

# **Administrative Reform, Political Ideology, and Bureaucratic Effort: The Case of Performance Management in the Bush Administration**

Stéphane Lavertu  
Assistant Professor  
John Glenn School of Public Affairs  
The Ohio State University  
[lavertu.1@osu.edu](mailto:lavertu.1@osu.edu)

David E. Lewis  
Professor  
Department of Political Science  
Vanderbilt University  
[david.lewis@vanderbilt.edu](mailto:david.lewis@vanderbilt.edu)

Donald Moynihan  
Professor  
Robert M. La Follette School of Public Affairs  
University of Wisconsin – Madison  
[dmoynihan@lafollette.wisc.edu](mailto:dmoynihan@lafollette.wisc.edu)

**Draft: 10/15/2012**

## **Abstract**

There is a long-standing concern about how differences in political beliefs between political officials and bureaucrats affect the administration of policy. This paper examines how ideological divergence between political principals and public agencies relates to the time and effort government agencies expend in implementing ostensibly non-partisan administrative reforms. Our empirical focus is on the George W. Bush administration's Program Assessment Rating Tool (PART) evaluation process, which required federal agencies to participate in evaluating their own programs. We find that respondents who worked in liberal agencies reported greater agency-wide effort than those who worked in conservative agencies, regardless of respondents' own political preferences and regardless of the perceived budgetary consequences associated with poor PART reviews. Further study reveals that liberal agencies' greater effort is partly attributable to their smaller size, the greater number of PART reviews they were required to conduct, and their administration of regulatory programs. Overall, this study indicates that the ostensibly non-partisan PART review process imposed a greater administrative burden on liberal agencies, and the results obtain using multiple and qualitatively different measures of ideology and administrative burden from multiple data sources.

Divergence between the ideological preferences of political officials and public administrators has long been of concern in a variety of governmental settings (Aberbach, Putnam, and Rockman 1981; Lee and Raadschelders 2008). Such gaps have motivated concerns about the excessive politicization of public programs or public agencies' unresponsiveness to the democratic process, depending on where one's ideological sympathies lie. In contrast to existing research, this study examines how ideological divergence between political principals and public agencies (as captured by agencies' programs, employees, and political constituencies) affects the time and effort that agencies expend in their implementation of administrative reforms. One might expect that ideological divergence leads to lower agency effort—that is, agency “shirking”—if it is difficult for political principals to observe the implementation of administrative reforms. This is the classic principal-agent perspective. We hypothesize, however, that ideological divergence results in greater agency-wide effort when political principals are able to monitor implementation, as ideologically divergent agencies may need to do more to satisfy skeptical or demanding political overseers.

The specific case we examine is the implementation of performance management reforms in the United States under the George W. Bush administration. President Bush, a conservative president, required that all programs undergo performance evaluations using the Program Assessment Rating Tool (PART). Whereas existing studies indicate that political ideology was a factor in how the OMB implemented the administrative reform—specifically, that liberal agencies and programs fared worse in the PART review process (Gallo and Lewis 2012; Gilmour and Lewis 2006a 2006b)—much less is known about how the politics of PART affected the behavior of agency personnel. (For an exception, see Lavertu and Moynihan 2012.) PART also is unique in that it involved significant OMB monitoring of agency activities, which enables us to

test our hypothesis that ideological divergence results in greater agency-wide implementation effort when the information asymmetry is minimal between the political principal and agency.

Using multiple data sources, including original survey data of agency employees and objective measures of agency-level characteristics, we examine the relationship between ideological divergence (specifically, agency liberalism, which is in contrast to President Bush's conservatism) and the time and effort that agencies expended in fulfilling their responsibilities in the PART evaluation process. The empirical analysis reveals that respondents who worked in liberal agencies reported greater agency-wide effort than those who worked in conservative agencies, even when controlling for the respondents' own political preferences and the perceived budgetary consequences of poor PART scores. Further study of this relationship reveals that liberal agencies' greater reported effort is partly attributable to their smaller size, the greater number of PART reviews they were required to conduct, and their tendency to house regulatory programs, which are associated with more effort-intensive PART reviews. These results indicate that PART reviews imposed a greater burden on liberal agencies, and they obtain using multiple and qualitatively different measures of ideology and administrative burden from multiple data sources.

Even if PART was intended to be a non-partisan "good government" effort, ideological differences between the presidential administration and agencies corresponded to agencies expending greater time and effort toward its implementation. Put simply, a reform promoted by a conservative president imposed a significantly greater burden on liberal agencies. Intentional or not, this disparate impact along ideological lines has political repercussions that are likely to undermine efforts to implement administrative reforms. Indeed, the more liberal President Obama discontinued PART immediately after assuming office. In the following pages, we

describe the politics of effort allocation in administrative reform, discuss PART and its politics, outline some hypotheses, describe and analyze our data, and, finally, discuss our results.

### **The Politics of Effort Allocation in Administrative Reform**

This study is focused on the politics of agency effort allocation during the implementation of administrative reforms. Observers in a wide variety of settings have noted that political officials often are concerned about bureaucrats who do not share their beliefs and, consequently, that political officials seek to control the behavior of such bureaucrats. Strategies for exercising such control have included the appointment of ideological allies in key leadership positions and otherwise promoting partisanship among public servants (Aucoin 2012; Dierickx 2003; Kickert 2011; Lewis 2008; Rouban 2008). Some of the literature on ideological divergence has documented extreme and overt cases of political principals changing the policy goals of an agency, engendering bureaucratic resistance (Durant 1987; Golden 2000). However, there has been less attention to more common administrative reforms whose explicit purpose is to improve government effectiveness (e.g., fiscal management, IT improvements, etc.), rather than alter policy. One possibility is that such reforms have truly non-political impacts—that is, agency employees experience the reform similarly regardless of the political proclivities of an agency's programs, employees, and constituency. As we discuss below, however, this is unlikely to be the case.

Scholars who take a principal-agent approach to the study of public bureaucracy are centrally concerned with the amount of effort bureaucrats expend and the end towards which effort is expended (see e.g., Brehm and Gates 1999; Gailmard and Patty 2007; Ting 2003). For example, bureaucrats may be imbued with public service motivation that makes them willing to exert exceptional levels of effort in pursuit of organizational goals in which they believe. But

public administration and political science scholars also focus on whether that effort is directed at goals that elected political principals want them to pursue and whether less effort is exerted in pursuit of goals in which bureaucrats do not believe. Thus, one might say that the allocation of effort across administrative goals and tasks is a primary concern.

Public agencies and their employees have resource constraints. Time spent on one task, for example, is time that cannot be spent on another. Thus, administrative reforms that require that agencies contribute resources—such as employee time and effort—may divert resources from other goals that agencies and their employees prioritize. Agency employees might resist contributing time and effort to a program they believe is inconsistent with their agencies' priorities. If agencies and employees possess sufficient discretion—for example, if effort itself is difficult or costly to observe, or if punishment is unlikely if they fail to exert sufficient effort—then they may not exert the level of effort that political reformers desire, thereby undermining administrative reform. For example, the US Department of Defense expended limited effort in the implementation of President Clinton's Reinventing Government reforms partly because of distrust among members of this conservative department toward their liberal commander-in-chief (Durant 2008). This is the classic “shirking” scenario, in which administrative agents are able to resist exerting effort to pursue goals with which they disagree—either because of a principal's inability to monitor agent behavior or the principal's inability or unwillingness to correct the behavior.

