

No. 10-21
May 2010

WORKING PAPER

TEN YEARS OF RESULTS FROM THE RESULTS ACT

By Jerry Ellig

MERCATUS CENTER
George Mason University

The ideas presented in this research are the author's and do not represent official positions
of the Mercatus Center at George Mason University.

Ten Years of Results from the Results Act

Jerry Ellig*
Senior Research Fellow
Mercatus Center at George Mason University
3301 North Fairfax Dr., Suite 208, Arlington, Virginia 22201, USA
jellig@gmu.edu

Prepared for presentation at *Stronger Nations, Stronger Relations: New Prospects for Asia-Pacific Regional Integration* 1st Annual IDAS International Conference, Chengchi National University, Taipei, Taiwan, June 1–2, 2010

ABSTRACT

The Government Performance and Results Act (GPRA) of 1993 directed U.S. federal agencies to produce strategic plans with outcome-oriented objectives, annual performance plans with performance goals, and annual performance reports that measure progress toward those goals. The legislation sought to improve internal management of programs and congressional decision making by making better information available about the effectiveness and efficiency of federal programs and spending. How well has GPRA accomplished those goals?

This paper summarizes lessons learned from a ten-year research project that evaluated the quality of annual performance reports produced under GPRA by the 24 U.S. federal agencies that account for more than 95 percent of all federal spending. The *Mercatus Performance Report Scorecard* evaluated agency reports based on 12 principal criteria found in GPRA. GPRA has significantly improved the quality of performance information. On average, the quality of agency performance reports improved by about 75 percent between fiscal year 1999 and fiscal year 2008. There is still substantial variation in quality, with only a few reports each year employing best practices on each of the *Scorecard* evaluation criteria. Factors like agency size, program structure, and ideology have little or no correlation with the quality of agency performance reports. A focused strategic plan with outcome-oriented goals and measures is one necessary condition for a high-quality GPRA performance report.

GPRA has improved the availability and use of performance information in agencies. The quality of agency GPRA initiatives is positively correlated with surveys of federal managers on the availability and use of performance measures.

Finally, there is little evidence that GPRA has altered congressional budget decisions. Linkage of results to costs is the weakest aspect of agency performance reports. Results information affected some presidential budget proposals, but Congress has shown little interest in using results information to make budget decisions.

*I would like to thank Marcus Peacock, Robert Shea, and Richard Williams for comments on earlier drafts.

Introduction

The Government Performance and Results Act of 1993 (GPRA) requires U.S. federal agencies to produce strategic plans with performance measures, annual performance plans, and annual performance reports. Strategic plans must explain the outcomes agencies seek to produce for citizens and establish measures the agencies will use to track progress. Annual performance reports must report on the measures and explain the agency's plans to improve performance in the future.

Congress enacted GPRA in part because "Federal managers are seriously disadvantaged in their efforts to improve program efficiency and effectiveness, because of insufficient articulation of program goals and inadequate information on program performance." The legislation also notes that "congressional policymaking, spending decisions and program oversight are seriously handicapped by insufficient attention to program performance and results" (GPRA Sec. 2a). The legislation sought to improve program management and congressional decision making by making better information available about the effectiveness and efficiency of federal programs and spending.

The congressional findings cited above suggest at least three ways to assess GPRA's effects:

1. Has the quality of performance information produced by agencies improved?
2. Has GPRA led to greater availability and use of performance information by federal managers?
3. Has GPRA led to greater use of performance information in budget decisions?

This paper summarizes results from the Mercatus Center's *Performance Report Scorecard*, U.S. Government Accountability Office (GAO) reports, and academic studies to answer those questions.

1. Quality of Performance Information

Mercatus Center Performance Report Scorecard

In 1999, the Mercatus Center at George Mason University initiated a ten-year research project that evaluated the quality of annual performance reports produced under GPRA by the 24 U.S. federal agencies that account for more than 95 percent of all federal spending. The *Mercatus Performance Report Scorecard* evaluated agency reports based on 12 principal criteria found in GPRA.

Table 1 lists the criteria. On each criterion, a report could achieve a score ranging from 1 point (no useful content) to 5 points (potential best practice). Thus, the possible report scores range from 12 to 60 points. The *Scorecard* did not offer an opinion on the quality of agencies' performance, nor did it express views on what activities the government should or should not undertake. It assessed the quality of disclosure, not the quality of results.

Table 1: Mercatus *Scorecard* Criteria

Transparency: How easily can a non-specialist find and understand the report?

1. Accessibility: Is the report easily accessible via the Internet and easily identified?
2. Readability: Is the report easy for a layperson to read and understand?
3. Verification and Validation: Are the performance data valid, verifiable, and timely?
4. Baseline and Trend Data: Did the agency provide baseline and trend data to put its performance measures in context?

Public Benefits: How well does the report document the outcomes the agency produces for the public and compare them with costs?

5. Outcome Goals: Are the goals and objectives stated as outcomes?
6. Outcome Measures: Are the performance measures valid indicators of the agency's impact on its outcome goals?
7. Agency Affected Outcomes: Does the agency demonstrate that its actions have actually made a significant contribution toward its stated goals?
8. Linkage to Costs: Did the agency link its goals and results to costs?

Leadership: How well does the report demonstrate that agency managers use performance information to make decisions?

9. Vision: Does the report show how the agency's results will make this country a better place to live?
 10. Explain Failures: Does the agency explain failures to achieve its goals?
 11. Major Management Challenges: Does the report adequately address major management challenges?
 12. Improvement Plans: Does it describe changes in policies or procedures to do better next year?
-
-

Source: McTigue et. al. (2009, 5)

Scoring Improvement Over Time

The quality of GPRA reports improved substantially during the ten years of the *Scorecard* project. Table 2 shows the change in each report's score between fiscal year 1999 and fiscal year 2008. For the 17 reports whose scores improved, the average increase was 8.94 points, almost double the average increase of 4.84 points for all 24 reports. Nine reports achieved double-digit increases in their scores. Figure 1 shows that average scores increased by about 15 percent between fiscal 1999 and fiscal 2008.

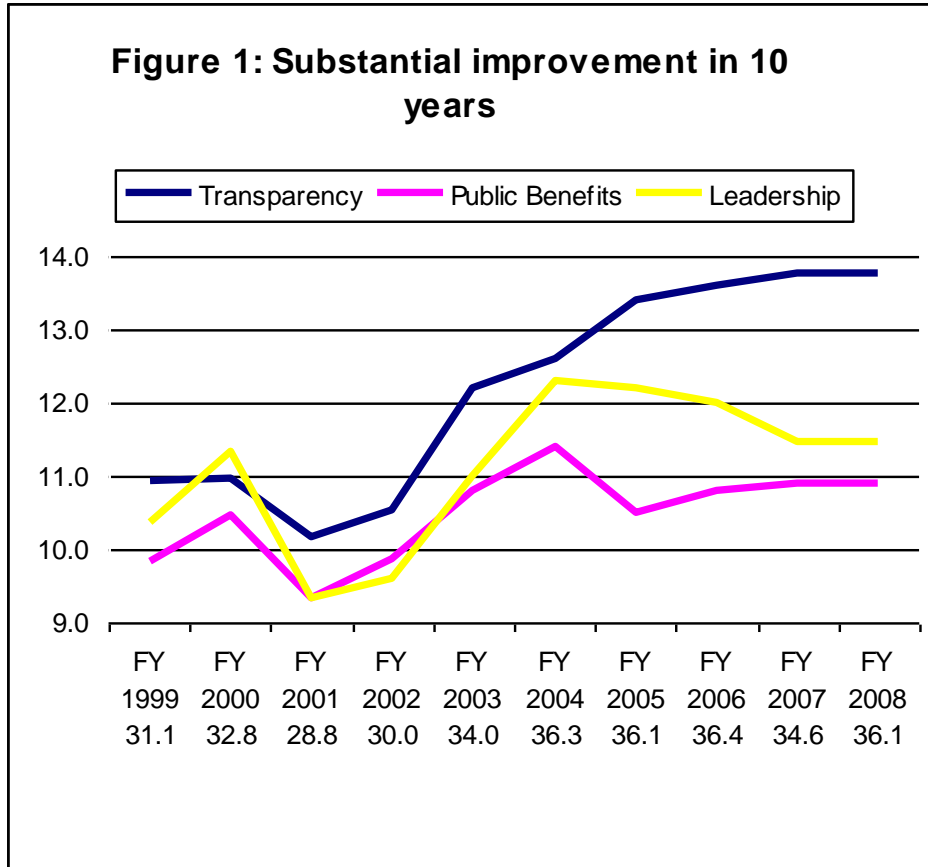
Table 2: Fiscal Year 2008 Scores & Ranks vs. Fiscal 1999

