.

The results are suggestive with regard to managerial tenure, education, and political experience. First, the estimates reveal that the longer a manager has been in their position at the time that a program is evaluated, the higher the PART score. The coefficient estimates are positive and significant at the .10 level or better in both models. Each additional year of managerial tenure is estimated to increase a program's PART score by 0.6–0.7 points.¹⁹ One explanation for this finding is that programs experiencing more frequent executive turnover perform worse. Programs run by appointees from the campaign may earn lower PART scores because these appointees serve shorter tenures on average. This finding is consistent with previous work demonstrating that executive turnover generated poorer performance (Boylan 2004, Lewis 2008).

Of course, another explanation is that poor performance generates short tenures rather than the other way around. More generally, it is possible that the same factors that lead programs to have low PART scores also determine whether managers have long or short tenures. To account for the possibility that tenure is endogenous, we estimated a simple instrumental variables regression.²⁰ The results were substantively similar and we could not

19 We have also estimated models allowing for different functional forms for the relationship between tenure and PART score and could not reject the null of a linear relationship.

20 For these models to be identified, exogenous regressors must be identified. In this case we employ two exogenous regressors, the year of the manager's hire and the year the program was assessed. Both regressors predict managerial tenure but should not influence PART score directly. The latter regressor is uncorrelated with the subset of best PART scores with one notable exception. The first set of PART scores was notably lower than other years. In these models, we exclude the programs assessed in the first year (2002).

Table 3
Federal Program PART Score by Appointees Associated w/Campaign

	(1)	(2)	(3)	(4)
<i>Appointment Authority</i>				
Appointee (0,1)	-4.75** (1.96)	-6.74** (2.34)	-3.63* (2.01)	-4.59* (2.38)
Appointee associated with campaign in press (0,1)	-1.15 (1.66)	-0.30 (1.85)	-8.53** (3.32)	-8.14** (3.63)
<i>Program Characteristics</i>				
Ln(Program Budget)	1.18** (0.31)	0.56 (0.39)	0.91** (0.36)	0.11 (0.46)
Block/Formula Grant (0,1)	-5.91** (2.96)	-5.97 (3.69)	-6.21* (3.31)	-8.00** (4.01)
Capital Assets and Service Acquisition (0,1)	2.99 (3.16)	2.60 (4.16)	3.25 (3.49)	0.03 (4.59)
Competitive Grant (0,1)	-2.75 (2.73)	-5.64 (3.50)	-2.91 (3.04)	-7.97** (3.76)
Credit (0,1)	0.81 (3.54)	0.45 (4.21)	2.56 (3.66)	-1.09 (4.18)
Direct Federal (0,1)	1.54 (2.56)	3.52 (3.05)	0.37 (2.85)	0.71 (3.17)
Research and Development (0,1)	8.53** (2.70)	4.34 (3.90)	8.85** (3.02)	2.05 (4.22)
<i>Bureau Characteristics</i>				
Fixed Term (0,1)	3.72 (3.43)	-1.91 (5.03)	4.17 (3.17)	-1.61 (4.72)
Commission (0,1)	5.62** (2.61)	3.10 (4.33)	5.47** (2.50)	2.34 (4.11)
# Programs Evaluated	0.13 (0.17)	0.36 (0.24)	-0.64** (0.25)	-0.91** (0.36)
<i>Political Characteristics</i>				
Agency Ideology (Liberal-Conservative)	2.26** (0.66)	3.65** (0.89)	2.23** (0.78)	3.85** (1.04)
Constant	58.63** (3.70)	61.06** (4.96)	60.94** (4.01)	66.28** (5.30)
Include Secretaries?	Yes	Yes	No	No
Only Most Effective PART Scores?	No	Yes	No	Yes
N	962	591	744	438
F(19, 942; 19, 571; 19, 693; 19, 387)	8.49**	6.13**	9.00**	7.07**
R ²	0.17	0.16	0.18	0.22

Note: Unstandardized regression coefficients and robust SEs reported. Omitted category for appointment authority is career member of the Senior Executive Service. Omitted category for program type is regulatory programs. Year assessed indicator estimates omitted.