This classic principal-agent perspective does not always apply, however. When a political principal is willing and able to monitor and enforce compliance with her or his policy directive, ideological divergence may motivate an administrative agent to expend more time and effort implementing the reform. There are at least two potential explanations for this. First, a political principal who is skeptical of an agency's mission and programs may be more vigilant and

demanding when it comes to making sure the agency implements his or her administrative reforms. Indeed, forcing program administrators to reallocate their time to pursuing administrative reforms itself could help a political principal undermine programs with which he or she disagrees. Second, administrators conscious of a political principal's skepticism vis-à-vis his or her agency and programs might expend relatively high levels of time and effort in an effort to protect an agency's programs, perhaps in the hope of appearing neutral and competent. It is important to note, then, that ideological divergence may lead to higher levels of agency effort even if the political principal seeks to apply administrative reforms in a politically neutral way, simply because an administrative agent may perceive a risk to his or her agency's programs.

Research on governments outside the United States has examined the issue of political-administrative ideological divergence, but not in the manner in which we do below. First, existing research, following Aberbach et al. (1981), tends to focus on ideological divergence between individuals—that is, divergence in the personal ideologies of political principals and senior bureaucrats. Our focus is on differences in the ideology of political principals and administrative agencies as a whole, as defined by their programs, political constituencies, and employees. Second, existing work has not examined the impact of ideological divergence on the implementation of administrative reforms, which often are less explicitly political. Moreover, though our claims may be applicable only to the performance management reform we examine, the ubiquity of such reforms means that the implications may have broad applicability (see, e.g., Askim et al, 2008; Cristofoli et al 2011; Van de Walle and Van Dooren 2008).

In summary, our focus is on the examination of the impact of political-administrative ideological divergence on the implementation of ostensibly non-partisan administrative reforms. Additionally, because of the nature of the administrative reform we examine (which we describe below) we suggest that ideological divergence between political officials and administrative

agencies may not lead to administrative “shirking”—that divergence might lead agencies to expend greater time and effort during implementation.

### **The Politics of PART**

PART is perhaps the best studied United States administrative reform effort since the Reinventing Government initiatives of the 1990s. A variety of studies have described its characteristics and impacts (e.g., Dull 2006; Frederickson and Frederickson 2006; Frisco and Stalebrink 2008; Gallo and Lewis 2012; Gilmour and Lewis 2006a,b; Heinrich 2012; Joyce 2011; Moynihan 2008; Moynihan and Lavertu 2012; Stalebrink and Frisco 2011; White 2012). PART was a questionnaire that the OMB used to evaluate federal programs. The questionnaire was broken into sections that reflected different evaluative dimensions: program purpose and design, strategic planning, program management, and program results/accountability. The OMB graded programs on an ineffective-to-effective summary scale based on scores across these dimensions. The OMB conducted PART reviews in waves between 2003 and 2008 until it had evaluated nearly all federal programs. This was an enormous project, consuming a great deal of the OMB’s analytical capacity during the Bush years (Moynihan 2008; Redburn and Newcomer 2008), perhaps at the expense of traditional and more fruitful forms of budgetary analysis (White 2012).

The OMB under Bush took significant steps to promote PART as a non-partisan administrative tool (Dull 2006). But there is a substantial literature that examines how politics came to bear in the implementation of PART. Gallo and Lewis’s (2012) study reveals, among other things, that programs administered by liberal-leaning agencies received lower PART scores than those administered by conservative-leaning agencies.<sup>1</sup> Gilmour and Lewis (2006a) found that PART scores related to the “purpose” of programs, as opposed to their “results,” had a

greater impact on OMB budgetary decisions, suggesting that the reform sought to alter what goals agencies pursued in addition to how well they pursued them. Moreover, Gilmour and Lewis (2006b) found that only PART scores for Democratic programs informed presidential budget proposals—in other words, that the budgets of traditionally conservative programs were relatively insulated from negative PART scores—although PART scores had no such impact on congressional appropriations (Heinrich 2012; Frisco and Stalebrink 2008). Finally, it appears that programs that were redistributive in nature received systematically lower scores (Greitens and Joaquin 2010).

The above studies make clear that ideological divergence between President Bush and agencies and their programs—essentially, the divergence in their policy and programmatic preferences—was related to PART implementation. In other words, politics played a role in how principals in the executive branch implemented the administrative reform. (Indeed, this is consistent with what Bertelli and John (2010) found in their study of performance measurement in local English governments.) But little is known about how the politics of PART affected agency behavior. This study of agency effort allocation seeks to address this gap.

### **PART and Agency Effort Allocation**

OMB budget examiners conducted PART evaluations, but the evaluations also required considerable effort from agency personnel who were charged with completing the PART questionnaire. Additionally, in designing PART, OMB officials deliberately placed the burden of proof on agencies (Moynihan 2008, 140). Thus, any claim an agency made about the merits of a program had to be supported by documented evidence. While poor performing programs were unlikely to be given high scores, strong programs could be scored poorly if agencies did not invest enough resources into the process. As Gilmour (2006, 17) notes, “It takes hard, careful



work for a program to get a score that reflects its true merit.” Gilmour also identified a series of strategies that agencies used to improve their scores, such as relying on senior staff involvement, providing ample documentation, hiring professional help, and measuring as much as possible—all of which required the investment of significant agency resources. Additionally, he identifies programs that improved their original PART scores by investing more effort in their response to PART reviews, rather than in actually changing their programs. Yet another PART-generated burden came at the conclusion of the assessment process, when the OMB directed agencies to implement a set of management recommendations that it would later assess.

The nature of the process means that the principal (the OMB) could compel participation in the process, could observe whether the agent was providing effort in response, and could reward such effort with higher grades. Thus, it is clear there was the potential for agencies to expend significant time and effort on PART evaluations. Considering both our theoretical framework of effort allocation and what we know about PART and its politics, we theorize that the amount of effort agencies expended likely depended on 1) ideological divergence in the executive branch, 2) the potential budgetary consequences associated with PART evaluations and 3) the administrative burden that PART evaluations imposed on agencies. We consider these in turn.

### *Ideological Divergence*

This study’s primary focus is on how agency-wide priorities and goals relate to agency-wide time and effort expended on the PART review process. We treat ideological divergence as capturing divergence in priorities and goals between the OMB and federal agencies.

PART entailed significant top-down oversight and, as existing research indicates, there was likely an ideological bias in PART reviews that worked against liberal agencies and

programs. A lack of effort could create risks for these agencies or programs. For example, a liberal employee managing a liberal program in a liberal agency might have devoted significant time and effort to cooperate in PART evaluations in order to compensate for the administration's bias and, thus, protect his or her program's reputation. Indeed, if the OMB demanded greater evidence from and provided more direction to liberal agencies, employees from those agencies might have had little choice but to expend greater effort than their counterparts in conservative agencies. Thus, we offer the following hypothesis.

*Hypothesis 1:* Agencies whose personnel and programs link them to politically liberal constituencies were more inclined to expend time and effort on PART reviews than agencies whose personnel and programs link them to politically conservative constituencies, all else held equal.

### *Budgetary Consequences*

Agencies' concerns about their programs should lead them to seek to protect them. If agency personnel believe that there are programmatic consequences from poor PART evaluations, then, as we discuss above, they have incentives to expend time and effort to substantiate a program's worth—provided that such concerns trump their competing priorities and that they possess the capacity to invest the required time and effort. Budgetary officials and members of Congress freely admit that there are a number of factors that influence the president's budget and congressional appropriations and that performance is only one of them. For example, some programs' budgets are insensitive to performance information because they are considered essential even if working poorly. And some programs may be performing poorly *because* of small budgets. Additionally, as we discuss above, research has found that presidential budget requests for programs administered by liberal agencies varied with PART scores (though

appropriations were not related to those scores) whereas the budgets of programs administered by conservative agencies were relatively immune from PART scores. Thus, we offer the following hypothesis:

*Hypothesis 2:* Agencies expended more time and effort in conducting and responding to PART reviews if their personnel believed that these reviews had an impact on program budgets, all else held equal.