Highest Rank = 1; Lowest = 24. Maximum Possible Score = 60; Minimum = 12

	Fiscal Year 2008		Fiscal Year 1999		Change in Score	Change in Rank
	Total Score	Rank	Total Score	Rank		
Labor	56	1	36	5	+20	+4
Veterans	54	2	48	3	+6	+1
Transportation	53	3	51	2	+2	-1
DHS*	40	4	27	22	+13	+18
NRC	40	4	25	17	+15	+13
Education	37	6	37	4	0	-2
Interior	37	6	31	11	+6	+5
State	37	6	25	17	+12	+11
Treasury	37	6	36	5	+1	-1
Energy	36	10	27	14	+9	+4
EPA	36	10	31	11	+5	+1
HHS	36	10	24	20	+12	+10
USAID	36	10	52	1	-16	-9
Commerce	35	14	22	22	+13	+8
Justice	34	15	23	21	+11	+6
Agriculture	33	16	22	22	+11	+6
GSA	32	17	32	9	0	-8
NSF	32	17	21	24	+11	+7
Social Security	32	17	33	8	-1	-9
NASA	31	20	27	14	+4	-6
OPM	28	21	27	14	+1	-7
HUD	27	22	28	13	-1	-9
Defense	26	23	34	7	-8	-16
SBA	22	24	32	9	-10	-15
Average	36.13		31.29		4.83	
Median	36.00		29.50		6.50	

*Since DHS did not exist in 1999, the chart shows its score and rank from fiscal year 2004, the first year its report was included in the Scorecard.

Source: McTigue et. al. (2009, 10)



Source: McTigue et. al. (2009, 8)

All of these score data understate the full extent of improvement because the research team tightened the scoring criteria over time as new best practices emerged. The ideal way to measure improvement would be to reevaluate all of the fiscal year 1999 reports using fiscal year 2008 standards. Resource constraints precluded this. However, for the final *Scorecard* in 2008, the research team re-examined the top four reports from fiscal year 1999 using the same standards applied in fiscal year 2008. Table 3 shows the results. Evaluated by fiscal year 2008 standards, the best fiscal year 1999 report (from USAID) would have ranked 16th in fiscal year 2008, with just 33 points out of a possible 60. The other fiscal year 1999 reports would have ranked even lower.

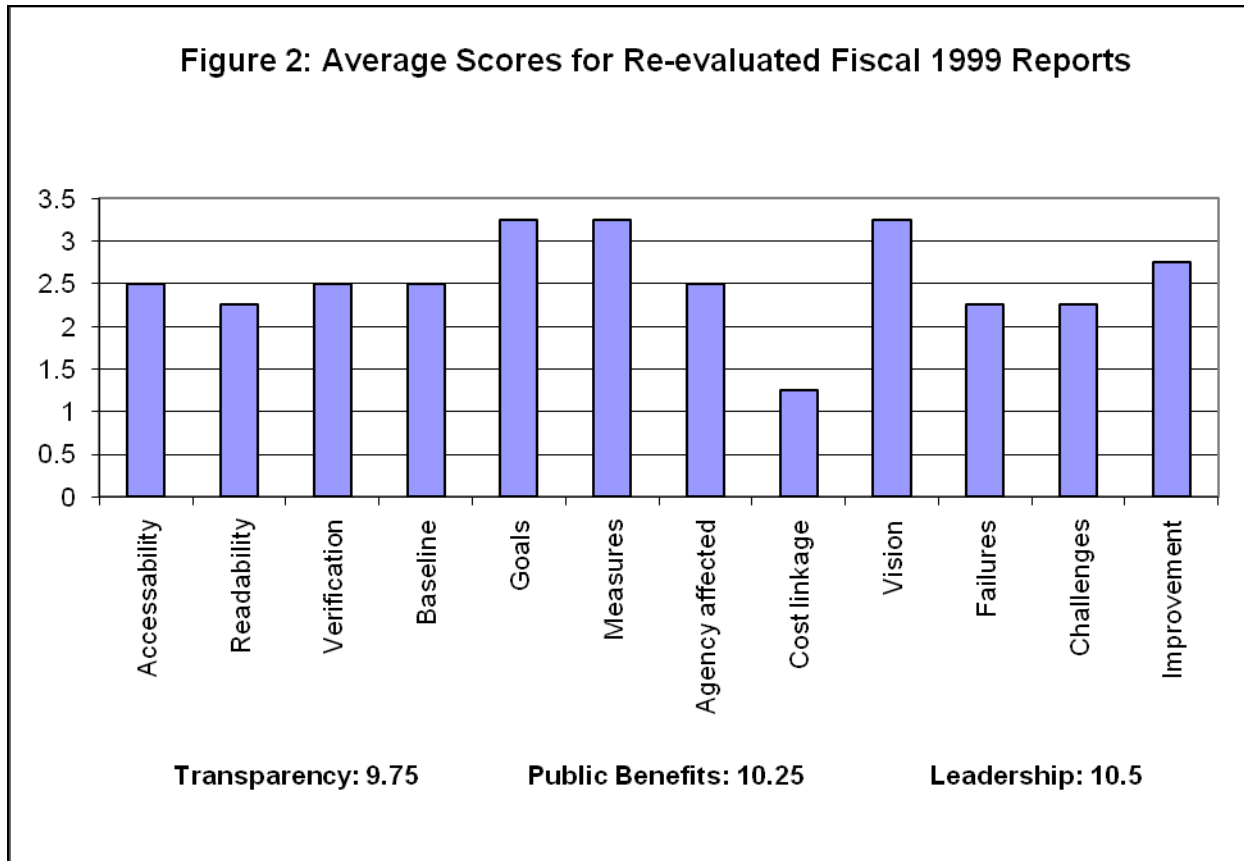
Table 3: Top Four 1999 Reports Would Rank Low in 2008

Highest Rank = 1; Lowest = 24. Maximum Possible Score = 60; Minimum = 12.

	<u>Transparency</u>	<u>Public Benefits</u>	<u>Leadership</u>	<u>Total</u>	<u>Rank</u>
Labor	20	19	17	56	1
Veterans	19	16	19	54	2
Transportation	16	20	17	53	3
DHS	15	13	12	40	4
NRC	15	13	12	40	4
Education	14	12	11	37	6
Interior	16	10	11	37	6
State	15	10	12	37	6
Treasury	14	10	13	37	6
Energy	13	11	12	36	10
EPA	13	11	12	36	10
HHS	13	13	10	36	10
USAID	15	10	11	36	10
Commerce	15	10	10	35	14
Justice	15	8	11	34	15
Agriculture	12	10	11	33	16
USAID 1999	9	11	13	33	16
GSA	11	12	9	32	17
NSF	15	7	10	32	17
Social Security	12	8	12	32	17
NASA	11	8	12	31	20
Transportation 1999	9	12	10	31	20
Veterans 1999	11	10	10	31	20
OPM	11	8	9	28	21
HUD	11	8	8	27	22
Education 1999	10	8	9	27	22
Defense	11	7	8	26	23
SBA	8	8	6	22	24
Fiscal Year 2008					
Average	13.8	10.9	11.5	36.1	
Fiscal Year 2008					
Median	14.0	10.0	11.0	36.0	

Source: McTigue et. al. (2009, 12)

Figure 2 shows how these four reports scored on individual criteria. Average scores on criteria 5 (outcome-oriented goals), 6 (outcome-oriented measures) and 9 (vision) all exceed the satisfactory score of 3—even when evaluated by fiscal year 2008 standards. This suggests that even in the early days of GPRA, the higher-ranking agencies got off to a good start in formulating outcome-oriented goals and measures. However, the average score on criterion 8 (linkage of results to costs) barely exceeds 1. This indicates that, compared to current practice, even the top scorers had little cost-related content in fiscal year 1999.



Source: McTigue et. al. (2009, 11)

We can measure how these specific agencies' reports have improved by comparing the scores for their fiscal year 2008 reports with the scores on their fiscal year 1999 reports evaluated under fiscal year 2008 standards. Using the scores reported in Table 3, USAID's report improved by about 9 percent over ten years (from 33 to 36 points), Education's report improved by 37 percent (from 27 to 37 points), Transportation's report improved by 71 percent (from 31 to 53 points), and Veterans Affairs' report improved by 74 percent (from 31 to 54 points).

Extrapolating from these four reports, report quality may have improved by about 75 percent on average.¹ Of course, that means the quality of some agencies' reports increased by even more than that amount, and others by less.

