

**AN ANALYSIS OF THE PROGRAM ASSESSMENT RATING TOOL:
MEASURING THE PERFORMANCE OF FEDERAL ENVIRONMENTAL AND NATURAL RESOURCE
PROGRAMS**

by

Nicole Buell and John Tortorella

Dr. Lori Bennear, Advisor

May 2011

Masters project submitted in partial fulfillment of the requirements for the Master of Environmental
Management degree in the Nicholas School of the Environment of Duke University

2011

Lori Bennear

Abstract:

Interest in, and use of, performance evaluation tools in the federal government has increased over the past two decades as a mechanism to gauge the effectiveness of federal programs. An effective evaluation system can improve government performance in many ways including boosting outcomes, strengthening accountability, and enhancing process transparency. A well-developed and consistently applied performance evaluation system can also assist in the identification of programs that excel and those that do not achieve their intended results. This research aims to analyze one such performance evaluation tool, the Office of Management and Budget's Program Assessment Rating Tool (PART), by examining a subset of programs concerned with the environment and natural resources. The objective of this study is to evaluate the effectiveness of PART as a means to hold agencies accountable for their performance. This objective is accomplished by examining the quality of goals, measures, and evidence provided by programs in their PART reviews as well as the relationship between PART rating and budgetary allocations. Focusing on the set of 167 natural resource and environmental program PARTs, no correlation was found between improved PART scores and budgetary allocations. Significant correlations between PART ratings and evidence and program measures were found. We conclude from this research that future performance evaluation mechanisms should focus more closely on program outcomes, while increasing understanding within agencies of the performance evaluation process and improving transparency for all parties.

Table of Contents:

Introduction:..... 4
Background: 4
 Performance Evaluation 4
 History of Performance Evaluation in the Federal Government 7
 The Performance Assessment Rating Tool 8
 Critiques of PART 10
 Research Objectives 10
Methods:..... 12
Results: 14
 Data and Summary Statistics: 14
 Research Theme: Utilization of Performance Information: 17
 Research Theme: PART as a Performance Budgeting Tool..... 25
Discussion: 27
 Recommendations..... 27
 Implications for the Future of Performance Evaluation 29
Literary Citations:..... 30
Appendix:..... 31
Acknowledgements: 35

Table of Figures:

TABLE 1: Agency Trends..... 14
TABLE 2: Total PART Ratings for all assessed programs compared to environment and natural resource programs 16
FIGURE 1: Assessment Years..... 16
TABLE 3: Program Types (A program may be more than one type) 17
FIGURE 2: Section Two Score Spread by Final Rating 17
FIGURE 3: Section Four Score Spread by Final Rating..... 18
TABLE 4: Pairwise Correlations Between Final Scores and Section Scores 18
TABLE 5: Model of Section Two Scores and Performance Measure Characteristics 19
TABLE 6: Model of Section Four Scores and Performance Measure Characteristics..... 20
TABLE 7: Model of Final Assessment Rating and Program and Agency Characteristics 21
TABLE 8: Model of Final Assessment Rating and Program and Agency Characteristics 24

Introduction:

The interest in, and use of, performance evaluation tools in the federal government has increased over the past decades in order to improve efficiency and enhance results (Metzenbaum 1998). An effective evaluation system can increase government performance in many ways including boosting outcomes, strengthening accountability, and enhancing process transparency (Metzenbaum 1998). The environmental field, in particular, has observed an increased use of performance information in order to better address current environmental issues. A well-developed performance evaluation system can assist in the identification of whether and when environmental problems exist and how best to create lasting solutions. Evaluation can be particularly challenging in the environmental field partly because the time scale for outcomes is often beyond the scale captured in an evaluation process.