### *Administrative Burden*

We expect effort to increase as agencies find reform implementation more burdensome. Reform implementation requires the provision of time, effort and expertise, and it may be that an agency does not possess sufficient capacity to administer the reform, perhaps because reallocating such capacity from other programs is prohibitively difficult. It might also be the case that contributing to PART reviews required a more substantial reallocation of resources—in other words, posed a more significant administrative burden—in some agencies than it did in others. Some agencies, for example, had over one hundred programs evaluated through PART and others had only a handful. This may have political implications as poor PART scores (which are more likely to be associated with liberal programs), for example, can result in the requirement that agencies take on significant remedial actions. Additionally, administrative burden might vary with the types of programs agencies administer. For example, some programs better lend themselves to performance measurement because it is easier to observe programmatic outcomes. Thus, we offer the following hypothesis.

*Hypothesis 3:* Agencies on which PART imposed more demands relative to their administrative capacities expended greater time and effort on PART reviews, all else held equal.

## EMPIRICAL ANALYSIS

### Overview of Data and Methods

The data analysis proceeds in two steps. First, we analyze data from a survey of federal agency executives in order to identify relationships between ideological divergence (at the agency and personnel levels), budgetary incentives, and administrative burden, and the reported time and effort that agencies spent on PART reviews. Second, we consider differences in the administrative burden that PART imposed on agencies across the ideological spectrum. To do so, we examine descriptive statistics by agency ideology. In the sections that follow we describe our data sources, briefly review our statistical methods, and present and discuss the results.

### Survey and Agency-Level Data

We rely on data from the 2007-8 *Survey on the Future of Government Service*. The principal investigators for this survey of federal executives fielded the survey in 2007-2008 and targeted 7,448 federal administrators and program managers (both career and appointed). The response rate to the survey was 33% and robustness checks reveal that the sample of career executives is representative of the population from which it was drawn (see Clinton et al. 2012 for a full discussion).<sup>ii</sup> The survey included a number of questions about the backgrounds, attitudes, and activities of federal executives. Importantly for our purposes, the survey included a number of questions about the PART process, including whether the respondent's agency had a program evaluated in the PART process, how much effort was put into the evaluations, the validity of the PART evaluations, and the impact of PART evaluations on the president's budget and the appropriations process. Among the sample of respondents, 1,000 indicated that their agency had programs "reviewed as part of the Program Assessment Rating Tool (PART) process". The data provided by these respondents provide the basis of the analysis that follows.<sup>iii</sup>

## Variables

The variables we employ are described and summarized in Table 1. We also briefly review them below.

[Insert Table 1 here.]

### *Agency Effort*

The dependent variable *effort* is the focus of this study. Respondents indicated on a five point scale “How much time and effort did your agency put into the Program Assessment Rating Tool (PART) process?” Thus, although the models are estimated using respondent-level data, our analysis is focused on the amount of agency-wide effort expended during PART implementation.

### *Ideological Divergence*

The Bush administration was relatively conservative. Consequently, to capture ideological divergence between the administration and agencies (and their employees) we employ measures of agency liberalism. In other words, each variable captures how liberal (or Democratic) agencies are. First, we use the percentage of agency executives that self-identify as Democrats. Second, we use Clinton and Lewis’s (2008) estimates of agency ideology to create the dichotomous variable *liberal agency*. The Clinton-Lewis measure of agency ideology uses responses from an expert survey about agency ideology to generate numerical estimates of agency ideology. We use this second measure because there are aspects of agency ideology that are distinct from the political views of agency executives (e.g., mission, history, culture).<sup>iv</sup> We have also estimated models using average respondent ideology and average respondent ideal point by agency, but these results confirm what is discussed below, so we relegate them to an appendix (Appendix A).<sup>v</sup>

### *Budgetary incentives*

We also measure the perceived budgetary incentives of expending effort in the PART review process using data from the survey. The measure captures respondent beliefs regarding the extent to which PART scores influenced presidential budget requests. Specifically, the survey asked respondents “To what extent did PART scores influence the president’s budget request for the program(s) evaluated?” Answers range from “Not much at all” (0) to “Tremendous amount” (4). As we discuss above, research indicates that PART scores had a discernible influence on budget requests only for programs administered by liberal agencies, but not on ultimate program appropriations (Heinrich 2012; Frisco and Stalebrink 2008). This is consistent with a number of accounts that members of Congress did not consider PART scores in their decision-making.<sup>vi</sup>

### *Administrative Burden*

The variables above are based on respondent perceptions of agency effort and budgetary consequences. The variables we employ to capture agencies’ administrative burden are based on objective measures of agency size and programmatic budgets (proxies for administrative capacity) and the scope of PART review activities within agencies. We capture agency size using the natural log of agency employees<sup>vii</sup> and the log of the average budget of programs in the respondent’s agency. We capture the administrative burden of PART activities with a count of the number of agency programs that underwent PART review, holding agency size and capacity constant. Additionally, to gain some insight regarding what types of programs are associated with the most agency effort, we employ counts of the number of regulatory and R&D programs reviewed, as the amount of work involved in measuring performance is argued to be higher for those types of programs than others (Radin 2009).<sup>viii</sup>

Finally, the analysis below considers additional measures to examine differences in administrative burden between liberal and conservative agencies. First, we consider the percentage of respondents from each agency who indicated some involvement in PART reviews. Second, we summarize responses to a 2007 GAO survey item that asks federal managers to estimate the resource burden that PART involvement imposed on their programs.<sup>ix</sup>

### **Methods and Results: The Impact of Ideological Divergence**

The results displayed in Table 2 are from models focused on the relationship between agency political ideology and the reported agency-wide time and effort expended on PART reviews. The dependent variable in all models is *effort*, an ordinal variable with five categories. All models estimating *effort* are ordered probit models; all results tables report standard errors clustered by agency; and cut point estimates are omitted in the interest of space.<sup>x</sup>

[Insert Table 2 about here.]

We hypothesized that liberal agencies expended more effort than conservative agencies, perhaps to protect their programs from a biased administration or because the PART process simply demanded more of them (*hypothesis 1*). As we discuss in our motivations for hypotheses 2 and 3, budgetary incentives and administrative burden are factors that could relate to ideological effects in various ways. Therefore, we present the results of models estimated with and without variables capturing these factors—and with and without control variables—in order to get a sense for the robustness of the ideological effects.

The results presented in Table 2 support *hypothesis 1*, which states that liberal agencies committed more effort to PART. There is a positive and statistically significant relationship between agency ideology and *effort*. These results obtain with and without controls for a host of agency and respondent-specific factors, including budgetary incentives, administrative burden,

and respondent ideology. The latter control provides some reassurance that these results are not attributable to bias in respondent perceptions. Thus, it does not appear that agencies whose political constituencies and ideological traditions diverge from those of the presidential administration “shirked” their responsibilities when it came to PART.<sup>xi</sup>

In Figure 1 we graph the estimated impact of the percentage of Democratic executives in an agency and the probability that an agency put a “significant” or “tremendous” amount of effort into the PART process. As the figure illustrates, increasing the percentage of Democrats from one standard deviation below the mean percentage (34%) to one standard deviation above the mean (76%) is associated with an estimated increase of 14 percentage points in the probability that respondents reported that their agencies expended a “significant” or “tremendous” amount of time and effort on PART process. In the second set of models employing the Clinton-Lewis measure, respondents in liberal agencies are 10 percentage points more likely to report that their agency put a “significant” or “tremendous” amount of time and effort into the PART process.

[Insert Figure 1 about here.]