Qualitative analysis of best practices also reveals substantial improvements since fiscal year 1999. Table 4 shows how the "state of the art" has advanced during the past decade. Except for Criterion 1 (accessibility), only a few reports in each year used the best practices. This qualitative description of best practices is consistent with the quantitative score assessment: performance reporting made significant progress between 1999 and 2008.

Improvements in scores and best practices on the *Mercatus Scorecard* are of course just rough indicators of improvements in the quality of useful performance information. As agencies became more familiar with the *Scorecard* criteria, some may have sought to "game" the scoring system by searching for the easiest ways to improve their scores rather than the most useful ways to improve their performance information. By the end of the project, approximately half of the 24 agencies each year were seeking more detailed advice and feedback from the Mercatus Center research team. Though some gaming surely occurred, the size of the score improvements and the nature of the improvement in best practices suggest that agencies also accomplished some genuine improvements in the quality of performance information.

¹ Scores for the four re-evaluated fiscal year 1999 reports averaged one-third lower under the fiscal year 2008 standards than under the fiscal year 1999 standards. If we assume that using the fiscal year 2008 scoring standards would have reduced all fiscal year 1999 scores by one-third, the average fiscal year 1999 score using fiscal year 2008 standards would have been 20.65 instead of the 31.29 shown in Table 2. An increase from 20.65 to the average fiscal year 2008 score of 36.13 implies that the average quality of performance reports improved by at least 75 percent. If the average fiscal year 1999 score using fiscal year 2008 scoring standards would have been 20.65, the percentage improvement is calculated by subtracting this score from the actual fiscal year 2008 score (36.13), then dividing this difference by 20.65.

Table 4: Evolution of Best Practices

Fiscal Year 1999	Fiscal Year 2008
<p>Criterion 1: Accessibility</p> <ul style="list-style-type: none"> • Report available online • Findable via an obvious link from agency home page 	<p>Those features, plus</p> <ul style="list-style-type: none"> • Report posted in a timely fashion after due date • Direct (single click) link from home page • Downloadable as single or multiple files • Contact information provided for questions/comments
<p>Criterion 2: Readability</p> <ul style="list-style-type: none"> • Relative lack of jargon • Lengthy reports easy to scan due to headings, sidebars, tables, and charts 	<p>Those features, plus</p> <ul style="list-style-type: none"> • Performance section focuses on key measures • Goals, measures, and results in tables • Secretary's letter describes fiscal year 2008 results for each goal as well as achievements over time • Citizens' reports include concise summaries of results and reader-friendly links to more information
<p>Criterion 3: Verification and validation</p> <ul style="list-style-type: none"> • Discussion of validation procedures • Acknowledgement of data deficiencies • Outline of plans to develop more or better data 	<p>Those features, plus</p> <ul style="list-style-type: none"> • Assessment of data quality for each goal • Data source provided for each measure • Data definitions, verification and validation information, and limitations discussed for each measure
<p>Criterion 4: Baseline and trend data</p> <ul style="list-style-type: none"> • 5 to 10 years of trend data • Several years of actual data combined with goals for next several years 	<p>New standard</p> <ul style="list-style-type: none"> • Multiple years of data include targets, actual results, and costs • Long-range targets or forecasts provided for each measure
<p>Criterion 5: Outcome-oriented goals</p> <ul style="list-style-type: none"> • Most or all strategic goals are outcome-oriented 	<p>New standard</p> <ul style="list-style-type: none"> • Most goals are clear, measurable outcomes • Outcomes are intuitively meaningful to the public
<p>Criterion 6: Outcome measures</p> <ul style="list-style-type: none"> • Most measures are outcomes or related to outcomes 	<p>New standard</p> <ul style="list-style-type: none"> • Most measures are final or intermediate outcomes

Fiscal Year 1999	Fiscal Year 2008
<p>Criterion 7: Agency affected outcomes</p> <ul style="list-style-type: none"> • Report explains how activities contribute to results in specific cases • Acknowledges other factors that might affect results 	<p>Those features, plus</p> <ul style="list-style-type: none"> • Report consistently describes how activities led to observed results • Performance metrics are highly outcome-oriented
<p>Criterion 8: Results linked to costs</p> <ul style="list-style-type: none"> • Charts show costs, personnel counts, and performance measures for program areas 	<p>New standard</p> <ul style="list-style-type: none"> • Costs broken down by strategic goal and most individual performance measures • This information is provided for several years
<p>Criterion 9: Vision</p> <ul style="list-style-type: none"> • Each strategic goal identifies a result of interest to citizens and states how the department intends to accomplish it 	<p>Those features, plus</p> <ul style="list-style-type: none"> • Narratives cite major accomplishments that affect citizens' quality of life • Narratives linked to outcome-oriented performance measures demonstrating that the narratives describe typical results
<p>Criterion 10: Explanation of failures</p> <ul style="list-style-type: none"> • Failures to meet targets explained • Plans to remedy failures explained 	<p>New standard</p> <ul style="list-style-type: none"> • All performance shortfalls identified, along with plans and a timeline to remedy them • Improvement plans offered even when targets were met
<p>Criterion 11: Management challenges</p> <ul style="list-style-type: none"> • Thorough discussions of major management challenges identified by the agency inspector general and GAO 	<p>Those features, plus</p> <ul style="list-style-type: none"> • Inspector general's report lists major management challenges and assesses agency's progress on them • Agency provides self-assessment of progress and timeline for resolving each challenge • Report explains how each challenge affects strategic goals
<p>Criterion 12: Improvement plans</p> <ul style="list-style-type: none"> • Discussion of each measure includes fiscal year 1999 results, projected fiscal year 2000 performance, and actions planned for fiscal year 2001 	<p>New standard</p> <ul style="list-style-type: none"> • Improvement strategies presented for all shortfalls and major management challenges. • Report describes broader challenges the agency faces and plans for addressing them

Source: McTigue et. al. (2009, 13–16)

What Factors Affect Scores?

Federal agencies have different sizes, missions, and means of achieving their missions. Federal agencies that provide direct services to the public might find it easier to define and gather data on outcomes than agencies that mainly make grants to states or other third parties (Frederickson and Frederickson 2008, Radin 2006, 159–80, Metzenbaum 2005, 285–86, Kettl 1988, 18). Larger agencies might produce better reports because they can afford to put more resources into performance measurement, or they might do worse because they have more programs and are inherently harder to manage due to their size (Kettl 1998, 2–3). Ideology could also play a role (Radin 2006, 91–114, 120–21, 189). Finally, theories of bureaucracy suggest that an agency’s tendency to provide good performance information depends in part on whether agency managers believe that the president or Congress wants them to do so (Downs 1967, Tullock 2005 [1965]).

Table 5 sheds light on these issues by reporting the results of ordinary least squares and tobit regressions that model the 2008 *Scorecard* scores and 1999–2008 improvement as a function of agency size (measured by net cost of operations); percent of budget devoted to block and formula grants, competitive grants, and direct services; ideology; and managers’ perceptions of whether lack of congressional interest or fear of OMB micromanagement are barriers to better performance management in their agencies (see McTigue et. al. 2009, 27–31, for a description and sources of data).² The sample size is very small (24 observations), so these results should be taken with a larger than usual grain of salt.

Neither the structural variables nor agency size are correlated with *Scorecard* scores. (The regressions include the net cost of operations squared because three very large outliers—Defense, Social Security, and HHS—have much larger budgets than the rest of the departments, and worse scores.) Ideology is marginally significant in the tobit regressions, providing very weak evidence that more liberal departments might produce better GPRA reports.

The regressions provide some evidence that perceived lack of congressional interest inhibited improvement in agency scores between 1999 and 2008. The tobit regressions suggest that fear of OMB micromanagement led to lower scores in 2008 and perhaps inhibited improvement between 1999 and 2008.

² Tobit is arguably the more appropriate method to use when the dependent variable always falls within a specified range.

Table 5: Few Structural Factors Influence Scorecard Scores

Dependent variable	2008 Score		1999–2008 Score Change	
	OLS	Tobit	OLS	Tobit
Explanatory variables				
Net cost of operations	.007 [.09]	.007 [.12]	-.06 [.73]	-.06 [-.92]
Net cost of operations Squared	-.00004 [-.33]	-.00004 [-.41]	.00006 [.56]	.00006 [.70]
% Block and formula Grants	.13 [1.30]	.13 [1.64]	.10 [.93]	.10 [1.18]
% Competitive Grants	-.15 [-1.00]	-.15 [-1.27]	-.07 [-.41]	-.07 [-.52]
% Direct federal Services	.08 [1.06]	.08 [1.34]	.10 [1.25]	.10 [1.58]
Ideology	-3.6 [-1.64]	-3.6 [-2.07*]	-3.8 [-1.61]	-3.8 [-2.04*]
Lack of congressional Interest	-.25 [-.90]	-.25 [-1.14]	-.54 [-1.84*]	-.54 [-2.33**]
Fear of OMB Micromanagement	-.39 [-1.70]	-.39 [-2.15**]	-.38 [-1.54]	-.38 [-1.95*]
Constant	46.99 [6.46***]	46.99 [8.17***]	23.06 [2.98***]	23.06 [3.77***]
Adjusted (OLS) or Pseudo (Tobit) R-squared	.18	.09	.12	.08

T-statistics are in brackets.