This research aims to analyze one such performance evaluation tool – the Office of Management and Budget’s (OMB) Program Assessment Rating Tool (PART) – used to enhance natural resource program efforts. PART was a Bush Administration initiative that operated from 2002 – 2008 and evaluated approximately 98% of all federal programs (expectmore.gov). We analyzed 167 federal programs identified by OMB as natural resource and environment programs to examine how PART was used to encourage program improvement. Our research is focused around two themes: first, to examine how the performance information was used to determine a final rating and secondly, if PART was successfully used as a performance budget tool. We will explore if the quality of results and the demonstration of results are correlated with final ratings, if the PART guidance was followed and few quality measures of outcome versus outputs were encouraged, and if performance ratings influenced budget decisions.

Background:

Performance Evaluation

Performance-focused systems utilize information on program outcomes, as opposed to program processes, to encourage improvement and inform decision-making. There are many approaches, but typically these systems have three primary characteristics: the use of performance measures, the use of these measures to motivate improved performance and inform management and resource allocation decisions, and broad dissemination of the performance results all interested parties beyond just those assessed. Performance-focused systems allow flexibility to the programs to decide approaches to

implement measurement and improvement as long as performance levels are achieved. (Metzenbaum 1998)

Performance-based systems have been identified as useful tools to improve performance outcomes, boost accountability, enhance transparency, increase flexibility, and support fairness. Evaluations can boost outcomes by increasing awareness of problems (“what gets measured, gets done”), clarifying goals and measures to improve direction and expectations, and motivating improved performance particularly through rewards or penalties. Outcomes can also be improved through innovation, as the system allows the flexibility to decide the method for achieving results, and through learning as programs gather information to improve their activities and processes. (Metzenbaum 1998)

Another benefit of performance-focused systems is their promotion of accountability for results promised by a program. As was mentioned with the mechanisms to improve outcomes, programs are responsible for the consequences of their performance or lack of performance. Furthermore, transparency can improve accountability by improving the accessibility of performance information. Information can allow for public engagement in decision-making and redirection of government priorities through public education, awareness, and heightened responsibility. Lastly, performance-based systems are also characterized by flexibility and fairness by focusing on results instead of process and decreasing inconsistencies in inspections. (Metzenbaum 1998)

Performance-focused systems are not without their limits and weaknesses. Simply collecting measurement data alone will not lead to improvements. It is imperative that performance measures be carefully chosen, data collected accurately, and information be linked to a motivational mechanism (Metzenbaum 1998). The performance information must be well managed and actively and sensibly used to promote improvements.

There are three primary ways that performance measures are typically employed: as targets, for comparative and benchmarking purposes, and to facilitate experiential learning. Targets define a goal for achievement. Target-based systems utilize two steps – first a performance metric is selected and then target is selected from the possible range of outcomes of the metric. Benchmarking does not require that targets be set, but instead evaluates current performance based off of past performance. This approach is considered particularly valuable because measures have meaning only with context.

This method is often preferred for its dynamics, as a target-based system requires constant updating, whereas a comparative system automatically generates a new goal once a performance level is achieved. A final performance-based approach is through learning, which supports the previous two approaches. This system uses information about performance measures to apply lessons learned about what does or does not improve performance. (Metzenbaum 1998)

There are many components of an ideal evaluation system that are challenging to achieve in real life. Historically, evaluations of environmental agencies have focused more on whether the agency complied with administrative process instead of focusing on performance, which would more effectively judge the effect of the agencies' activities (Bennear and Dickinson 2005). Bennear and Dickinson (2005) explored several characteristics of an ideal performance system that involved a cyclic process of learning, acting, and evaluating where the end of one cycle marks the beginning of the next. However, actual government deviates from the ideal model frequently. There are many instances where knowledge of a problem doesn't motivate change and where outputs are incorrectly used instead of outcomes (Bennear and Dickinson 2005).