It is worth noting that in one model respondents who self-identify as Democrats are marginally more likely to report that their agencies put a significant amount of effort into the PART process. They were about 4 percentage points more likely to report a “significant” or “tremendous” amount of effort was required based upon estimates from the second set of models. The positive correlation between respondent partisanship and perceptions of effort could reveal partisan bias among respondents.<sup>xii</sup>

Table 2 also allows us to examine whether the perceived budgetary consequences of PART explain the results. Specifically, it includes estimates of respondent perceptions regarding the impact of PART scores on presidential budget requests. Recall that we hypothesized a



positive relationship between respondent concerns about the impact of PART scores on agency budgets and the amount of agency time and effort expended. The results indicate a positive and statistically significant relationship between *budget* and *effort*. Indeed, a respondent that indicates that PART scores influence the president's budget only a "small amount" are 10 percentage points less likely to report a "significant" or "tremendous" amount of effort than a respondent that indicates PART scores influence the president's budget a "significant amount." This result suggests that concerns about low PART scores influencing program budgets motivated agencies to expend effort in the PART review process, whereas agencies insulated from budgetary concerns were freer to allocate less time and fewer resources to the PART process.

There is also evidence that the administrative burden of the PART process influenced the amount of time and effort agencies expended. We find that that agencies with larger programs report expending less effort in the PART process. While we characterize program budget size as a proxy for capacity, other work suggests that larger programs were less vulnerable to the budgetary risks posed by PART scores (Gilmour and Lewis 2006a). It is possible, therefore, that respondents working in larger programs committed less effort because they perceived less budgetary risk, in which case the finding could be interpreted as support for *hypothesis 2* regarding budgetary incentives.

We also find that agencies that had more programs evaluated put more effort into PART. Model estimates indicate that increasing the number of programs evaluated from one standard deviation below the mean (15 programs) to one standard deviation above the mean (85 programs) increases the probability by 12 percentage points that respondents report a "significant" or "tremendous" amount of effort. This raises the question of whether politics

played a role in the definition of what constitutes a program and how decisions were made about how many programs to evaluate.

### **Methods and Results: Supplementary Analysis of Administrative Burden**

Table 3 presents some descriptive statistics that further expand on these results. The first two columns present summary statistics divided between agencies with 50 percent or fewer respondents indicating that they are Democrats and those with over 50 percent indicating that they are Democrats. The next two columns summarize statistics for non-liberal and liberal agencies, as identified by Clinton and Lewis (2008). The table clearly illustrates that liberal agencies are smaller, had more programs undergo PART review, and had more regulatory programs and fewer R&D programs undergo PART review. Additionally, the table reveals that a higher proportion of survey respondents reported PART involvement if they worked in liberal agencies. Finally, the table indicates that in a 2007 GAO survey of federal agency managers, a higher percentage of managers in liberal agencies agreed to a “moderate,” “great,” or “very great” extent that PART “imposed a significant burden on management resources.”

[Insert Table 3 about here.]

The differences in results related to regulatory and R&D programs are worth analyzing further. Liberal agencies are more likely to administer regulatory programs, so one might contend that the count of regulatory programs evaluated is merely a proxy for agency ideology. This is true. But it may also be that demonstrating the value of regulatory programs is more difficult under a performance management framework (Frederickson and Frederickson 2006). Similarly, that the count of PART-reviewed programs is positively associated with agency effort may in part be an indication that omitted program types (block/formula grants, capital assets and

service acquisition, competitive grant, credit, direct federal) are similarly unsuited to performance management (see, e.g., Radin 2006).

In Table 4 we present the results of a set of models that includes counts of the number of regulatory and R&D programs in each agency.<sup>xiii</sup> The results indicate that agencies with large numbers of regulatory programs are estimated to be more likely to put forth greater effort and those with large numbers of R&D programs less effort. That the number of R&D programs reviewed is associated with less agency effort seems to undermine the argument that different types of programs are unsuited to performance management, as R&D program outcomes should be difficult to estimate. However, qualitative studies of the PART process might help to explain the result. Frederickson and Frederickson (2006) note that the R&D agency they studied, the National Institutes for Health, enjoyed greater success in persuading OMB examiners to allow them to use alternative ways to present program quality. Additionally, they found that R&D programs bucked a more general trend of poor PART assessments for services performed by third-party actors. The OMB acknowledged that R&D programs would have an especially difficult time complying with PART assessments and guided budget examiners to exempt them from requirements to demonstrate outcomes (Gilmour 2006, 12). The OMB seemed willing to allow a more procedural assessment of quality for these programs, examining if grant awards followed appropriate peer-reviewed processes, and deferring to scientific expertise (Gilmour 2006, 22). Our findings seem to confirm what previous studies suggest, which is that R&D programs were exempted from those criteria in a way that other types of programs were not. Indeed, Gallo and Lewis (2012) show that R&D programs even received systematically higher PART scores relative to regulatory programs.

Finally, parameter estimates for the control variables indicate that respondents in cabinet departments are significantly less likely to report that their agencies put significant effort into the

PART process. This may reflect that programs in the more politicized cabinet agencies are better protected from potential negative repercussions than agencies that lack a seat at the cabinet table.

## **Discussion**

This study reveals that agency personnel from ideologically liberal agencies—those whose employees are more liberal and that administer traditionally liberal programs—reported that their agencies expended greater effort in the PART review process than those in conservative agencies. It identifies partly why this result obtains: liberal agencies are smaller, were forced to evaluate more PART programs, and were more likely to administer regulatory programs, which are thought to be more difficult to justify using a performance management framework. Thus, whether greater agency-wide effort is due to explicitly political considerations by the Bush administration or simply to the nature of liberal agencies and the programs they administer, it is clear that PART imposed a greater administrative burden on liberal agencies.

It would be interesting to determine whether or not the relatively greater administrative burden placed on liberal agencies was deliberate. Alas, this is not something we can do empirically, so we briefly sketch competing arguments that one might make.

One argument is that the greater burden on liberal agencies is incidental—that liberal agencies simply administer more programs with fewer employees, and that these programs are more likely to be regulatory in nature. The OMB cannot change these basic characteristics of an agency. It cannot significantly influence the number of employees an agency has and it is unlikely to demand more PART assessments simply because an agency has more employees. Someone making this argument might also point to the fact that 98% of all federal programs were evaluated, thereby undercutting claims that the OMB selectively required liberal agencies to participate in PART.

A counter-argument might be that ideological divergence between the Bush administration and liberal agencies drove the OMB to impose a greater administrative burden on liberal programs. Those making such an argument might point to significant empirical evidence that liberal agencies suffered under PART, receiving systematically lower PART scores and facing greater risk to their budgets (Gallo and 2012; Gilmour and Lewis 2006a, b). They might also focus on the large difference in the number of programs reviewed in liberal and conservative agencies. It is true that almost all federal government programs were reviewed, but the OMB had a good deal of discretion in determining what constituted a program. Gilmour (2006) portrays the decision about what constituted a program as something negotiated between the OMB and agency leadership. Liberal programs might have been defined more narrowly, perhaps to permit closer scrutiny.

An illustration of this second argument may have occurred at the (liberal) Department of Education, for which programs were defined so narrowly that it had more programs than the (conservative) Department of Defense, despite having less than one tenth of the budget (Joyce 2011). Functions within the Department of Education with budgets as low as \$1 million were defined as programs. And functions that arguably represent a single program were subdivided. For example, activities undertaken as part of the Individuals with Disabilities Education Act were disaggregated into seven separate programs for PART purposes (Gilmour 2006, 11). This narrow definition of programs appears to have been a deliberate strategy of agency leadership who doubted the quality of much of what the Department did: “the leaders at the Education Department believe the department is burdened with many ill-conceived, poorly designed programs, and see the PART process as a means of shining a light on those deficiencies” (Gilmour 2006, 16).