Statistical significance: *10 percent **5 percent ***1 percent

One other factor appears to have a noticeable effect on an agency's *Scorecard* score but could not easily be quantified to put into regressions: the extent to which the agency developed a strategic plan with outcome-oriented goals and performance measures. Over the years, various agencies' scores have risen or fallen substantially when the outcome orientation of their underlying strategic plans changed.

Several prominent examples occurred in fiscal year 2007. Two departments—HHS and Homeland Security—improved their scores considerably. HHS’s fiscal 2007 strategic goals covered some of the same topics as the fiscal 2006 goals, but they indicated more clearly the outcomes HHS was trying to achieve. HHS articulated no strategic objectives in fiscal 2006, but in fiscal 2007 there were 16 strategic objectives, most of which were outcome-oriented. Two-thirds of performance measures were outcome-oriented in fiscal 2007, compared to only one-third in fiscal 2006. All of Homeland Security’s fiscal 2007 strategic goals and objectives were expressed as outcomes. A majority of the fiscal 2007 performance goals were outcome-oriented, compared to one-fifth in fiscal 2006.

Two other departments—State and USAID—saw their scores fall significantly. Their fiscal 2006 strategic goals identified many specific outcomes; their fiscal 2007 strategic goals read more like a statement of principles. The Mercatus research team could not find performance goals for either agency in fiscal 2007. Fewer measures for fiscal 2007 were related to intermediate or final outcomes (McTigue et. al. 2008, 35–39).

Why some agencies chose to produce higher quality strategic plans or GPRA reports than others did is something of a mystery. But there is little evidence that any agency has an inherent advantage due to its size or program structure.

2. Availability and Use of Performance Information in Agencies

One of GPRA’s major purposes was to improve federal managers’ ability to manage by providing them with information about results. Numerous scholars suggest that this could be one of GPRA’s most important contributions (Frederickson and Frederickson 2006, 185; Hatry et. al 2005, 200; Joyce 2005). Periodic Government Accountability Office (GAO) surveys suggest that the availability and use of performance information in federal agencies has improved since GPRA. Mercatus Scorecard scores are also correlated with GAO survey results on the availability and use of performance information. Thus, the available evidence suggests that GPRA has indeed improved the availability and use of performance information in some federal agencies.

Baseline Indicators

In 1997, GAO began surveying individual managers on the availability and use of performance information for the programs and activities for which they are responsible. The surveys went to managers in agencies covered by the Chief Financial Officers’ Act. These agencies account for the vast majority of federal spending, and they are the same agencies covered by the Mercatus Center’s Scorecard. The GAO surveys ask whether managers have and use five types of performance measures related to GPRA:

- Outcomes – Direct results achieved through the provision of goods and services by your organization
- Outputs – Products or services produced, distributed, or provided to service population
- Efficiency – Cost per unit, productivity measures, ratios of direct to indirect costs, etc.
- Customer satisfaction – Measures of quality and timeliness from external sources

- Quality – Measures of quality from internal sources

The 1997 survey asked managers to recall whether they had and used performance measures in 1994, the year after passage of GPRA. Responses to these questions, shown in Table 6, indicate that availability and use of performance information was not widespread one year after the passage of GPRA. The percentages in the table (and in subsequent figures and tables) are the percentages of managers who said they have or use these measures “to a great extent” or “to a very great extent.”

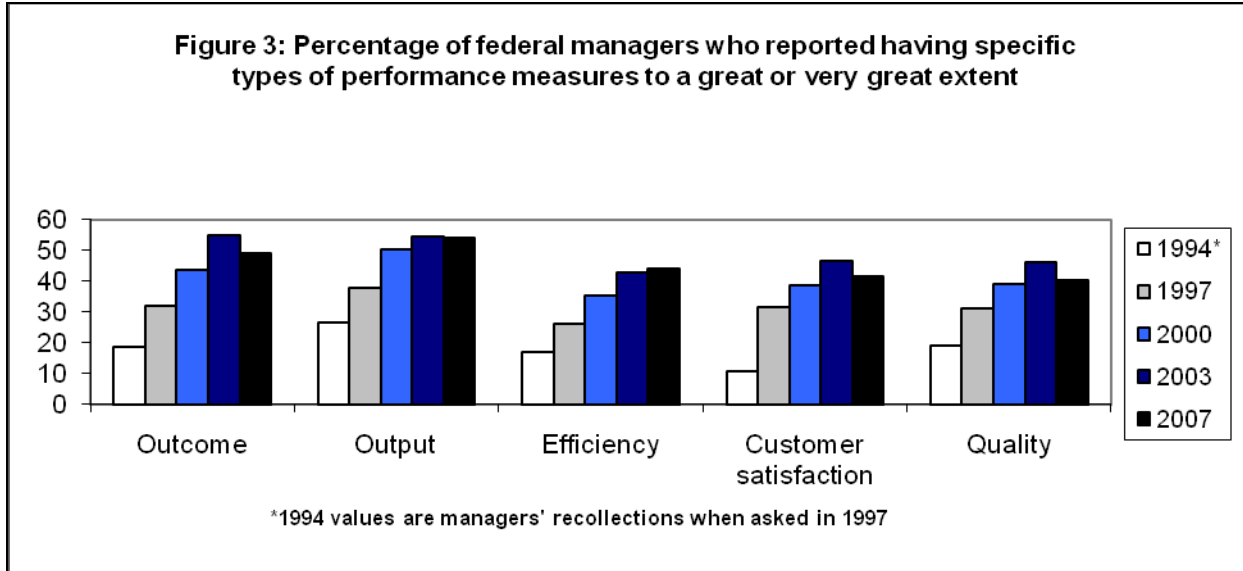
Table 6: Availability and Use of Performance Measures, 1994

Availability of performance measures	
Outcome	18.6
Output	26.6
Efficiency	16.9
Customer satisfaction	10.6
Quality	18.9
Uses of performance information from their programs	
Develop agency budget	16.2
Make funding decisions for the program	14.1
Make changes to the program by managers above my level	8.7

Source: GAO (1997).

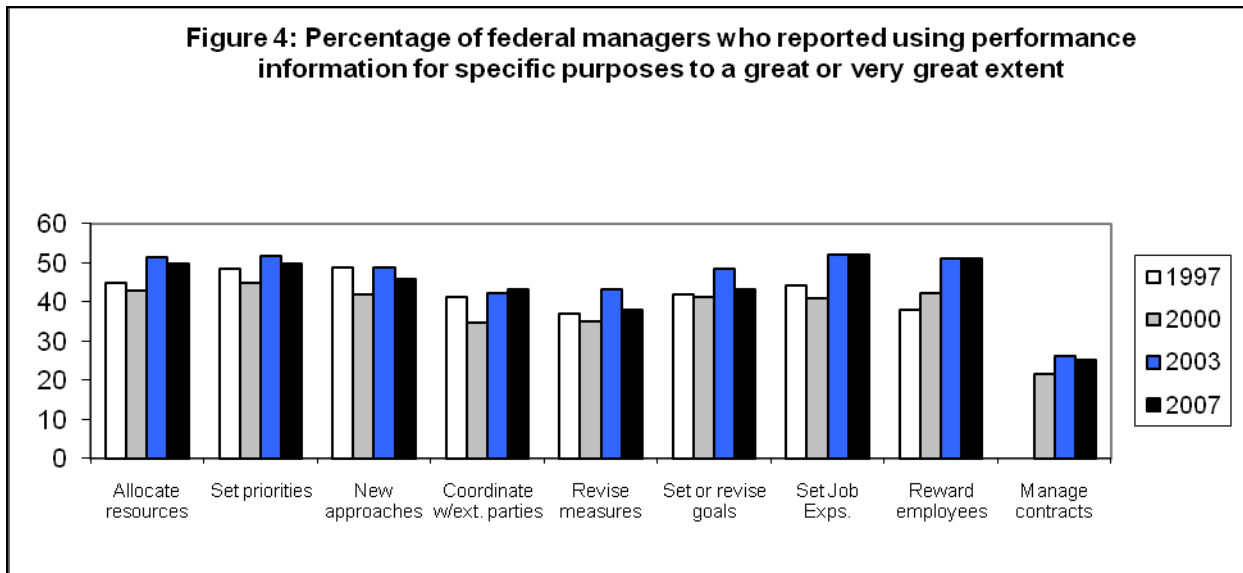
Measuring Improvement Over Time

GAO continued to survey federal managers every several years since 1997. Figure 3 shows the percentage of federal managers who said they had various types of performance measures in 1994, 1997, 2000, 2003, and 2007. The latter years show marked improvements from 1994 and 1997. For example, only 18.6 percent of managers said they had outcome measures for their programs in 1994; the figure rose to 31.8 percent in 1997. By 2003, 55 percent of managers said they had outcome measures for their programs. The number receded to 48.9 percent in 2007—still well above its level in either 1994 or 1997. For each type of performance measure, differences between 1997 and 2007 are statistically significant (GAO 2008, 4).