For example, a 2000 GAO report found that in fiscal year 2000, only 16 percent of EPA's goals and 12 percent of its measures were based on outcome-based performance measures. This report describes the difficulties within the EPA of devising systems of outcome-oriented performance. For the purposes of evaluating the effectiveness of programs, outcomes of policies and programs are generally preferable to other forms of data because they provide concrete evidence of results. A primary issue with creating performance measures based on outcomes, specifically for the EPA, is the nature of the programs involved. For programs that focus on pollution and other quantifiable media outcome variables are easy to develop and report, but for programs such as enforcement or research and development outcome measures are very difficult to devise. A number of EPA concerns were cited in this report. Specifically, the fact that environmental problems are long-term and highly interrelated makes a causal link between programs and outcomes difficult to establish. Data collection is also a problem because EPA relies on incomplete and incompatible information from states to fill its databases. (GAO 2000)

Program evaluations should be a built in part of the policy process much as the government has required a benefit-cost analysis prior the deployment of many policies. Bennear and Coglianese (2005) point out that requiring accurate analyses of the effectiveness of programs would create accountability to constituencies and a stream of feedback that would inform policymakers of what exactly is working

and what is not. This would allow for the development of a robust literature on evaluations where there is currently a relative scarcity of well-designed and executed program evaluations. The statistical tools that have been developed to do program evaluations or applied effectively to evaluations like differences-in-differences and its derivatives, propensity score matching, and instrumental variables make it possible for environmental policy analysts to determine effectiveness of programs even when randomization is not possible. (Benneer and Coglianesse 2005)

Benneer & Coglianesse discuss the increase in environmental regulations and programs in recent decades that are fit to be evaluated for causal impacts following the initial wave of command-and-control environmental legislation in the United States. Program evaluation mechanisms fit with many types of policy issues and developing a comprehensive literature on evaluation methodologies now will benefit society in the future through better designed policies and further refinement in evaluation techniques. Eventually program evaluation will lead to policies that are more flexible, adaptable and accountable to those they are intended to benefit. (Benneer and Coglianesse 2005)

History of Performance Evaluation in the Federal Government

While the private sector has utilized performance-based systems for many years, the integration of performance evaluations into the government was partially instigated because of a decreasing confidence in government (Metzenbaum 1998). One of the first major initiatives to hold government agencies more accountable was the Government Performance and Results Act (GPRA) enacted in 1993 by Congress. Prior to this Act, most efforts to implement performance programs were initiated through an executive order and consequently, efforts typically only lasted the length of the administration that introduced them (Metzenbaum 1998).

However, GPRA codified performance measurement and encouraged a shift from primarily focusing on outputs to looking at outcomes as well (Metzenbaum 1998). GPRA put an emphasis on creating outcome-oriented strategic plans with performance targets and required annual progress reports to the public (Heinrich 2010). A third component of the Act allowed agencies to waive certain administrative procedural requirements, such as staffing or pay levels, to allow flexibility to achieve greater performance results (Benneer and Dickinson 2005). While there is support that GPRA has contributed to improved government collection and use of performance information, the Act, however, lacks any

significant “teeth” (US Government Accountability Office 2008). GPRA is not explicit in what it considers acceptable evidence of performance, nor does it provide consequences for not demonstrating performance (Benneer and Dickinson 2005). Furthermore, GPRA does not tie performance demonstrations to budgeting decision-making (Benneer and Dickinson 2005).

Almost a decade after the introduction of GPRA, the second Bush administration introduced the Program Assessment Rating Tool (PART) in 2002. The program was administered by the Office of Management and Budget (OMB) and sought to link program performance to budget decisions. The focus of PART, as opposed to GPRA, shifted from agencies to specific agency programs and was first rolled out with the fiscal year (FY) 2004 budget (Gilmour and Lewis 2006). However, PART is seen as a vehicle to achieve the goals established in GPRA. PART strengthens and reinforces performance measures recommended under GPRA, by guiding the development of ambitious and achievable outcome oriented measures and requiring consistency of the measures across GPRA and PART reports and assessments (PART Guidance 2007). PART goes one step further than GPRA and links budget decisions to performance ratings. Programs are required to justify resource requests based off performance achieved and goals set (PART Guidance 2007). To date, OMB reports to have assessed 98% of the programs in the federal government, accounting for around 1015 programs in total (expectmore.gov).