*Significant at the $p < .10$ level; **significant at the $p < .05$ level in two-tailed test.

reject the null hypothesis that the coefficient estimates in the simple OLS regression were the same as those in the two-stage least squares.²¹

Although there were notable differences in education between appointees from the campaign and other managers, education was estimated to have mixed effects on performance. Although three of four coefficients were in the expected direction, in one case having a master's level education was estimated to decrease a program's PART score relative to managers with only a BA or high school education. Programs administered by persons with doctorates were estimated to get PART scores 3.69 points higher than other programs. One explanation for the mixed results is that only the very most talented people from the set of employees with low levels of education make it into management positions. This would dampen the influence of master's level education on performance. Managers with a master's level of education may be able to get management positions without the same type of screening mechanism that applies to those without degrees.

The estimates for coefficients on previous work in the White House or in Congress are negative in all cases, and the coefficient on work for the White House is significant and large in the second model. Programs administered by appointees with previous work in the White House are estimated to earn PART scores 8.44 points lower than other programs. This result suggests that patronage-type appointees may be more attentive to political concerns or patrons than management in their agency. If they are focused more on advancing their careers or the interests of the party or patrons, management in their agency may suffer. The results suggest that the skills required to do the political work of the White House do not directly translate to federal program management.

An alternative explanation is that presidents may prefer to send appointees from the White House to run programs with low PART scores to improve their performance. We cannot rule out this possibility, particularly since we know anecdotally that Margaret Spellings left the White House to run the Department of Education, a department with very low PART scores. That said, the President also asked Condoleezza Rice to head the State Department, an agency with significantly higher overall evaluations. Indeed, if presidents are concerned about poor performance, loyalty is important but so is competence. Presidents need appointees who know how to effectively run a government agency to turn poor-performing programs around. Given these facts, it is unlikely that patronage-type appointees are selected to run the poorest performing programs to improve their performance.

Interestingly, and contrary to previous work, we could not reject the null that previous agency experience had no influence on PART score (Lewis 2008). Previous management experience in either the private or public sector was also negatively correlated with PART scores in the two models. There are a couple of possible explanations for these findings. First, almost all managers in the sample had previous public management experience which suggests the measures of agency or management experience do not adequately parse out different types and levels of public management experience. Indeed, those managers with greater tenure in their agency management positions (i.e., public management experience) earned significantly higher PART scores, which suggests that such experience helps. The lack of variation in the public management variable also means that the few cases without public management experience exert tremendous leverage in model estimation, and these cases may not be representative of the overall population of managers. The results

21 See Appendix 3 for these estimates.

Table 4
Federal Program PART Score by Manager Background

	All PART Scores	Most Effective Part Scores Only
<i>Manager Background Characteristics</i>		
Bureau Experience (0,1)	-0.56 (1.33)	-1.51 (1.88)
Months Serving as Bureau Chief (0 to 182)	0.71** (0.28)	0.67* (0.37)
Previous Public Management Experience (0,1)	-4.68 (3.40)	-7.02* (4.05)
Masters (0,1)	1.49 (1.43)	-3.34* (1.78)
Doctorate (0,1)	3.69** (1.85)	4.12* (2.45)
Worked in Another Department (0,1)	0.99 (1.23)	1.53 (1.76)
Private Management Experience (0,1)	-4.33** (1.30)	-5.40** (1.93)
Worked in White House (0,1)	-2.27 (2.52)	-8.44** (2.96)
Worked in Congress (0,1)	-2.96* (1.71)	-2.82 (2.44)
<i>Bureau Characteristics</i>		
Fixed Term (0,1)	3.22 (3.21)	1.47 (4.24)
Commission (0,1)	4.57* (2.76)	2.36 (5.13)
# Programs Evaluated	0.16 (0.17)	0.44* (0.24)
<i>Political and Program Characteristics</i>		
Agency Ideology (Liberal-Conservative)	2.61** (0.66)	4.50** (0.93)
Ln(Program Budget)	1.45** (0.33)	0.82* (0.42)
Constant	70.57** (4.44)	75.60** (5.48)
N(Observations, Managers)	878	531
F(25, 282; 25, 49)	8.42**	6.40**
R ²	0.19	0.23

Note: Unstandardized regression coefficients and robust SEs reported. Program type and year assessed indicator estimates omitted. *Significant at the .10 level; **significant at the .05 level in two-tailed test.

regarding private management experience are consistent with some past research which suggests that private management experience does not necessarily translate in the public sector and inconsistent with other research which suggests it is helpful (see, e.g., Lewis 2008; Oh and Lewis 2008). Differences in management environment may help explain why private sector management experience helps in some cases like the Defense Department (Oh and Lewis 2008) and not in others.