Source: Author's calculations based on data in GAO (1997, 2004a) and spreadsheets furnished by GAO for 2000 and 2007.

Figure 4 shows less-sanguine results for uses of performance information. The average percentage of managers who said they use performance information to a great or very great extent increased slightly, but not by much. Use of performance information to allocate resources, for instance, increased by just 5 percentage points between 1997 and 2007, from 44.8 percent to 49.8 percent. The largest increase occurred in the use of performance information to reward employees who report to the manager, which rose from 38 percent in 1997 to 51.1 percent in 2007. This is the only improvement in the use of performance information that was statistically significant (GAO 2008, 6).



Source: Author's calculations based on data in GAO (1997, 2004a) and spreadsheets furnished by GAO for 2000 and 2007.

Survey results for individual agencies vary widely. Results on availability and use of performance information change if we calculate an average response for managers in each agency and then average the agency responses. By counting each agency's average equally, this method implicitly gives more weight to responses from smaller agencies with fewer managers. But it helps us identify whether an appreciable number of agencies have experienced improvements.

For two years—2000 and 2007—GAO surveyed a large enough sample of managers to calculate valid average responses for each individual agency. Table 7 shows averaged agency responses in 2000 and 2007.³ Availability and use of performance information improved for every type of performance measure and every type of use. Improvements in the availability of outcome and efficiency measures are highly statistically significant; improvements in the other uses are marginally significant. Improvements in almost all of the uses of performance information are highly statistically significant.

Table 7: Agency Average Change in Percent of Managers Who Have and Use Performance Information

	2000	2007	Difference	T-statistic
Availability of performance measures				
Outcome	45.5	53.9	8.4	2.80***
Output	53.4	59.4	6.0	1.92*
Efficiency	36.2	45.2	9.0	3.34***
Customer satisfaction	36.6	43.2	6.6	1.97*
Quality	36.9	43.1	6.2	1.98*
Uses of performance information in their programs				
Allocate resources	45.6	50.5	4.9	1.99*
Set priorities	46.4	53.2	6.7	2.48***
Adopting new approaches/work processes	42.5	51.3	8.8	3.18***
Coordinate with external parties	35.7	45.4	9.7	4.17***
Refine program performance measures	38.2	44.5	6.3	2.37***
Set or revising performance goals	43.3	50.2	6.9	2.67***
Set job expectations for employees I manage	42.8	55.0	12.3	4.80***
Reward employees I manage or supervise	43.6	52.7	9.1	3.19***
Manage contracts	24.2	28.8	4.5	1.51

Source: Author's calculations based on spreadsheets for 2000 and 2007 furnished by GAO. Statistical significance levels: ***1 percent **5 percent *10 percent

Did GPRA Play a Role?

In another paper (Ellig 2010), I correlated the quality of agency GPRA reports (measured by the agency's *Mercatus Scorecard* score) with the availability and use of performance information (measured by the GAO survey responses). The regressions included variables controlling for

³ The GAO survey covers all 24 CFO Act agencies, plus separate breakouts for the Centers for Medicare and Medicaid Services (HHS), Federal Emergency Management Agency (DHS), Federal Aviation Administration (DOT), Forest Service (USDA), and Internal Revenue Service (Treasury). I included the responses from managers in those five sub-components in the average for their parent departments.

agency leadership's perceived commitment to performance management, agency size, mix of program types in the agency, complexity of the agency's missions, agency ideology, and elected officials' interest in performance management as perceived by agency managers. Because GAO survey results by agency are available only for 2000 and 2007, the sample size was quite small (46 observations). The results should be taken as suggestive rather than definitive.

The regression coefficients indicate that a 1 point increase in *Scorecard* score is usually associated with a 0.3–0.65 percentage point increase in managers reporting that they have or use performance information for various purposes. Compared to an agency that produces no GPRA report, an agency receiving the average *Scorecard* score of 34 for the years covered in the study would have 10–16 percentage points more managers responding that they have the types of performance measures listed in Table 8. An agency with an average report would have 10–22 percentage points more managers saying that they use performance information for various purposes. Between 40 and 57 percent of managers said they had the types of performance measures listed in Table 8 or used performance information for purposes listed in the table. Therefore, merely producing an average GPRA report appears to make a relatively big contribution to the availability and use of performance information.

One other significant caveat should accompany these results. The GAO surveys do not link the use of performance information with actual improvement in results. Therefore, we do not know whether, or to what extent, the increased availability and use of performance information has improved “program efficiency and effectiveness,” as GPRA intended.

3. GPRA and Budgeting

One of GPRA's other major purposes was to improve congressional budgeting decisions by providing better information about program results. The legislation mentions three possible means of increasing the availability of this information to Congress. First, the measures in annual performance reports should show agency progress in achieving their strategic goals. Second, GPRA requires agencies to list program evaluations in their annual performance reports, which should help point legislators toward detailed evaluations of individual programs. Third, the legislation established pilot programs on “performance budgeting.” Performance budgeting matches proposed expenditures with outcomes and shows how the amount of outcome is expected to vary with changes in the level of spending.

Over time, one would expect that effective congressional use of outcome information would lead to reallocation of expenditures from programs that do not produce results to programs that do. As several scholars note, this is not a rigid, automatic process. Some programs may fail to produce results because they are poorly structured, have vague goals, or receive insufficient funding (Moynihan 2008, 127–29; Joyce 2005, 93–94). Fixing those problems could transform some ineffective programs into effective ones. Nevertheless, one would expect that better outcome information would lead legislators to terminate or shrink at least some programs in order to reallocate resources to more effective ones that seek to achieve similar goals. Certainly some OMB staff hoped this would happen (Moynihan 2008, 128).

GPRA could affect budget decisions in three ways: by altering agency budget recommendations and management of resources after they receive their budgets, by altering the president's budget recommendations to Congress, or by altering actual budget decisions made by Congress (Joyce 2005, 96).

Agency Budget Recommendations and Budget Execution

The GAO surveys of federal managers ask several questions about uses of performance information that appear related either to budget recommendations or to “budget execution”—that is, allocation and management of financial resources where the agency has discretion. The agency averages in Table 8 above show that between 2000 and 2007, about 5 percent more managers said they use performance information to allocate resources in their programs, and about 7 percent more said that they use performance information to set priorities. These are not huge changes, but they indicate some progress.

Case studies suggest that performance budgeting has become more widespread in federal agencies and affects managers' decisions. In 2005, GAO conducted a series of performance budgeting case studies. Officials at OMB, NASA, EPA, and Veterans Affairs all said that consolidating budget requests based on strategic goals and performance measures had improved coordination among different parts of agencies that had to work together to accomplish the strategic goal. “OMB staff explained that there is more coordination among EPA's program offices because programs that support common goals and objectives have to ‘sell’ themselves together under the new planning and budget structure” (GAO 2005a, 68). Commerce and EPA told GAO that they used performance budgets for internal management even though Congress continued to appropriate funds to individual programs rather than strategic goals or outcomes (GAO 2005a, 87–88).

Another possible indicator of agencies' progress in using performance information for budgeting would be changes in agency evaluations on “Budget and Performance Integration,” one of five management priorities in the “President's Management Agenda” articulated by the G.W. Bush administration in 2001. The goal of budget and performance integration was succinctly stated: “Over time, agencies will be expected to identify high quality outcome measures, accurately monitor the performance of programs, and begin integrating this presentation with associated cost. Using this information, high performing programs will be reinforced and non-performing activities reformed or terminated” (OMB 2001a, 29). The President's Management Council, in consultation with experts in government and academia, developed a set of standards for evaluating agencies' success in budget and performance integration. OMB issued a quarterly scorecard indicating each agency's achievement and progress, using color codes of red (unsatisfactory), yellow (mixed results), and green (success). A 2001 baseline evaluation of 26 federal agencies that account for virtually all federal spending awarded just three agencies with yellow for mixed results; none achieved green (OMB 2001b). By the end of 2008, 19 agencies had “gotten to green” on budget and performance integration (now renamed “performance improvement”); the remainder were rated yellow (OMB 2008).