The Performance Assessment Rating Tool

The PART is a diagnostic tool developed to drive improvements in federal program performance. The program generates program ratings in four primary areas: program purpose and design, strategic planning, program management, and program results and accountability. Responses are required to be evidence based. To earn a high PART rating, programs must implement performance measures, justify resource requests by achieved results, and continually improve efficiency. (PART Guidance 2007)

It is important to note that a program as defined in PART, is not always reflective of how a program is managed in practice. Programs are selected for review in consultation with the agency and OMB. In many cases the PART programs may be a combination of programs (as recognized by existing structures, including the Budget) with similar goals that would not be logical to review individually. (PART Guidance 2007)

The quality of a program's performance goals and measures of performance are recognized as the primary determinants of an overall PART rating. OMB requires programs to define and select program measures that reflect a program's mission and priorities, be few in number, and provide information to inform decisions. Measures should reflect outcomes, but there may be instances where output measures are also acceptable as proxy outcomes. In addition to outcomes and outputs, PART also requires efficiency measures to encourage improvement of agency resource use. Furthermore, PART requires that once measures have been developed that ambitious and achievable targets should be set. The guidance also recommends the use of baseline measures from which to measure improvements. Performance goals are distinguished between long-term (covering a multi-year period) and annual (generally more near-term) goals. (PART Guidance 2007)

PART ratings are calculated from the totals of the section scores, each of which have different weights: program purpose and design – 20%, strategic planning – 10%, program management – 20%, and program results and accountability – 50%. Sections two and four deal with performance information. Raw scores are not reported to the public as they would give a false degree of precision, but are instead translated into 5 qualitative ratings: effective, moderately effective, adequate, ineffective, and results not demonstrated. The latter score is received when programs do not have acceptable performance measures or lack baseline and performance data. (PART Guidance 2007)

Responses for each question are scored in a yes/no format. Section four allows for additional responses of large and small extent. Not applicable responses are permitted on specific questions, but not all. The first section of the PART contains five questions on the clarity and soundness of the program's purpose and design. The following planning section contains eight questions that focus on the quality of the measures and targets to ensure that planning, priority setting, and resource allocation are strategic and focused. The third section employs seven questions to gauge whether a program is effectively managed to improve and meet its performance goals. The final section, section four, contains 5 questions to rate overall program performance achievement on long-term and annual measures and targets. Additional questions may be added to each section for certain program types including capital asset and research and development programs. (PART Guidance 2007)

Each question is given equal weight, but programs may consult with OMB prior to beginning the assessment to change the question weights to emphasize key components of the program. Guidance is

provided on what constitutes the necessary evidence to receive a yes response for each question. A number of responses are linked. For example, if a program does not answer that they have adequate annual measures in section two, they must also receive a “no” response in section four when responding to questions about whether annual targets are achieved. Similarly, responses for certain questions might automatically generate a “results not demonstrated” rating. (PART Guidance 2007)

Critiques of PART

Gilmour and Lewis (2005) examined the impact of performance on budget decisions for the first round of PART assessments completed in 2004 and 2005. They looked at the difference in 2004 and 2005 budgets to control for political influence of programs in 2004 PART ratings and found that the section four scores are not as important as the three other sections of PART in determining funding, although all are correlated. They also break down the programs by size to evaluate changes in budgeting because smaller programs do not have the same kind of permanence as large and medium programs and can be terminated by OMB through budget decisions. (Gilmour and Lewis 2005)