DISCUSSION

There are two key findings in the empirical analysis of the PART scores. First, programs run by political appointees perform worse on average than programs run by career managers. Second, among programs administered by appointees, those run by appointees from the campaign or party received the lowest scores. These results are robust to specifications including controls for a host of agency or program specific factors.

Given the controversy of the PART scores themselves, it is worth considering whether observed differences in PART scores are meaningful (Gilmour 2006; Metzenbaum 2009; Moynihan 2006, 2008; Posner and Fantone 2007; Radin 2005; US Government Accountability Office 2004, 2005, 2008). That is, do the differences in PART scores measure real differences in program performance, and, thus, managerial performance? If the PART scores are flawed by factors such as uneven application, problems in rating, uneven

resources for assessment, then the scores arguably should be noisy, uncorrelated with much of anything of theoretical interest. Random problems emerging in PART scores should confound any analysis of the scores, resulting in imprecise and inconsistent estimates across models. Yet, predictable patterns emerge in the scores across program types and agencies.

What is the source of these regularities? Is it systematic bias working its way into the scores in ways that are confounded with key variables of interest? This is possible but a few aspects of this analysis give us confidence that differences in PART scores reflect real differences in performance. First, the analysis accounted for a host of agency and program-specific factors. For example, the analysis accounted for the most visible form of ideological bias in the scores. One of the most persistent critiques of PART is that the scores are politicized. However, if ideology is properly accounted for in model estimation, observed differences in PART scores can be meaningful. Second, we examined only those PART scores that career executives report as being the most valid and the results remain. Overall, career executives were not sanguine about the validity of the PART scores. This makes drawing inferences about real performance difficult, but career executives thought PART scores were more credible in some agencies than others. Presumably, if career professionals believe the PART scores reflect real differences among programs in their agencies, they contain meaningful performance information and the systematic differences between appointees from the campaign and other executives derive from differences in management performance.

A second possible concern is that appointees from the campaign or party were less concerned about PART scores than other appointees (or career executives) and this explains the observed performance gap. Although possible, this is unlikely since PART was a high priority for the administration and presumably particularly so for the appointees most loyal to the president such as those from the campaign. The analyses above also evaluate the PART scores judged best by career executives and thereby eliminate PART scores that are unreliable due to a lack of effort. In addition, if the Bush Administration sought to manipulate PART scores for its own benefit, one would expect them to inflate the scores of programs run by appointees, particularly those from the campaign, rather than programs run by careerists. If the Bush Administration manipulated PART scores in this way, then the differences we see in PART scores actually *underestimate* the true differences in program performance.

A final concern is that appointees most concerned with change management would earn the lowest PART scores. The disruption caused by trying to change a program's direction would result in lower PART scores. This is unlikely, however, since effective means for implementing change are measured in the first three portions of the PART worksheet. For example, the PART evaluates the clarity of program purpose and mission, the use of long term goals connected to the program mission, strategic planning, and the adequacy of performance measures. Managers bringing change need to clearly articulate a vision, connect that vision to identifiable metrics, and develop a plan for reaching targets. Managers effectively bringing change should arguably earn *higher* PART scores on at least three of the four portions of the PART worksheet (program purpose and design, strategic planning, and strategic management). When models were estimated on each section independently, the results confirm that programs administered by appointees from the campaign or party earned significantly lower scores.

CONCLUSION

One of the charges raised by critics of Obama Administration appointments is that these appointments will hinder US policymaking and management. (see, e.g. “Political patronage alive and well at the USTR” *Entrepreneur* (2009)) By contrast, some scholars argue that the political skills that can come from experience working for the campaign or party are necessary for any agency head (Maranto 2008). Appointed executives do *political* work such as negotiating with state and local officials and dealing with the media. Although political skills are undoubtedly important, the evidence here suggests that appointees given jobs partly due to campaign or party work perform worse on average than other appointees and career managers. If persons are given jobs for reasons other than their ability to manage a program or agency well, this decreases the chances they will succeed in that task. The political skills and experience gained from work on the campaign or for the party do not translate into effective governance or management. Presidents must sometimes trade-off their desire for accountability with their concern for effective management (Edwards 2001; Hecl 1977; Mackenzie 1981).