Presidential Budget Recommendations

In addition to requiring agencies to produce performance budgets, the G.W. Bush administration undertook another systematic initiative intended to link performance information with budget recommendations. In February 2003, the administration released its system for reviewing the performance of most federal programs, called the Program Assessment Rating Tool (PART).

PART was a framework “used to evaluate a program’s purpose, design, planning, management, results, and accountability to determine its overall effectiveness.” It was also intended to help OMB and Congress make performance budgeting decisions. PART questionnaires contained questions divided into four categories: program purpose and design, strategic planning, management, and results. Each section received a score between 0 and 25 points. The program’s total score was a weighted average of the four scores: purpose and design (20 percent), strategic planning (10 percent), management (20 percent), and results (50 percent). If information on results was available, a program could be rated Effective (85 points and above), Moderately Effective (70–84 points), Adequate (50–69 points), or Ineffective (0–49 points). Regardless of the numerical score, a program could also be rated “Results Not Demonstrated” if it had not established goals and measures and collected data to evaluate performance (OMB undated).

PART sought to link measurement of program results with GPRA’s requirements for measurement of the agency’s overall results. OMB Circular A-11 instructed agencies to use the same performance measures for GPRA and PART when plans and reports include programs that have been PARTed (Brito and Ellig 2009, 39).

A GAO analysis of the first year of PART data found that PART scores were positively correlated with the recommended funding changes in the president’s fiscal 2004 budget, but only for small “discretionary” programs—the programs that require a congressional appropriation decision each year. A one point increase in a small discretionary program’s PART score was associated with a 1.07 percent recommended funding increase in the president’s budget. However, PART scores explained only about 15 percent of the variation in the president’s budget requests; other factors likely had a larger impact (GAO 2004b, 42–46).

Gilmour and Lewis (2006a) found that the effect of PART scores on administration budget proposals for fiscal 2004 depended on the political orientation of the program’s department. The administration proposed larger budget increases for programs with higher PART scores in “Democratic” departments, but PART scores had either no effect or a negative effect on recommended funding in “Republican” departments. For fiscal 2005, however, Gilmour and Lewis (2006b) found that PART scores had a positive, statistically significant effect on budget recommendations, and political factors had little effect. A one-point increase in the PART score was correlated with a 0.40–0.47 percent increase in recommended funding. Consistent with GAO (2004b), the authors found that the effect was concentrated in small programs, where a one-point increase in the PART score was associated with a 1.28 percent increase in recommended budget.

Norcross (2005) and Norcross and McKenzie (2006) examined the relationship between PART ratings and presidential budget requests for fiscal years 2006 and 2007. Table 8, drawn from these studies, shows the percent of programs with various ratings that the president’s budget

recommended for funding increases, decreases, or no change. The president's budget usually recommended funding reductions for programs rated ineffective. Effective programs were most likely to get funding increase recommendations, followed by moderately effective and then adequate programs. However, the table also reveals that the relationship between PART ratings and funding recommendations was far from automatic or mechanical. Funding increases were recommended for more than one-third of effective, moderately effective, and adequate programs. Since the results not demonstrated programs were programs for which insufficient information about results was available, clearly something other than PART ratings affected the budget recommendations for those programs.

**Table 8: PART Ratings and Funding Changes
Recommended in the President's Budget**

	Ineffective	Adequate	Moderately Effective	Effective	Results Not Demonstrated
Fiscal 2006					
Increase	5%	43%	51%	61%	30%
No change	9%	22%	11%	4%	28%
Decrease	86%	36%	38%	35%	41%
Fiscal 2007					
Increase	11%	37%	56%	61%	26%
No change	14%	21%	13%	10%	31%
Decrease	75%	42%	31%	28%	42%

Source: Norcross (2005, 19); Norcross and McKenzie (2006, 22).

Proposed program terminations provide another way to search for links between PART ratings and presidential budget recommendations. Thirty-two of the 99 programs proposed for termination in the fiscal 2006 budget had undergone PART reviews. Ten of these programs were rated ineffective, six were rated adequate, and 16 were results not demonstrated (Norcross 2005, 19–20). In the fiscal 2007 budget, seven programs proposed for termination were ineffective, eight were adequate, two were moderately effective, and 15 were results not demonstrated (Norcross and McKenzie 2006, 24). For fiscal 2008, 37 programs recommended for termination had been PARTed. Five were ineffective, six were adequate, four were moderately effective, and 22 were results not demonstrated (Norcross and Adamson 2007, 29). Thus, most programs recommended for termination were either ineffective or results not demonstrated. The administration did not recommend terminating any effective programs. For all three years, the budget's *Major Savings and Reforms* document claims that PART ratings were one factor affecting many of the termination recommendations, but not the only factor.

Thus, the available studies suggest that PART ratings affected the G.W. Bush administration's budget recommendations to some extent, though not to a great degree.

Congressional Budget Decisions

In general, Congress has displayed less interest in performance-based budgeting than the president. In the early years of GPRA, House Majority Leader Dick Armey took an active role in

assessing agencies' strategic plans and performance plans. Apparently few appropriators or committee chairs shared his enthusiasm. In her study of GPRA implementation by the Department of Transportation, for example, Curristine (2002, 42) notes, "Indications from interviews with appropriators show that they will not use performance measures in making funding decisions on highways."

The congressional response to the administration's attempt to reformat agency budget justifications submitted to Congress for fiscal 2004 and 2005 is instructive. In the late 1990s, OMB began to discuss the need to restructure budget accounts to better align resources with results. Some agencies began experimenting with performance budgets. In July 2003, OMB directed agencies to develop performance budgets that integrated their GPRA-mandated annual performance plans into their congressional budget justifications beginning with fiscal year 2005.

A major goal of this change was to better link costs with information about goals and outcomes. Federal appropriations accounts and programs do not necessarily match up with agency strategic goals, measures, or outcomes. Some performance goals cut across multiple accounts or programs, and appropriations for an individual program may not measure the full cost of achieving that program's goals (GAO 2005a, 24). OMB also sought to restructure appropriations accounts so that managers responsible for particular outcomes would have greater ability to reallocate resources to achieve those outcomes, since managers normally lack authority to reallocate funds between appropriations accounts (GAO 2005a, 43; 72–73).

Congress rejected most of the performance budgeting formats. Moynihan (2008, 124) quotes one OMB budget examiner on the congressional response:

The good government types and the government oversight committees are supportive, but they have very little clout. The appropriations committees have been much less enthusiastic. My committee was outright hostile. They think that the performance information in the budget produced a lot of paper but nothing they found useful.

Extensive GAO case studies revealed that committees usually preferred to use information organized by program and categories of expenditures rather than strategic goals:

Congressional appropriations subcommittee staff for the most part continued to state a preference for and rely on previously established budget structures. Appropriations subcommittees and staff said that the changes in budget accounts and presentations shifted the focus away from programs and items of expenditures of interest to congressional appropriators and instead highlighted strategic and performance goals. While these staff expressed general support for budget and performance integration, they objected to changes that replaced information, such as workload and output measures, traditionally used for congressional appropriations and oversight with the new performance perspective (GAO 2005a, 7).

The Environmental Protection Agency (EPA) had structured its budget requests around strategic goals since fiscal 1999. Nevertheless, Congress required the EPA to continue to break budget requests down by program as well, and this is the information Congress used to make appropriations. Appropriations subcommittee staff generally did not use the performance-based budget to conduct their work but rather the program-based information they requested from EPA (GAO 2005a, 94). In 2004, the House and Senate appropriations subcommittees requested that the EPA reformat its budget justification using appropriations accounts and programs rather than strategic goals (GAO 2005a, 14–17, 78).

Similarly, appropriations committee staff did not use the performance-based information from the Labor Department, but consulted budget justifications from earlier years and requested supplementary information (GAO 2005a, 95). In 2004, the House Appropriations Committee directed Veterans Affairs “to refrain from incorporating ‘performance-based’ budget documents in the 2005 budget justification submitted to the Committee, but keep the Performance Plan as a separate volume.” When the department submitted a restructured performance-based budget for 2005, the committee responded, “If the Department wishes to continue the wasteful practice of submitting a budget structure that will not serve the needs of the Congress, the Congress has little choice but to reject that structure and continue providing appropriations that serve its purposes” (GAO 2005a, 79). The committee directed HUD “not to submit or otherwise incorporate the strategic planning document or its structure into its fiscal year 2005 Budget Justification submission to the Committee” (GAO 2005a, 80). After telling the Departments of Transportation, Treasury, and independent agencies to revert to the traditional budget justification format, the committee warned, “If the Office of Management and Budget or individual agencies do not heed the Committee’s direction, the Committee will assume that individual budget offices have excess resources that can be applied to other, more critical missions” (Moynihan 2008, 123).