Gilmour and Lewis (2006) conducted a study to examine the role of merit and political considerations on PART ratings. OMB has claimed a relationship between PART scores and budget allocations and the work of Gilmour and Lewis (2006) supported this claim. They found a positive relationship between scores and budget allocations, suggesting that the administration takes into consideration performance when preparing budget requests. However, they also found that the PART scores were influenced by politics and that the impact of the performance information on budgets reduced when taking into consideration this influence. In particular, the authors found that programs established under Democratic presidents received on average 5.5 points lower than programs established under Republican presidents. They also found that Democratic programs were more influenced by merit evaluations, whereas other program budgets were less influenced by performance results. Ultimately, the authors found that PART was operating somewhat successfully as a performance budget tool as performance measurement was correlated with the budget allocations, but that the performance data might not be used in a completely impartial manner. (Gilmour and Lewis 2006)

Research Objectives

For a performance evaluation system to be effective the measures must be appropriate and the information utilized in way that encourages improvement and innovation. We aim to evaluate the

effectiveness of the PART program as a tool to conduct evaluations and hold agencies accountable for results. In 2005, Benneer and Dickinson explored PART's ability to overcome the barriers to performance-focused environmental policy through the use of incentives (budget increases or decreases). They examined the performance ratings of over 397 programs from 2003-2005 and concluded that these incentives might encourage some programs to design and conduct more thorough evaluations, but PART can also create incentives to manipulate performance metrics, which might not result in improved government performance. Gilmour and Lewis (2006) conducted a similar study with 234 programs' PART scores over the same time period. They found that there was a positive relationship between an increase in PART scores and proposed budget increases but when decomposing the PART score into its four parts there was no relationship between the results score and the program budget (Gilmour and Lewis 2006).

Our study aims to build off these researchers' work and the work of Heinrich (2010), which examined 92 programs in the Department of Health and Human Services for the relationship between the quality of the measures and evidence supplied to the PART assessment and the assigned performance ratings to the amount of funding received by the programs. We will utilize a similar methodology to explore 167 environmental programs that completed PART assessments during the period of 2002-2008 in order to inform and make recommendations for improved government environmental performance evaluation. Furthermore, we hope to also inform the new performance evaluation systems being put in place in the Obama Administration. The 2010 Budget proposed to build off the performance data collection of the PART system but to put less emphasis on ratings and more on demonstration of results (Brodsky 2009). The Office of Management and Budget is currently developing a public site, Performance.gov, in order to increase the use of performance information. Our research aims to make recommendations to improve the collection and utilization of performance information for this new system.

The objective of this study is to evaluate the effectiveness of PART as a tool to conduct evaluations and hold agencies accountable for results by examining the quality of goals, measures, and evidence as well as the relationship between PART rating and budgetary funding. The study will specifically focus on PART as applied to natural resource and environmental programs.

The hypotheses to be tested are as follows:

Research Theme: Utilization of Performance Information

H1: The overall PART rating will be will positively correlated to the quality of measures (Section two score) and the demonstration of results (Section four score).

H2: The overall PART will be positively correlated to the quality of the evidence provided to support program results questions as measured by the characteristics of the program measures.

Research Theme: PART as a Performance Budgeting Tool

H3: Increases in federal funding received by the program will be positively correlated to the final PART rating.

In addition to the analysis of the three hypotheses presented above, we seek to provide suggestions for how to assess environmental program performance going forward.

Methods:

The analysis performed in this study relies on data and information provided by a number of programs focused on environment, natural resources, and energy which performed PART reviews for the time period 2002-2008 and which are federally funded. The 167 programs included in the study are based within twenty-eight agencies. Agencies with the most programs included in the sample are the Departments of Agriculture and Interior as well as the Environmental Protection Agency. For a complete listing of programs please refer to the appendix. While PART reviews were not conducted each year (2002-2008) for all programs there is a large dataset available. Additionally, the inclusion of programs from multiple departments and agencies allows for the comparison of performance across multiple sectors of the government.