One of the ironies of the process is that while PART generated greater effort on the part of liberal agencies, there is little evidence that this effort was associated with a stronger embrace of performance management by these managers. Some research has found that managerial involvement in the PART review process spurred the use of performance information in conservative and, to a lesser extent, moderate agencies, but that managerial involvement in PART reviews had no such effect in liberal agencies (Lavertu and Moynihan 2012). This tells us that some aspects of implementation remain impossible to observe, and effort is not always equivalent to the desired outcome of the principal.

## **Conclusion**

On balance, the evidence indicates that liberal agencies expended greater effort in the PART review process than conservative agencies. This is at least partly because PART reviews posed a greater administrative burden on liberal agencies. Even as liberal agencies were working harder under PART, they benefited less from it, as they were subjected to lower PART scores and the threat of budget cuts (Gallo and Lewis 2012; Gilmour and Lewis 2006b). Whether or not the enhanced burden derives from political motivations (e.g., the desire of a conservative administration to apply greater scrutiny to liberal programs) this study provides another example of how ostensibly non-partisan administrative reforms may have political implications.

## References

- Aberbach, Joel, Robert D. Putnam and Bert A. Rockman. 1981. *Bureaucrats and Politicians in Western Democracies*. Cambridge, MA: Harvard University Press.
- Askim, Jostein, Åge Johnsen, and Knut-Andreas Christophersen. 2008. Factors behind organizational learning from benchmarking: Experiences from Norwegian municipal benchmarking networks. *Journal of Public Administration Research and Theory* 18:297–320.
- Aucoin, Peter. 2012. New Political Governance in Westminster Systems: Impartial Public Administration and Management Performance at Risk. *Governance* 25(2): 177–199
- Bertelli, Anthony M. and Peter John. 2010. Government Checking Government: How Performance Measures Expand Distributive Politics. *The Journal of Politics* 72(2): 545–558.
- Brehm, John, and Scott Gates. 1999. *Working, Shirking, and Sabotage: Bureaucratic Responses to a Democratic Public*. Ann Arbor: University of Michigan Press.
- Daniela Cristofoli, Greta Nasi, Alex Turrini and Giovanni Valotti. 2011. Civil Service Reforms in Italy: The Importance of External Endorsement and Administrative Leadership. *Governance* 24(2): 261–83.
- Clinton, Joshua D., Anthony Bertelli, Christian R. Grose, David E. Lewis, and David C. Nixon. 2012. Separated Powers in the United States. *American Journal of Political Science* 56(2): 341–54.
- Clinton, Joshua D. and David E. Lewis. 2008. Expert Opinion, Agency Characteristics, and Agency Preferences *Political Analysis* 16(1): 3–20.
- Dierickx, Guido. 2003. Senior Civil Servants and Bureaucratic Change in Belgium. *Governance* 16: 321–348

- Dull, Matthew. 2006. Why PART? The Institutional Politics of Presidential Budget Reform. *Journal of Public Administration Research and Theory* 16(2): 187–215.
- Durant, Robert F. 1987. Toward Assessing the Administrative Presidency: Public Lands, the BLM, and the Reagan Administration. *Public Administration Review* 47: 180–89.
- . 2008. Sharpening a Knife Cleverly: Organizational Change, Policy Paradox, and the "Weaponizing" of Administrative Reforms. *Public Administration Review* 68: 282-294.
- Frederickson, David G., and H. George Frederickson. 2007. *Measuring the Performance of the Hollow State*. Washington, DC: Georgetown University Press.
- Frisco, Velda and Odd J. Stalebrink. 2008. Congressional Use of the Program Assessment Rating Tool. *Public Budgeting and Finance* 28(2): 1-19.
- Gailmard, Sean, and John W. Patty. 2007. Slackers and Zealots: Civil Service, Policy Discretion, and Bureaucratic Expertise. *American Journal of Political Science* 51(4): 873-89.
- Gallo, Nick and David E. Lewis. 2012. The Consequences of Presidential Patronage for Federal Agency Performance. *Journal of Public Administration Research and Theory* 22(2): 219-43.
- Gilmour, John B. 2006. *Implementing OMB's Program Assessment Rating Tool (PART): Meeting the Challenges of Integrating Budget and Performance*. Washington D.C.: IBM Center for the Business of Government.
- Gilmour, John B., and David E. Lewis. 2006a. Assessing performance budgeting at OMB: The influence of politics, performance, and program size. *Journal of Public Administration Research and Theory* 16(2): 169-86.
- . 2006b. Does Performance Budgeting Work? An Examination of the Office of Management and Budget's PART Scores. *Public Administration Review* 66(5): 742-52
- . 2006c. Political appointees and the competence of Federal Program Management.



- American Politics Research* 34(1): 22-50.
- Golden, Marissa. 2000. *What Motivates Bureaucrats?* New York: Columbia University Press.
- Greitens, Thomas J., and M. Ernita Joaquin. 2010. Policy Typology and Performance Measurement: Results from the Program Assessment Rating Tool (PART). *Public Performance and Management Review* 33(4): 555–70.
- Heinrich, Carolyn J. 2012. How Credible is the Evidence, and Does It Matter? An Analysis of the Program Assessment Rating Tool. *Public Administration Review* 72(1): 123-134.
- Joyce, Philip G. 2011. The Obama Administration and PBB: Building on the Legacy of Federal Performance-Informed Budgeting. *Public Administration Review* 71(3): 356-67.
- Kickert, Walter. 2011. Distinctive of Administrative Reform in Greece, Italy, Portugal and Spain. Common Characteristics of Context, Administrations and Reforms. *Public Administration* 89(3): 801-818.
- Lavertu, Stéphane and Donald P. Moynihan. 2012. The Effect of Agency Political Ideology on Reform Implementation: Performance Management in the Bush Administration. Forthcoming at *Journal of Public Administration Research and Theory*.
- Lazear, Edward P., and Michael Gibbs. 2008. *Personnel Economics in Practice*, 2<sup>nd</sup> Ed. New York: Wiley.
- Lee, Kwang-Hoon and Jos C. N. Raadschelders. 2008. Political-Administrative Relations: Impact of and Puzzles in Aberbach, Putnam, and Rockman, 1981. *Governance*, 21(3), 419-438.
- Milgrom, Paul, and John Roberts. 1992. *Economics, Organization and Management*. Upper Saddle River, NJ: Prentice Hall.
- Moynihan, Donald P. 2008. *The Dynamics of Performance Management: Constructing*

- Information and Reform*. Washington, DC: Georgetown University Press.
- Moynihan, Donald P. and Stéphane Lavertu. 2012. Does Involvement in Performance Reforms Encourage Performance Information Use? Evaluating GPRA and PART. *Public Administration Review* 72(4): 592-602.
- Pratt, John W., and Richard J. Zeckhauser, eds. 1985. *Principals and Agents: The Structure of Business*. Boston: Harvard Business School Press.
- Radin, Beryl. 2006. *Challenging the Performance Movement: Accountability, Complexity, and Democratic Values*. Washington, DC: Georgetown University Press.
- 2009. What Can We Expect from Performance Measurement Activities? *Journal of Policy Analysis and Management* 28: 505-512.
- Redburn, F. Stevens, and Kathryn Newcomer. 2008. *Achieving Real Improvement in Program Performance and Policy Outcomes: The Next Frontier*. Washington, DC: National Academy of Public Administration.
- Rouban, Luc. 2007. Public Management and Politics: Senior Bureaucrats in France. *Public Administration* 85(2): 473-501
- Stalebrink, Odd and Velda Firsco. 2011. PART in Retrospect: An Examination of Legislator's Attitudes toward PART. *Public Budgeting and Finance* 31(2): 1-21.
- Ting, Michael M. 2003. A Strategic Theory of Bureaucratic Redundancy. *American Journal of Political Science* 47(2): 274-292.
- U.S. General Accounting Office. 2004. *Performance Budgeting: Observations on the Use of OMB's Program Assessment Rating Tool for the Fiscal Year 2004 Budget*. Washington, DC: General Accounting Office (Available at <http://www.gao.gov/assets/250/241315.pdf>, last accessed January 6, 2012).