Similarly, the Senate Appropriations Committee told the Labor Department that it should use performance information for management purposes but should submit its budget requests in the traditional appropriations format rather than a performance budget format (GAO 2005a, 81). Congress accepted only NASA’s proposed revisions to its appropriations accounts (GAO 2005a, 78).

Different committee staff cited different reasons for rejecting the administration’s performance-based budget formats. Committees often preferred to appropriate funds by functional area or program, sometimes disagreed with the agency’s strategic goals, expressed concern that strategic goals would change when the agency’s strategic plan changed, and questioned whether some agencies could track expenditures by strategic goal. Some staff noted that the new format omitted some useful information, such as unit cost, workload, and output measures; historical spending trends; and funding levels broken down by program or state. Some said there was too much performance information, too much narrative, or that the information was poorly organized and formatted (GAO 2005a, 81–85).

Some of these reasons suggest that the committees simply did not want to appropriate funds based on outcomes. Others imply that committees saw the prospective value of performance

budgeting but did not think the agencies' performance budgets provided the right information in the right way.

A content analysis of appropriations documents suggests Congress had little interest in using performance information at this time. Moynihan (2008) examined appropriations bills, accompanying conference reports, and oversight and appropriations hearings in search of performance discussions. In 3,257 single-spaced pages of text, he found that "performance" was mentioned just 57 times in reference to expected or projected program performance, 21 times when committees urged agencies to use performance information, 109 times when legislators asked agencies for more data, and 47 times in reference to actual program achievement. Only nine of these latter instances involved citation of quantitative performance indicators. "The documents examined show no discussions of legislators using the [performance] information themselves" (Moynihan 2008, 131–33).

Another possible indicator of congressional receptivity to performance budgeting would be congressional reaction to PART. A 2005 GAO study cited several examples of committee hearings or legislation that related to PART. Nevertheless, GAO (2005b, 49) concluded, "Despite its efforts, OMB has had limited success in engaging Congress in the PART process."

Even if Congress did not use PART, it may have acquiesced in presidential recommendations that were informed by PART. Table 9 shows that congressional budget decisions on PARTed programs were often consistent with the president's proposals, with two exceptions. First, Congress was much less likely than the president to cut funding for ineffective programs and much more likely to increase funding for these programs. Second, for fiscal 2007, Congress increased funding for a much smaller percentage of the effective programs that the president recommended, and decreased funding for a much larger percentage.

Table 9: PART Ratings and Funding Changes Approved by Congress

	Ineffective	Adequate	Moderately Effective	Effective	Results Not Demonstrated
Fiscal 2006					
Increase	18%	47%	53%	59%	34%
No change	4%	13%	12%	5%	25%
Decrease	79%	39%	35%	36%	42%
Fiscal 2007					
Increase	30%	33%	43%	48%	24%
No change	41%	29%	23%	14%	42%
Decrease	30%	38%	34%	38%	33%

Source: Norcross and Adamson (2007, 28); Norcross and McKenzie (2006, 23)

There is, however, little evidence that Congress considered PART ratings when it made these decisions. An analysis of committee reports in the 109th Congress, which approved the fiscal year 2006 and 2007 budgets, revealed that only about 6 percent of them had PART-related content, which leads the authors to conclude that Congress used PART "on a limited basis." One

subcommittee even banned departments under its jurisdiction from including PART information in its fiscal 2008 budget submission (Frisco and Stalebrink 2008, 16).

Why the Divergence?

The available studies suggest that performance information has had some influence on presidential budget recommendations and agency management, but very little influence on congressional budgeting decisions. The relative lack of congressional interest can be explained by the different political incentives congressional appropriations committees and the president face to monitor agency outputs, efficiency, and outcomes.

Members of the U.S. Congress represent particular geographical constituencies. The appropriations committees and subcommittees make most budget decisions. Committee members tend to be “high demanders” of the services provided by the agencies over which the committee has jurisdiction (Niskanen 1994 [1975], 250–51). High demand members may have many constituents who benefit from the agencies’ program outcomes (such as, for example, a large number of blue collar factory workers who need vocational retraining). Alternatively, a high demand member may have constituents who benefit greatly from the expenditures even if they are not direct consumers of the services (such as employees of a large military base in the member’s district).

In both cases, the member’s district or state receives concentrated benefits from the expenditures or outcomes, while the entire nation pays the costs of those particular programs. “Each legislator will want to procure a project that is larger than optimal because he or she no longer internalizes the full marginal cost of the project. An example from everyday life is the tendency of restaurant bills that are split equally among diners (by previous agreement or norm) to be larger in total than separate checks would have been” (Primo 2007, 44). Legislators and voters as a whole, however, would be better off if the government funded only those programs that produced benefits to the entire nation that exceed the costs to the entire nation, and if those programs operated at maximum efficiency and effectiveness.

To the extent that their constituents benefit from program outcomes, committee members have some reason to monitor agencies’ performance. But they also face two countervailing incentives that discourage them from monitoring. First, some of their constituents may profit personally from agencies’ inefficient or less effective expenditures. Appropriations for weapons systems the Defense Department says it doesn’t need are an extreme example of this type of expenditure. Improved accountability for outcomes would likely reduce or eliminate those kinds of expenditures, thus reducing benefits that flow to some individual districts. Second, individual members of Congress must decide how to divide their time and staff resources between activities that benefit their own constituencies almost exclusively (such as answering mail, speaking at community events, and helping constituents get federal money) and activities whose benefits are spread across the entire nation. Monitoring the efficiency and effectiveness of federal programs is often a good example of the latter activity (Niskanen 1994 [1975], 251–54). Thus, committee members will likely devote less time and effort to monitoring the efficiency and effectiveness of government programs than the typical or “median” voter would like.

The president, on the other hand, is elected by the entire nation. To win in a two-party system, the president usually has to appeal to the median voters. Therefore, the president has a stronger incentive than the members of appropriations committees to reflect the preferences of the median voter (Niskanen 1994 [1971], 227). Given these differing political incentives, it is no surprise that the U.S. executive branch shows more interest than congressional appropriations committees in using performance information to make budget decisions.

At the outset of GPRA, many hoped that transparent disclosure of performance information would itself increase the political benefits and reduce the costs of monitoring agencies' performance. Annual performance reports would make voters more aware of government performance and reduce congressional monitoring costs. With some segment of voters better informed and more vigilant, members of Congress would find that improved efficiency and effectiveness of programs attracts votes. Thus far, this has not happened to any great extent.

Another possible way to increase congressional focus on efficiency and effectiveness would be process reforms that promote budget decisions based on benefits and costs to the nation as a whole rather than just benefits and costs to individual members' constituencies. Several options include:

- **Budget amendment rules.** Primo (2007) finds that enforceable spending limits prompt “agenda setters” to propose funding for programs that are more efficient, or at least less inefficient, in the sense that they better balance cost to the nation with benefits to the nation. A budget rule that allows spending bills complying with a pre-set spending ceiling to proceed to a floor vote without amendments would likely induce committees to make more efficient spending recommendations (Primo 2007, 74–81).
- **Supermajority requirements.** Scholars have suggested that supermajority requirements to pass appropriations bills could, at least in some cases, make it harder for legislators to enact budgets that contain many inefficient programs, since they need to get “buy-in” from a larger number of representatives whose constituents pay the costs (Primo 2007, 53–54; Niskanen 1994 [1971], 227–28).
- **Spending and performance commission.** Brito (2010) suggests a commission to examine and terminate ineffective discretionary spending programs modeled on the Base Realignment and Closure Commission. In 1988, Congress created this commission of independent experts to identify military bases for closure based on military need. The commission's recommendations were implemented unless Congress voted to disapprove the entire list of base closures. A spending commission, composed of independent experts examining programs according to performance-based criteria specified by Congress, would issue recommendations that became operative unless Congress approved a joint resolution of disapproval. This process would allow legislators to vote in favor of performance-based budgeting without having to explicitly vote against individual programs that may be politically popular even if they are ineffective.

A skeptical reader might interpret these kinds of recommendations as mere statements of ideological preference for smaller government. That interpretation misses the point entirely. Like

diners who order too many bottles of wine because they're splitting the check equally, legislators are led by current institutional incentives (as if by an invisible hand!) to approve more programs that are less efficient than either they or their constituents would prefer in more sober moments under different rules. Changing the "rules of the game" to focus legislators more on efficiency and effectiveness and less on bringing home rewards to their individual constituencies could make legislators and the public better off by prompting legislators to choose a mix of programs and spending that gives the public greater value for its money.