Each detailed PART report can be accessed via www.expectmore.gov. Data that can be acquired from the detailed assessments of PART reviews includes the program type, assessment year, assessment rating, scores for each section (Program Purpose and Design, Strategic Planning, Program Management, Program Results/Accountability). In addition program funding level for fiscal years 2008 and 2009 are included. Questions asked in each section of the PART are available along with answers and supporting documentation, if submitted. This study will focus on the information reported for sections dealing with performance information, sections two and four. The program performance measures reported include the timeframe established to achieve a goal (either multi-year or annual), measures of outputs,

outcomes, or efficiencies. Each PART also includes a description of the unit of analysis used to assess a program's performance, performance targets and actual targets achieved.

From the available data as described new variables were constructed measuring the quality of the review process and evidence presented in the PART process. Data coding follows from the studies by Heinrich (2010) and Gilmour and Lewis (2006). Variables were coded to measure the type of information presented in support of programs, documentation of the types and category of performance measures (annual, long-term etc.) and whether or not performance outcomes were reported.

Budget data for each program was collected from agency budget documents from fiscal years 2000-2010. The President's proposed budget values and the Congressional enacted values were collected for each year, as the information was available. As mentioned earlier, it is important to note that programs assessed for PART are not perfect reflections of programs as they appear in the budget. There are many instances where like programs were combined for a single PART assessment. Budget numbers were matched based off of the reported fiscal year 2008 and 2009 numbers in the PART details on www.expectmore.gov.

Additional data on characteristics of the agencies implementing the program are included from the work of Lee, Rainey, and Chun (2009). These agency-level data provide information on the type of agency such as the salience of its mission in the congressional, presidential and media realms. These metrics are important because congress controls the purse strings and the PART is designed to be directly applied to the budgetary decision process. Higher levels of presidential and media salience may have an impact on budget decisions as well, although this link may be less direct than with congressional salience. The additional agency characteristics also include directive, goal, and evaluative ambiguity which Lee et al (2009) discuss as having the effect of reducing an agency or program's focus on outcomes and shifting attention more to processes and outputs. Data on the size of agencies, age of agencies, financial publicness, professional staffing and managerial capacity are also included. While these data are at the level of agencies and not programs they are still informative in determining which characteristics influence budget decisions.

The analysis in this study focuses on multiple regression to evaluate the relationships between the strength of the evidence presented in the PART reports and PART scores both in the aggregate and on the results of the various sections of PART. Regressions will be estimated to determine the relationship between the answers and evidence provided in the PART and the final PART rating. The dependent

variables in the regression models will be PART rating and PART sections two and four ratings. Independent variables in these models will be sourced from those constructed from the information presented in the PART process and characteristics of the department and/or agency conducting the program. Due to the nature of the data, with many program measures being associated with a single program the standard errors are clustered at the program level. To estimate the relationship between funding levels and final PART scores, dummy variables for “performing” (effective, moderately effective, and adequate) and “not performing” (ineffective and results not demonstrated) were created. These dummy variables were coded beginning two years after the assessment had been made to coordinate when the assessment evidence would appear in the budget. A second model was also created to assess the effect of individual PART ratings and funding levels coded by the same specifications.

Results:

Data and Summary Statistics:

Presented here are summary tables of the data collected from the PART reports on environment and natural resource programs.

TABLE 1: Agency Trends

Agency	# of Programs	Range of Assessment Rating	Average # of Program Measures
Agency for International Development	1	Adequate	3
Bureau of Indian Affairs and Bureau of	3	Adequate-Moderately Effective	5
Bureau of Land Management	7	Results Not Demonstrated-Adequate	8
Bureau of Reclamation	12	Results Not Demonstrated-Moderately Effective	4
Corps of Engineers-Civil Works	10	Results Not Demonstrated-Moderately Effective	6
Department of Agriculture	10	Adequate- Effective	6
Department of Commerce	2	Moderately Effective	7
Department of Defense--Military	1	Adequate	7
Department of Energy	2	Adequate- Moderately	13