- U.S. Government Accountability Office. 2008. *Government Performance: Lessons Learned for the Next Administration on Using Performance Information to Improve Results*. Washington, DC: Government Accountability Office (available at <http://www.gao.gov/assets/130/120846.pdf>, last accessed January 6, 2012).
- Van de Walle, Steven, and Wouter Van Dooren (eds). 2008. *Performance Information in the Public Sector: How it is Used*. Houndmills, UK: Palgrave.
- White, Joseph. 2012. Playing the Wrong PART: The Program Assessment Rating Tool and the Functions of the President's Budget. *Public Administration Review* 72(1): 112-21.

**Table 1. Variables based on Lewis's survey data**

<i>Name</i>	<i>Description</i>	<i>N</i>	<i>Min, Max</i>	<i>Mean (SD)</i>
Effort	Answer to "How much time and effort did your agency put into the Program Assessment Rating Tool (PART) process?" Answers range from "Not much at all" (0) to "Tremendous amount" (4).	891	0, 4	2.92 (0.91)
Percentage Democratic	Percentage of agency executives that self-identify as Democrats by agency. The average is based on all respondents, whether or not they indicate involvement with PART reviews.	1000	0.00, 1.00	0.58 (0.18)
Liberal Agency	Measure of agency ideology from Clinton & Lewis (2008). Indicates whether agencies are liberal (1) or not (0). Some agencies in Table A1 are coded according to larger agencies in which they reside. Thirty-eight agencies coded.	991	0, 1	0.31 (0.46)
Budget	Answer to "To what extent did PART scores influence the president's budget request for the program(s) evaluated?" Answers range from "Not much at all" (0) to "Tremendous amount" (4).	694	0, 4	1.25 (1.24)
Employment	Number of employees working in agency in September 2007.	982	9, 664,779	80,780 (125,086)
Average Program Budget	The average program size for programs evaluated in each agency (in millions of dollars).	973	63, 206,842	8137 (34,378)
# of Programs Evaluated in PART	Number of agency programs evaluated in the PART process	1000	0, 115	49.89 (34.69)
Cabinet	Whether (1) or not (0) respondent's agency is a cabinet department or located within a cabinet department.	1000	0, 1	0.76 (0.43)
Independent Commission	Whether (1) or not (0) respondent works in a commission outside the cabinet.	1000	0, 1	0.07 (0.25)
Democrat	Whether (1) or not (0) the respondent identified himself or herself as a Democrat or a Democrat-leaning independent.	977	0, 1	0.60 (0.49)
Regional Office	Whether (1) or not (0) the respondent works in a regional office.	1000	0, 1	0.16 (0.37)
Number of Years in Federal Government	Number of years the respondent has worked for the federal government.	994	0, 47	26.38 (9.60)

**Note:** Summary statistics are restricted to respondents who are career civil servants and who indicated some level of involvement with PART reviews. "Don't know" is coded as missing for all variables.

**Table 2. Reported agency time and effort expended for PART review process**

<b>Agency Ideological Divergence</b>					
Percentage Democratic	1.28** (0.45)	1.04** (0.43)	--		
Liberal Agency	--	--	0.40** (0.14)	0.37** (0.11)	
<b>Budgetary Incentives</b>					
Budget	--	0.19** (0.04)	--	0.18** (0.03)	
<b>Administrative Burden</b>					
Ln (Employment)	--	0.08 (0.06)	--	0.12 (0.07)	
# of Programs Evaluated in PART	--	0.01** (0.00)	--	0.01** (0.00)	
Ln(Average Program Budget)	--	-0.08** (0.04)	--	-0.13** (0.04)	
<b>Controls</b>					
Cabinet	--	-0.45** (0.14)	--	-0.41** (0.15)	
Ind. Commission	--	0.15 (0.22)	--	0.23 (0.22)	
Democrat	--	0.07 (0.08)	--	0.12* (0.07)	
Regional Office	--	-0.24 (0.19)	--	-0.26 (0.18)	
Number of Years in Federal Government	--	-0.00 (0.00)	--	-0.00 (0.00)	
	N	891	654	883	652
# of Agency Clusters	45	38	40	37	
Wald Chi <sup>2</sup>	8.26**	137.91**	8.61**	107.76**	
Pseudo R <sup>2</sup>	0.01	0.05	0.01	0.05	
<b>Note:</b> The results are from ordered probit models that estimate the amount of reported agency time and effort devoted to PART reviews. The dependent variable is <i>effort</i> . The inclusion of the “other” variables greatly reduces the sample size. Standard errors are clustered by agency and reported in parentheses below regression coefficients. Cut point estimates are not reported. Significance levels are based on two-tailed z-tests or chi-square tests: **p<0.05 and *p<0.10 (so that *p<0.05 for a one-tailed test).					

**Table 3. Agency-level measures averaged and reported by agency ideological categories**

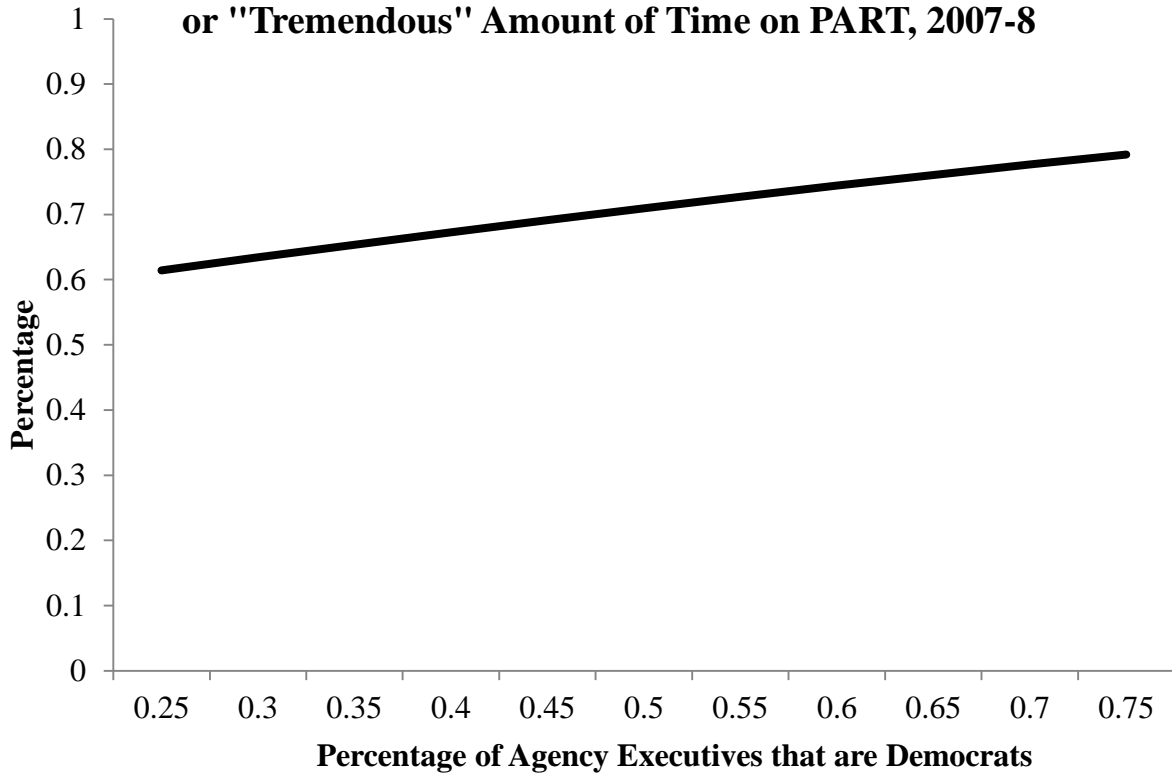
	Agencies for which the percentage of Democratic employees is less than or equal to 50%	Agencies for which the percentage of Democratic employees is greater than 50%	Moderate and conservative agencies, according to Clinton & Lewis (2008)	Liberal agencies, according to Clinton & Lewis (2008)
Average agency time and effort put into the PART process (1-5 scale)	3.66 (0.50)	3.82 (0.50)	3.64 (0.49)	4.07 (0.18)
Average proportion of respondents reporting PART involvement	0.09 (0.11)	0.13 (0.13)	0.14 (0.10)	0.18 (0.07)
Average number of employees in 2007	61,087.50 (146,294.60)	26,490.03 (51,286.39)	85,213.50 (141,461.50)	18,943.00 (26,698.84)
Average number of programs "PARTed"	9.72 (18.30)	18.23 (29.45)	26.00 (25.81)	34.80 (40.95)
Average number of "PARTed" regulatory programs	0.96 (2.20)	1.48 (3.05)	2.35 (3.02)	2.60 (4.56)
Average number of "PARTed" R&D programs	1.92 (6.36)	1.63 (3.87)	3.65 (7.17)	2.90 (5.04)
Percent of agency managers that agreed to a "moderate," "great," or "very great" extent that PART "imposed a significant burden on management resources."	55.74	62.01	56.31	66.78
N	29	40	26	10
<b>Note:</b> The above statistics result from averaging agency-level data. The standard deviation is reported in parentheses below each agency-level average. The measure of time and effort reported by individuals was averaged at the agency level before the reported means and standard deviations were calculated. The number of agencies used to create the measures ranges from 10 to 40.				