What then to make of the common sense belief that political skills are necessary for agency management? One conclusion that can be drawn is that political skills must be accompanied by proper experience and subject area expertise. Persons rewarded with a job from the campaign or party, however, often do not have accompanying work experience or subject area expertise. The effect of patronage factors on management performance might also be smaller in the context of fewer appointees. The United States has significantly more political appointments than other developed democracies by a significant amount. The White House is involved in the selection of 3,000–4,000 persons to policy or confidential positions throughout the government (Lewis 2008; Patterson and Pfiffner 2001). Some existing research suggests that agencies perform best when there is an appropriate mix of appointees and career professionals (see, e.g., Dunn 1997; Golden 2000; Krause, Lewis, and Douglas 2006). Working in teams, appointees, and careerists can take advantage of the different skills they bring to management and work productively together. Reducing appointees might help agencies find the right balance between appointees and careerists that will improve performance.

Given the effect of patronage appointees on performance, it is worth asking why presidents appoint such persons to administration jobs. The obvious answer is that there is political benefit to doing so. Beyond that, however, presidents and their personnel officials try and place people into jobs for which they are qualified and where they can do the least damage to the president’s agenda (Lewis 2009). Indeed, the data suggest that very few of the persons from the campaign or party were placed into positions where they were directly managing federal programs. Those few that were placed into such jobs, however, did perform worse than other appointees. The persistence of large numbers of positions, particularly at the program management level, increases the chances that someone with few qualifications will be named to a key management position. Although presidents prefer to place such persons in positions where they can do the least damage to their public agenda, this is not always possible as the example of Michael Brown at FEMA illustrates.

APPENDIX 1

Federal Program PART Score by Appointees from Campaign

	(1)	(2)	(3)	(4)	(5)	(6)
<i>Appointment Authority</i>						
Appointee (0,1)	-5.72** (2.41)	-7.82** (2.46)	-0.36** (0.17)	—	-4.38 (6.30)	-11.50** (4.10)
Appointee from campaign or national party (0,1)	-5.00 (3.17)	-4.22 (2.63)	-0.57** (0.12)	-6.84 (14.77)	-4.97* (2.82)	-9.58** (4.66)
<i>Program Characteristics</i>						
Ln(Program Budget)	0.59 (0.53)	0.88** (0.40)	0.05** (0.02)	0.80* (0.45)	2.52** (0.72)	0.76 (0.74)
Block/Formula Grant (0,1)	-1.25 (4.62)	-4.73 (3.54)	-0.31 (0.21)	-1.06 (4.97)	—	—
Capital Assets and Service Acquisition (0,1)	3.70 (5.68)	2.98 (4.09)	-0.01 (0.26)	7.96 (5.83)	—	—
Competitive Grant (0,1)	-3.17 (4.52)	-2.22 (3.47)	-0.19 (0.20)	-2.39 (6.30)	—	—
Credit (0,1)	-0.38 (5.63)	2.53 (4.15)	-0.07 (0.25)	3.35 (4.56)	—	—
Direct Federal (0,1)	1.70 (3.89)	3.32 (3.01)	0.17 (0.18)	5.87* (3.49)	—	—
Research and Development (0,1)	5.46 (5.54)	5.53 (3.75)	0.29 (0.23)	5.07 (4.62)	—	—
<i>Bureau Characteristics</i>						
Fixed Term (0,1)	6.55 (4.72)	-2.06 (5.77)	0.03 (0.32)	-0.83 (5.76)	—	—
Commission (0,1)	10.75* (5.79)	5.11 (4.31)	0.53 (0.27)	2.34 (6.02)	—	—
# Programs Evaluated	-2.96** (0.75)	0.31 (0.24)	0.04** (0.01)	0.33 (0.38)	0.35 (0.31)	-2.24** (0.81)
Agency Ideology (Liberal-Conservative)	5.77** (1.22)	1.68 (1.53)	0.24** (0.06)	2.57** (0.91)	7.12 (7.27)	6.75** (3.20)
Constant	61.37** (5.19)	59.61** (4.89)	—	58.83** (4.71)	56.34** (11.49)	73.53** (4.52)
N	331	591	591	505	155	190
F(19, 311; 27, 563; 13,194)	9.12**	8.08**	—	78.83**	8.30**	5.39**
R ²	0.32	0.27	—	0.15	0.11	0.16