		Effective	
Department of State	1	Adequate	3
Department of the Interior	2	Results Not Demonstrated- Adequate	4
Environmental Protection Agency	53	Results Not Demonstrated- Moderately Effective	6
Farm Service Agency	2	Results Not Demonstrated- Moderately Effective	5
Federal Highway Administration	1	Moderately Effective	6
Forest Service	6	Results Not Demonstrated- Moderately Effective	8
Maritime Administration	1	Effective	6
Minerals Management Service	4	Moderately Effective- Effective	6
National Aeronautics and Space Administration	2	Moderately Effective	21
National Oceanic and Atmospheric Administration	8	Results Not Demonstrated- Effective	10
National Park Service	10	Results Not Demonstrated- Moderately Effective	6
Natural Resources Conservation Service	4	Adequate- Moderately Effective	11
Office of Surface Mining Reclamation and Reconservation	3	Moderately Effective- Effective	6
Other	1	Effective	7
Power Marketing Administration	1	Moderately Effective	3
Tennessee Valley Authority	2	Adequate	3
United States Coast Guard	3	Results Not Demonstrated- Moderately Effective	4
United States Fish and Wildlife Service	5	Results Not Demonstrated- Effective	7
United States Geological Survey	1	Moderately Effective- Effective	8
TOTAL	28	167	Overall average=7

Table 1 illustrates basic characteristics of the 167 programs analyzed over the 28 agencies. The number of programs assessed, range of final rating scores received, and average number performance measures for each agency. The Environmental Protection Agency has the largest representation in the dataset, accounting for approximately 32% of the programs examined.

TABLE 2: Total PART Ratings for all assessed programs compared to environment and natural resource programs

Rating	All Programs	Env. And Nat. Resource Programs
Effective	193 (19%)	14 (8%)
Moderately Effective	326 (32%)	62 (37%)
Adequate	297 (29%)	68 (41%)
Ineffective	26 (3%)	3 (2%)
Results Not Demonstrated	173 (17%)	20 (12%)
Number of Programs Assessed	1015	167

Table 2 compares PART ratings for the set of all assessed federal programs with those of the environment and natural resource programs in this study. While the set of programs identified as “environment and natural resource programs” are not a random sample, they are fairly reflective of the total set of federal reviewed programs. However, there are somewhat notable differences in two rating categories, as the total set of federal reviews has a higher percentage of “effective” programs and a lower percentage of “adequate” rated programs.

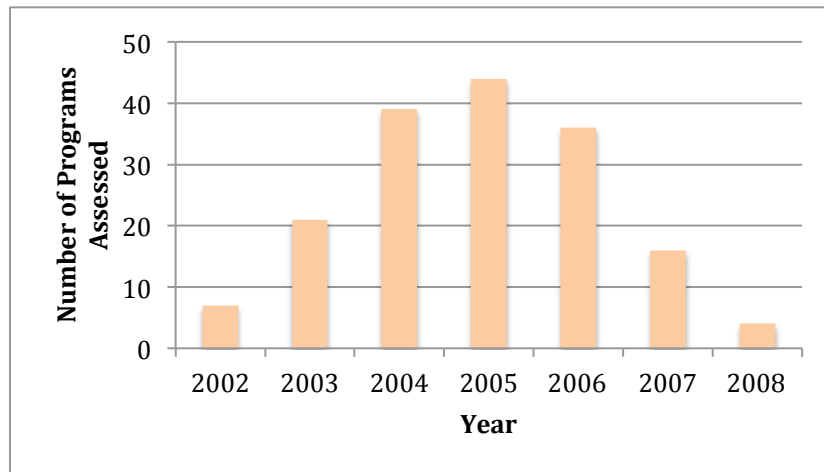


FIGURE 1: Assessment Years

Figure 1 demonstrates the ramp up of the PART program since its inception in 2002 through its declined use and ultimate conclusion in 2008. This figure is exclusive for the environment and natural resource programs.