**Table 4. Reported agency time and effort expended for PART review process**

<b>Agency Ideological Divergence</b>		
Percentage Democratic	0.67* (0.39)	
Liberal Agency	--	0.22** (0.11)
<b>Budgetary Incentives</b>		
Budget	0.19** (0.04)	0.18** (0.03)
<b>Administrative Burden</b>		
Ln (Employment)	0.02 (0.07)	0.04 (0.09)
# of Programs Evaluated in PART	0.01** (0.00)	0.01** (0.00)
Ln(Average Program Budget)	-0.04 (0.05)	-0.07 (0.05)
# of Regulatory Programs	0.03** (0.02)	0.03 (0.02)
# of R&D Programs	-0.01* (0.00)	-0.01** (0.01)
<b>Controls</b>		
Cabinet	-0.33** (0.16)	-0.33** (0.16)
Ind. Commission	0.07 (0.21)	0.11 (0.21)
Democrat	0.07 (0.08)	0.10 (0.07)
Regional Office	-0.24 (0.20)	-0.25 (0.19)
Number of Years in Federal Government	-0.00 (0.00)	-0.00 (0.00)
N	654	652
# of Agency Clusters	38	37
Wald Chi2	233.88**	175.51**
Pseudo R2	0.05	0.05

**Note:** The results are from ordered probit models that estimate the amount of reported agency time and effort devoted to PART reviews. The dependent variable is *effort*. Standard errors are clustered by agency and reported in parentheses below regression coefficients. Cut point estimates are not reported. Significance levels are based on two-tailed z-tests or chi-square tests: \*\*p<0.05 and \*p<0.10 (so that \*p<0.05 for a one-tailed test).

**Figure 1. Influence of Agency Partisanship on Probability that Respondent Reports that Agency Spent a "Significant" or "Tremendous" Amount of Time on PART, 2007-8**



Note: Estimated effects calculated with all values held at their means or modes (for discrete variables) using estimates from the second model in Table 2.



**Appendix A. Reported agency time and effort expended for PART review process**

<b>Agency Ideological Divergence</b>					
Average Agency Respondent Ideology <sup>1</sup>	0.43** (0.19)	0.42** (0.19)	--		
Average Agency Ideal Point <sup>2</sup>	--	--	0.73** (0.27)	0.51** (0.26)	
<b>Budgetary Incentives</b>					
Budget	--	0.18** (0.04)	--	0.18** (0.04)	
<b>Administrative Burden</b>					
Ln (Employment)	--	0.09 (0.07)	--	0.08 (0.07)	
# of Programs Evaluated in PART	--	0.01** (0.00)	--	0.01** (0.00)	
Ln(Average Program Budget)	--	-0.09** (0.04)	--	-0.08** (0.03)	
<b>Controls</b>					
Cabinet	--	-0.41** (0.15)	--	-0.41** (0.15)	
Ind. Commission	--	0.27 (0.23)	--	0.18 (0.24)	
Democrat	--	0.09 (0.07)	--	0.10 (0.07)	
Regional Office	--	-0.25 (0.19)	--	-0.25 (0.19)	
Number of Years in Federal Government	--	-0.00 (0.00)	--	-0.00 (0.00)	
N	885	651	891	654	
# of Agency Clusters	45	38	45	38	
Wald Chi2	4.96**	132.56**	7.16**	121.13**	
Pseudo R2	0.01	0.05	0.01	0.05	
<b>Note:</b> The results are from ordered probit models that estimate the amount of reported agency time and effort devoted to PART reviews. The dependent variable is <i>effort</i> . Standard errors are clustered by agency and reported in parentheses below regression coefficients. Cut point estimates are not reported. Significance levels are based on two-tailed z-tests or chi-square tests: **p<0.05 and *p<0.10 (so that *p<0.05 for a one-tailed test).					

<sup>1</sup> Whether a respondent identified herself or himself as very conservative (1), conservative (2), somewhat conservative (3), moderate (4), somewhat liberal (5), liberal (6), or very liberal (7).

<sup>2</sup> Estimate of the respondent's ideology based on his or her responses to items inquiring about his or her level of support for a number of bills (Clinton et al. 2012). Higher values indicate greater liberalism.

<b>Appendix B. Reported agency time and effort expended for PART review process</b>		
<b>Agency Ideological Divergence</b>		
Percentage Democratic	0.76* (0.46)	
Liberal Agency	--	0.19 (0.17)
<b>Budgetary Incentives</b>		
Budget	0.19** (0.03)	0.19** (0.03)
<b>Administrative Burden</b>		
Ln (Employment)	0.11 (0.08)	0.12 (0.08)
# of Programs Evaluated in PART	0.01* (0.00)	0.01** (0.00)
Ln(Average Program Budget)	-0.09* (0.05)	-0.12** (0.05)
<b>Controls</b>		
Cabinet	-0.56** (0.14)	-0.56** (0.18)
Ind. Commission	0.20 (0.32)	-0.10 (0.34)
Democrat	0.08 (0.08)	0.11 (0.07)
Regional Office	-0.22 (0.20)	-0.23 (0.19)
Number of Years in Federal Government	-0.00 (0.01)	-0.00 (0.00)
% of Regulatory Programs	0.50 (0.37)	0.32 (0.34)
% of R&D Programs	-0.06 (0.43)	-0.19 (0.47)
% of Block/Formula Grant Programs	-0.37 (0.65)	-0.25 (0.59)
% of Capital Assets & Service Acq. Programs	-0.34 (0.68)	-0.45 (0.80)
% of Competitive Grant Programs Evaluated	0.64 (0.47)	0.55 (0.54)
% of Credit Programs	0.37 (0.67)	0.01 (0.63)
N	654	652
# of Agency Clusters	38	37
Wald Chi2	244.22**	167.01**
Pseudo R2	0.05	0.05
<b>Note:</b> The results are from ordered probit models that estimate the amount of reported agency time and effort devoted to PART reviews. The dependent variable is <i>effort</i> . Standard errors are clustered by agency and reported in parentheses below regression coefficients. Cut point estimates are not reported. Significance levels are based on two-tailed z-tests or chi-square tests: **p<0.05 and *p<0.10 (so that *p<0.05 for a one-tailed test). Likelihood ratio tests of nested models (5.71, 6df; 4.03, 6df), respectively.		