Note: Unstandardized regression coefficients and robust SEs reported. Omitted category is a regulatory program run by a career member of the SES. Year assessed indicator estimates omitted. All models estimated with only most reliable PART scores. Model 1 excludes cases with no bureau listed on PART worksheet. Model 2 estimated with department fixed effects (fixed-effect estimates omitted). Model 3 is an ordered probit model of program categorical grades (cut point estimates omitted). Model 4 includes two-stage least-squares estimates on a model estimated only on appointees. We could not reject null that work for the campaign or party is exogenous ($p < .83$). Model 5 includes only liberal agencies and Model 6 includes only conservative agencies. Models 5 and 6 have sparse specifications because of the limited number of cases; estimating the full models would leave too few cases in cells including program type, year, and agency structure. Coefficient estimates in models with full specifications look similar for Model 5 but in Model 6 the coefficients are smaller and insignificant (Appointee [B(SE): -5.76 (4.09); Appointee from campaign or national party: -5.24 (4.14)]).

**Significant at the $p < .05$ level; *significant at the $p < .10$ level in two-tailed test.

APPENDIX 2
Instrumental Variables Details

Finding exogenous regressors that predict whether a manager of a specific program is an appointee or worked on the campaign but has no influence on PART scores is very difficult. To estimate these models, we examine only the most reliable PART scores. These scores exhibit no statistically distinguishable time patterns except that the scores are noticeably lower in 2002 (the first year of the PART) than the other years. We exclude programs assessed in 2002 from the models. The time component of the scores is important since we rely on the natural dynamics of appointment politics as the exogenous regressors. Specifically, we use indicators indicating that a “person” was hired right after an election (0,1; i.e., 2001 or 2005), year of the president’s term in which the “program” was assessed (1–4), and whether the program was assessed during the president’s second term (0,1) to predict the probability that a program was administered by an appointee (rather than, e.g., a careerist in an acting capacity). In the first case, the chances that a program was managed by an appointee are predicted by the date of hire, but the date of hire should not directly influence PART scores. In the second case, the chances that a program is run by an appointee are higher or lower depending upon when in the president’s term the program is assessed but when in a term a program is assessed should not influence the PART score. Finally, during second terms, there are more potential patronage appointees to draw from since there have been two campaigns. In the models estimated, we could not reject the null

APPENDIX 3

Federal Program PART Score by Manager Background (IV Estimates)

	Most Effective Part Scores Only
<i>Manager Background Characteristics</i>	
Bureau Experience (0,1)	-2.60 (1.96)
Months Serving as Bureau Chief (0–182)	0.74* (0.42)
Previous Public Management Experience (0,1)	-6.11 (4.13)
Masters (0,1)	-3.22* (1.76)
Doctorate (0,1)	5.45** (2.46)
Worked in Another Department (0,1)	1.20 (1.76)
Private Management Experience (0,1)	-5.26** (2.00)
Worked in White House (0,1)	-7.87** (3.02)
Worked in Congress (0,1)	-3.15 (2.44)
<i>Bureau Characteristics</i>	
Fixed Term (0,1)	-1.15 (4.37)
Commission (0,1)	2.58 (5.16)
# Programs Evaluated	0.21 (0.22)
<i>Political and Program Characteristics</i>	
Agency Ideology (Liberal-Conservative)	3.47** (0.89)
Ln(Program Budget)	1.00** (0.41)
Constant	69.22** (4.89)
N(Observations, Managers)	500
F(25, 282; 25, 49)	142.39**
R ²	0.21

Note: Unstandardized regression coefficients and robust SEs reported. Program type and year assessed indicator estimates omitted. We could not reject null that tenure is exogenous ($p < .14$).

**Significant at the .05 level, *Significant at the .10 level in two-tailed test.

that the coefficients were the same using standard methods. Specifically, we estimated the models with the `ivregress` command in STATA 10.0, followed by the `estat endogenous` command.