Conclusion

GPRA was intended to improve agency management and congressional budget decisions by improving the quality of government performance information. The Mercatus *Scorecard* project clearly indicates that the quality of performance information has improved. GAO surveys, in conjunction with *Scorecard* data, show that GPRA has also improved the availability and use of performance information by federal managers. The G.W. Bush administration undertook a major effort to integrate performance with budget information and use the former to inform the latter. But Congress rarely used GPRA-oriented performance information in budgeting.

Individual members of Congress win reelection by bringing federal expenditures to constituencies in their districts or states, while the costs are shared among all taxpayers in the nation. They face weaker incentives to monitor programs for efficiency or effectiveness, since these benefits are often not concentrated on constituents in specific states or districts but rather are shared with program beneficiaries and taxpayers nationwide. Given this reality, it is perhaps not surprising that appropriations debates focus more on distribution of expenditures than on effectiveness or the benefit-cost analysis of expenditures.

No one has yet invented a magic pill that turns politicians into statesmen. But the hard reality of resource constraints may yet counteract the customary political incentives to treat constituents' receipt of expenditures as the main "outcome" that matters.

The U.S. federal budget deficit skyrocketed from 3.2 percent of GDP in 2008 to 12.3 percent of GDP in 2009 and a projected 8 percent of GDP in 2010. Federal debt increased from 37.2 percent of GDP in 2008 to 48.8 percent in 2009, and it is projected to hit 64.6 percent of GDP in 2010 (OMB 2010, 114). Scholars point out that U.S. states tend to balance their budgets because market-based debt ratings limit their ability to borrow (Primo 2007, 128). In February 2010, the investment analysis firm Moody's suggested that the U.S. government's AAA bond rating might be in jeopardy (Burns 2010). Increased Social Security, Medicare, Medicaid, and interest payments will create growing pressure to cut other spending and increase taxes to record peacetime levels. Surely at some point, reforming or reallocating spending away from programs that perform poorly will become a politically attractive option.

References

- Brito, Jerry. 2010. "Running for Cover: The BRAC Commission as a Model for Federal Spending Reform." Working Paper, Mercatus Center at George Mason University, available at <http://mercatus.org/publication/running-cover>.
- Brito, Jerry, and Jerry Ellig. 2010. "Toward a More Perfect Union: Regulatory Analysis and Performance Management," *Florida State University Business Review* 8:1 (Spring-Summer), 1-55.
- Burns, Judith. 2010. "Moving the Market: Geithner on Defense of US Bond Rating," *Wall Street Journal* (Feb. 8), C2
- Curristine, Teresa. 2002. "Reforming the U.S. Department of Transportation: Challenges and Opportunities of the Government performance and Results Act for Federal-State Relations," *Publius* 32, 25-44.
- Downs, Anthony. 1967. *Inside Bureaucracy*. Boston: Little, Brown and Co.
- Ellig, Jerry. 2010. "Institutions, Incentives, and Performance Information in the US Government," unpublished manuscript.
- Frederickson, David G., and H. George Frederickson. 2006. *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* (Summer), 1-19.
- Gilmour, John B., and David E. Lewis. 2006a. "Does Performance Budgeting Work? An Examination of the Office of Management and Budget's PART Scores," *Public Administration Review* (Sept.-Oct.), 742-52.
- _____. 2006b. "Assessing Performance Assessment for Budgeting: The Influence of Politics, Performance, and Program Size in Fiscal Year 2005," *Journal of Public Administration Research and Theory*, 1-18.
- Government Accountability Office. 2008. "Government Performance: Lessons Learned for the Next Administration on Using Performance Information to Improve Results," Statement of Bernice Steinhardt, GAO-08-1026T (July 24).
- _____. 2005. "Managing for Results: Enhancing the Use of Agency Performance Information for Management Decision Making," Report GAO-05-927 (September).
- _____. 2005a. "Performance Budgeting: Efforts to Restructure Budgets to Better Align Resources with Performance," Report GAO-05-117SP (January).

_____. 2005b. "Performance Budgeting: PART Focuses Attention on Program Performance, but More Can Be Done to Engage Congress," Report GAO-06-28 (October).

_____. 2004a. "Results-Oriented Government: GPRA Has Established a Solid Foundation for Achieving Greater Results," Report GAO-04-38 (March).

_____. 2004b. "Performance Budgeting: Observations on the Use of OMB's Program Assessment Rating Tool for the Fiscal Year 2004 Budget," Report GAO-04-174 (January).

_____. 2001. "Managing for Results: Federal Managers' Views on Key Management Issues Vary Widely Across Agencies," Report GAO-01-592 (May).

_____. 1997. "The Government Performance and Results Act: 1997 Governmentwide Implementation Will Be Uneven," Report GAO/GGD-97-109 (June).

_____. 1992. "Program Performance Measures: Federal Agency Collection and Use of Performance Data," Report GAO/GGD-92-65 (May).

Hatry, Harry P., Elaine Morley, Shelli B. Rossman, and Joseph S. Wholey. 2005. "How Federal Programs Use Outcome Information: Opportunities for Federal Managers," in John M. Kamensky and Albert Morales (eds.), *Managing for Results 2005*. New York: Rowman & Littlefield Publishers, 197-274.

Joyce, Philip G. 2007. "Linking Performance and Budgeting: Opportunities in the Federal Budget Process," in Jonathan D. Breul and Carl Moravitz (Eds.), *Integrating Performance and Budgets*. Lanham, MD: Rowman & Littlefield, 2007, 19-70.

Kettl, Donald F. 1988. *Government by Proxy*. Washington, DC: Congressional Quarterly.

McTigue, Maurice, Henry Wray, and Jerry Ellig. 2009. *10th Annual Performance Report Scorecard*. Arlington, VA: Mercatus Center, George Mason University.

_____. 2008. *9th Annual Performance Report Scorecard*. Arlington, VA: Mercatus Center, George Mason University.

Metzenbaum, Shelley H. 2005. "Strategies for Using State Information: Measuring and Improving Program Performance," in John M. Kamensky and Albert Morales (eds.), *Managing for Results 2005*. New York: Rowman & Littlefield Publishers, 277-350.

Moynihan, Donald P. 2008. *The Dynamics of Performance Management*. Washington, DC: Georgetown University Press.

Niskanen, William A., Jr. 1971. *Bureaucracy and Representative Government*. Aldine-Atherton.

_____, 1994 [1975]. "Bureaucrats and Politicians," reprinted in William A. Niskanen, Jr., *Bureaucracy and Public Economics*. Brookfield, VT: Edward Elgar.

_____. 1994. *Bureaucracy and Public Economics*. Brookfield, VT: Edward Elgar.

Norcross, Eileen. 2005. "An Analysis of the Office of Management and Budget's Program Assessment Rating Tool," Working Paper, Mercatus Center at George Mason University, available at <http://mercatus.org/publication/analysis-office-management-and-budgets-program-assessment-rating-tool-fy06>.

Norcross, Eileen, and Kyle McKenzie. 2006. "An Analysis of the Office of Management and Budget's Program Assessment Rating Tool for Fiscal Year 2007," Working Paper, Mercatus Center at George Mason University, available at <http://mercatus.org/publication/analysis-office-management-and-budgets-program-assessment-rating-tool-fy-2007>.

Norcross, Eileen, and Joseph Adamson. 2007. "An Analysis of the Office of Management and Budget's Program Assessment Rating Tool for Fiscal Year 2008," Working Paper, Mercatus Center at George Mason University, available at <http://mercatus.org/publication/analysis-office-management-and-budgets-program-assessment-rating-tool-fy-2008>.

Office of Management and Budget. Undated. Program Assessment Rating Tool description, available at www.expectmore.gov.

_____. 2010. *Budget of the United States Government, Fiscal Year 2011*.

_____. 2008. "Executive Branch Management Scorecard."

_____. 2001a. *The President's Management Agenda, Fiscal Year 2002*, available at <http://www.whitehouse.gov/omb/assets/omb/budget/fy2002/mgmt.pdf>.

_____. 2001b. "Executive Branch Management Scorecard Baseline."

Primo, David M. 2007. *Rules and Restraint: Government Spending and the Design of Institutions*. Chicago: University of Chicago Press.

Radin, Beryl A. 2006. *Challenging the Performance Movement*. Washington, DC: Georgetown University Press.

Tullock, Gordon. 2005 [1965]. *The Politics of Bureaucracy*, reprinted in Charles K. Rowley, Ed., *The Selected Works of Gordon Tullock*. Indianapolis, IN: Liberty Fund.