<b>Appendix C. Reported agency time and effort expended for PART review process</b>		
<b>Agency Ideological Divergence</b>		
Percentage Democratic	0.70* (0.41)	
Liberal Agency	--	0.20 (0.20)
<b>Budgetary Incentives</b>		
Budget	0.18** (0.03)	0.18** (0.03)
<b>Administrative Burden</b>		
Ln (Employment)	0.03 (0.07)	0.05 (0.09)
# of Programs Evaluated in PART	0.01* (0.00)	0.01 (0.00)
Ln(Average Program Budget)	-0.04 (0.05)	-0.08 (0.06)
<b>Controls</b>		
Cabinet	-0.40** (0.15)	-0.39** (0.16)
Ind. Commission	0.07 (0.22)	0.09 (0.26)
Democrat	0.07 (0.08)	0.10 (0.07)
Regional Office	-0.23 (0.20)	-0.24 (0.20)
Number of Years in Federal Government	-0.00 (0.01)	-0.00 (0.00)
# of Regulatory Programs	0.04 (0.03)	0.03 (0.04)
# of R&D Programs	-0.01 (0.01)	-0.01 (0.01)
# of Block/Formula Grant Programs	-0.01 (0.02)	-0.01 (0.02)
# of Capital Assets & Service Acq. Programs	-0.00 (0.02)	-0.01 (0.02)
# of Competitive Grant Programs Evaluated	0.01 (0.01)	0.01 (0.02)
# of Credit Programs	0.03 (0.02)	0.01 (0.02)
N	654	652
# of Agency Clusters	38	37
Wald Chi2	274.79**	226.33**
Pseudo R2	0.06	0.05
<b>Note:</b> The results are from ordered probit models that estimate the amount of reported agency time and effort devoted to PART reviews. The dependent variable is <i>effort</i> . Standard errors are clustered by agency and reported in parentheses below regression coefficients. Cut point estimates are not reported. Significance levels are based on two-tailed z-tests or chi-square tests: **p<0.05 and *p<0.10 (so that *p<0.05 for a one-tailed test). Likelihood ratio tests of nested models (6.59, 6df; 5.61, 6df), respectively.		

---

<sup>i</sup> Typically, studies let survey respondents decide how to interpret what is meant by agency or program “ideology.” U.S. political ideology often is defined on a one-dimensional scale ranging from liberal (associated with the Democratic Party) to conservative (associated with the Republican Party) and typically is thought to reflect individuals’ political views about the proper role of government and the substance of public policy. The liberal left generally is thought to prefer greater government involvement in societal affairs—for example, via public programs that regulate the marketplace, redistribute societal resources, and promote social equality. The conservative right is thought to prefer policies promoting greater market freedom, smaller government, and greater national security. Though it is an imprecise construct, it is a heuristic that people in and out of government use to make inferences about the policy preferences and intentions of government actors, programs, and organizations.

<sup>ii</sup> Higher level political appointees are underrepresented in the sample. We focus here on the responses of career professionals who indicate involvement with the PART process. Investigators obtained a list of 7,448 federal administrators and program managers from Leadership Directories, Inc., the firm that publishes the *Federal Yellow Book*. Once this list was cleaned to remove names that were incorrectly included there were 7,151 names. In total 2,398 persons completed the survey. The original list also included 461 NSF executives since Leadership Directories, Inc. coded NSF program directors as program managers. If the NSF is excluded from the sample there are 2,250 respondents from 6,690 potential respondents. To verify the representativeness of the sample, the survey’s authors hired private firms to match unique names to home addresses, collect voter registration information, and compare party registration with self-reported partisanship. For complete details see Clinton et al. 2012.

<sup>iii</sup> Of the 1,737 respondents that answered this question, 1,000 indicated that their agency had a program evaluated, 113 indicated that their agency had not had a program evaluated, and 624 reported that they did not know.

<sup>iv</sup> Clinton and Lewis (2008) asked respondents in academia, think tanks, and specialized media outlets dealing with the federal government whether agencies “tended to be liberal, conservative, or neither consistently” and used responses to generate numerical estimates of agency ideology. The estimation technique allowed them to account for different definitions of liberal and conservative and the quality of the ratings. All agencies whose estimates were statistically distinguishable from 0 in a liberal direction were coded with a 1. All other agencies were coded with a 0.

<sup>v</sup> The ideology measure is based upon respondents’ answers to questions about whether the respondent identified herself or himself as very conservative (1), conservative (2), somewhat conservative (3), moderate (4), somewhat liberal (5), liberal (6), or very liberal (7). Responses are averaged by agency. The average ideal point measure is based upon respondent ideal points generated using survey questions about how respondents themselves would have voted on bills in Congress. Using these survey respondents as “votes” Clinton et al. (2012) generate ideal point estimates through standard techniques employed to generate such measures for members of Congress. The results confirm what is reported here. See Appendix A.

<sup>vi</sup> We have also estimated models using respondent beliefs regarding the extent to which PART scores influenced appropriations with similar results.

<sup>vii</sup> Source: Office of Personnel Management, Central Personnel Data File ([fedscope.opm.gov](http://fedscope.opm.gov)). We log this measure to account for extreme values.

<sup>viii</sup> It is possible that the composition of programs in different agencies is correlated with agency ideology and program budget. To account for this we have also estimated models including the percentage and number of programs in all different program types as identified by the Bush Administration. We could not reject the null in any case that the inclusion of these additional controls did not improve the fit of the model. We have included estimates from these models in Appendix B (percentages of programs of different types) and Appendix C (numbers of programs of different types), respectively.

<sup>ix</sup> The GAO administered the survey to a random, nationwide sample of mid- and upper-level federal employees in the agencies covered by the Chief Financial Officers Act of 1990, with an oversampling of managers from certain agencies to facilitate comparisons across 29 different agencies. The response rate to the survey was 70% overall, ranging between 55% and 84% across agencies. See Moynihan and Lavertu (2012) for a more detailed description of the data.

<sup>x</sup> Agencies can be aggregated at different levels, from sub-cabinet bureau to larger department. The decision to cluster standard errors matters little for the substantive conclusions of the models. In the models in Table 2 we cluster standard errors at a higher level of aggregation (referred to as *agcode2* in the SFGS data) to be consistent with the aggregation of other variables (employment, cabinet, etc.). When models are estimated clustering at a lower level of aggregation (i.e., at the bureau level rather than department level), the coefficient estimates are generally more precise. The primary differences between these models and the models in Table 2 are that the coefficients on persons working in regional offices become significant at the 0.10 level in both models. In addition, in the fourth model in Table 2, the coefficient on the log of agency employment becomes significant at the 0.10 level and the

---

coefficient on whether the respondent is a Democrat loses significance ( $p < 0.29$ ). Results are available from the authors upon request.

<sup>xi</sup> We have also estimated models only on non-Democratic respondents and the results are substantively similar. The coefficient estimate on the agency ideology is larger and significant at the 0.05 level in two-tailed tests.

<sup>xii</sup> In models estimated using respondent self-reported ideology, self-identified liberals were significantly more likely to report more agency effort in the PART process. We could not reject the null, however, in models using respondent ideal point that respondent ideal points were uncorrelated with *effort*.

<sup>xiii</sup> We have also estimated models with counts of the total number of other types of programs as well and include these in Appendix B. We do not include these models in the main text since the counts of these variables are highly correlated with the each other and the total number of programs evaluated—sometimes as high as 0.94. This multicollinearity increases the size of the standard errors and decreases the precision of the estimates.