REFERENCES

- Aberbach, Joel D. and Bert A. Rockman 2000. *In the web of politics: Three decades of the U.S. Federal Executive*. Washington, DC: Brookings.
- Bearfield, Domonic A. 2009. What is patronage? A critical reexamination. *Public Administration Review* 69:64–76.
- Bertelli, Anthony, Joshua D. Clinton, Christian Grose, David E. Lewis, and David C. Nixon. 2008. *The Ideology of bureaucrats, presidents, and legislators*. Paper presented at the Annual Meeting of the American Political Science Association. Boston, MA, April 28–31.
- Boylan, Richard T. 2004. Salaries, turnover, and performance in the federal criminal justice system. *Journal of Law and Economics* XLVII:75–92.
- Brewer, Gene A. and Sally Coleman Selden. 2000. Why elephants gallop: Assessing and predicting organizational performance in federal agencies. *Journal of Public Administration Research and Theory* 10:685–711.
- Chun, Young Han, and Hal G. Rainey. 2005. Goal ambiguity in U.S. Federal Agencies. *Journal of Public Administration Research and Theory* 15:1–30.
- Choi, Sungjoo, and Hal G. Rainey. 2010. Managing diversity in U.S. Federal Agencies: Effects of diversity and diversity management on employee perceptions of organizational performance. *Public Administration Review* 70:109–21.
- Clinton, Joshua D. and David E. Lewis. 2008. Expert opinion, agency characteristics, and agency preferences. *Political Analysis* 16:3–20.
- Collins, Melissa, Carra Glatt, David E. Lewis, and Michael Shapiro. 2006. *Politicization and federal management performance: Evidence from surveys of federal employees*. Paper presented at the Annual Meeting of the Association for Public Policy Analysis and Management. Madison, WI, November, 2–4
- Dunn, Delmer D. 1997. *Politics and administration at the top: Lessons from down under*. Pittsburgh, PA: Univ. of Pittsburgh Press.
- Edwards, George C. III. 2001. Why not the best? The loyalty-competence trade-off in presidential appointments. *Brookings Review* 19:12–6.
- Eggen, Dan, and Spencer S. Hsu. 2005. Immigration nominee's credentials questioned. *Washington Post*, September 20, 2005 (online edition, accessed August 7, 2009).
- Gailmard, Sean, and John W. Patty. 2007. Slackers and zealots: Civil service, policy discretion, and bureaucratic expertise. *American Journal of Political Science* 51:873–89.
- Gilmour, John B. 2006. *Implementing OMB's Program Assessment Rating Tool (PART): Meeting the challenges of integrating budget and performance*. Washington, DC: IBM Center for the Business of Government. <http://www.businessofgovernment.org/pdfs/GilmourReport.pdf> (accessed April 13, 2011).
- Gilmour, John, and David E. Lewis. 2006. Political appointees and the competence of Federal Program Management. *American Politics Research* 34:22–50.
- Golden, Marissa Martino. 2000. *What motivates bureaucrats? Politics and administration during the Reagan years*. New York, NY: Columbia Univ. Press.
- Hecl, Hugh. 1977. *A government of strangers: Executive politics in Washington*. Washington, DC: Brookings.
- Jung, Chansu, and Hal G. Rainey. 2008. *Developing the Concept of Program Goal Ambiguity and Exploring Federal Program Performance*. *Academy of Management Proceedings*, 1–6.
- . 2009. Program types, goal ambiguity, and Performance in U.S. Federal Programs. Paper presented at the 2009 Annual Meeting of the Public Management Research Association. Columbus, OH, October 1–3.

- Klinenberg, Eric, and Thomas Frank. 2005. Looting homeland security. *Rolling Stone*, December 15, 2005. http://www.rollingstone.com/politics/story/8952492/looting_homeland_security/.
- Krause, George A. and James W. Douglas. 2005. Institutional design versus reputational effects on bureaucratic performance: Evidence from U.S. Government macroeconomic and fiscal projections. *Journal of Public Administration Research and Theory* 15:281–306.
- . 2006. Does agency competition improve the quality of policy analysis? Evidence from OMB and CBO current year fiscal projections. *Journal of Policy Analysis and Management* 25:53–74.
- Krause, George A. David E. Lewis, and James W. Douglas. 2006. Political appointments, civil service systems, and bureaucratic competence: Organizational balancing and Gubernatorial revenue forecasts in the American States. *American Journal of Political Science* 50:770–87.
- Lewis, David E. 2008. *The politics of presidential appointments: Political control and bureaucratic performance*. Princeton, NJ: Princeton Univ. Press.
- . 2009. Revisiting the administrative presidency: Policy, patronage, and agency competence. *Presidential Studies Quarterly* 39:60–73.
- Light, Paul. 1995. *Thickening government: Federal hierarchy and the diffusion of accountability*. Washington, DC: Brookings.
- Mackenzie, G. Calvin. 1981. *The politics of presidential appointments*. New York, NY: Free Press.
- Maranto, Robert. 2005. *Beyond a government of strangers*. Lanham, MD: Lexington Books.
- . 2008. Career-political relationships: Going beyond a government of strangers. In *The trusted leader*. ed. Terry, Newell, Grant Reeher, Peter Ronayne. Washington, DC: CQ Press.
- Metzenbaum, Shelley. 2009. *Performance management recommendation for the new administration*. Washington, DC: IBM Center for The Business of Government. http://www.businessofgovernment.org/publications/grant_reports/details/index.asp?GID=330 (accessed April 13, 2011).
- Moe, Terry M. 1985. The politicized presidency. In *The new direction in American politics*, ed. John E. Chubb and Paul E. Peterson, 235–71. Washington, DC: Brookings.
- Moynihan, Donald P. 2006. What do we talk about when we talk about performance? Dialogue theory and performance budgeting. *Journal of Public Administration Research and Theory* 16:151–68.
- . 2008. *The dynamics of performance management: Constructing information and reform*. Washington, DC: Georgetown Univ. Press.
- . 2009. A Heckuva job: How management failures doomed the Bush administration. *Public Management Review* 11:121–33.
- Nakamura, David. 2009 D.C. Technology Office Scandal Revives Questions of Cronyism. *Washington Post*. March 14, 2009, B01.
- O’Connell, Anne Joseph. 2009. Vacant offices: Delays in staffing top agency positions. *Southern California Law Review* 82:913–1000.
- Oh, Paul S. and David E. Lewis. 2008. Management and leadership performance in the Defense Department: Evidence from Surveys of Federal Employees. *Armed Forces & Society* 34:639–61.
- Patterson, Bradley H. Jr., and James P. Pfiffner. 2001. The White House Office of Presidential Personnel. *Presidential Studies Quarterly* 31:415–38.
- Pfiffner, James P. 1996. *The strategic presidency: Hitting the ground running*. Lawrence, KS: Univ. Press of Kansas.
- Political patronage is alive and well at the USTR. *Entrepreneur*, May 18, 2009. <http://www.entrepreneur.com/tradejournals/article/200779825.html> (accessed July 20, 2010).
- Posner, Paul L. and Denise M. Fantone. 2007. Assessing federal program performance: Observations on the U.S. Office of Management and Budget’s Program Assessment Rating Tool and its use in the budget process. *Public Performance & Management Review* 30:351–68.
- Radin, B. A. 2005. *Testimony to the Senate Homeland security and Government Affairs Subcommittee on Federal Financial Management, Government Information, and International Security June 14, 2005*. http://hsgac.senate.gov/public/_files/Radin_testimony.pdf (accessed December 31, 2009).
- Smith, R. Jeffrey and Susan Schmidt. 2005. Bush official arrested in corruption probe. *Washington Post*, September 20, 2005. (online edition, accessed August 7, 2009).

- Tumulty, Karen, Mark Thompson, and Mike Allen. 2005. How many more Mike Browns are there out there? *Time*, September 25, 2005. <http://www.time.com/time/magazine/article/0,9171,1109345,00.html> (on-line edition, accessed July 25, 2009).
- U.S. Government Accountability Office. 2004. *Performance budgeting: Observations on the use of OMB's Program Assessment Rating Tool for the Fiscal Year 2004 Budget*. GAO-04-174. Washington, DC: Government Accountability Office.
- . 2005. *Performance budgeting: PART focuses attention on program performance, but more can be done to engage congress*. GAO-O6-28. Washington, DC: Government Accountability Office.
- . 2008. *Government performance: Lessons learned for the next administration on using performance information to improve results, Washington, DC*. Washington, DC: Government Accountability Office. GAO-08-1026T. <http://www.gao.gov/new.items/d081026t.pdf> (accessed December 31, 2009).
- Weisman, Jonathan, and Yuka Hayashi. 2009. Donors find a home in Obama's Ambassador Corps. *Wall Street Journal*. July 6, 2009. <http://online.wsj.com/article/SB124658149328689699.html> (online edition, accessed August 5, 2009).
- Weko, Thomas J. 1995. *The politicizing presidency: The White House Personnel Office*. Lawrence, KS: Univ. Press of Kansas.

Copyright of Journal of Public Administration Research & Theory is the property of Oxford University Press / UK and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.