TABLE 3: Program Types (A program may be more than one type)

Program Type	Count
Direct Federal	63
Competitive Grant	29
Block/Formula Grant	29
Regulatory-Based	28
Capital Assets and Service Acquisition	30
Credit	0
Research and Development	35

Table 3 shows that there is a fairly even distribution of program types within the dataset with the exception of the high count of direct federal programs and the lack of programs which are credit based..

Research Theme: Utilization of Performance Information:

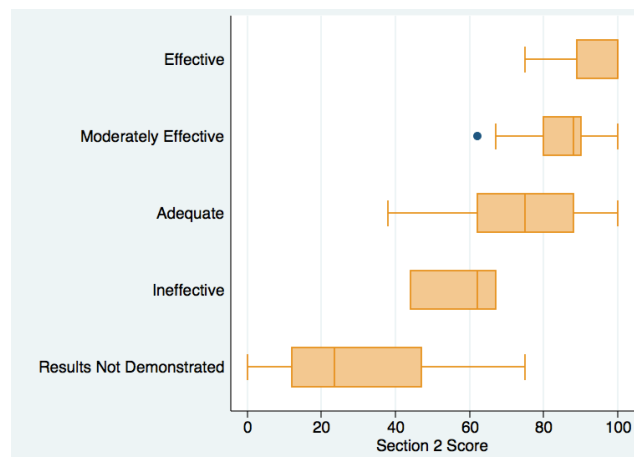


FIGURE 2: Section Two Score Spread by Final Rating

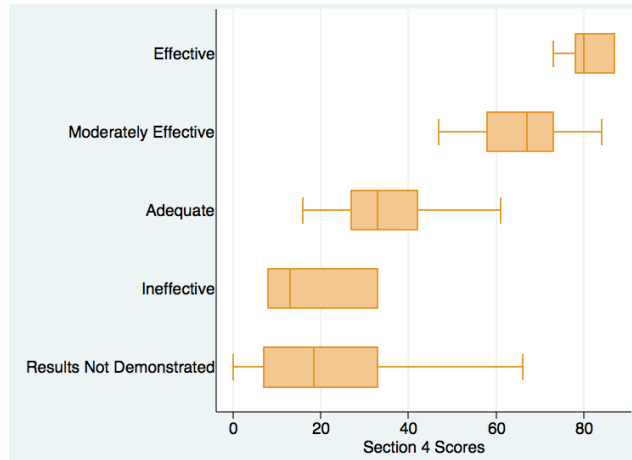


FIGURE 3: Section Four Score Spread by Final Rating

TABLE 4: Pairwise Correlations Between Final Scores and Section Scores

	Pairwise Correlation	P value
Section Two and Final Rating	.7684	~0
Section Four and Final Rating	.8087	~0

Figures 2 and 3 corroborate a positive correlation between the quality of measures (Section two) and demonstration of results (Section 4) and the final PART rating. This is supported by the correlations reported in Table 4, both significant at the 5% level. In Figures 2 and 3, most of the observations do not overlap across final ratings but there are several instances where this is not the case. This is indicative of the importance of the each section towards the final PART score as each is weighted differently to calculate the final PART rating. It is important to note the large spreads in the data for the “adequate” and “results not demonstrated” ratings in both Figures 2 and 3. The latter observation is likely a result of the four questions spread across Sections 2 and 4 that result in an automatic “results not demonstrated” rating regardless of the final overall score.

In order to evaluate the impact of the evidence provided in PART assessments on components of the assessment several regression models were estimated. The first models, presented below in Table 5 examine the impact of evidence and program type on Section 2 scores (quality of measures). The second set of models presented in Table 6 again examine the impact of evidence and program type but with Section 4 scores (demonstration of results) as the dependent variable. In each table the first column

presents regression results with independent variables for the total number of each measure type as well as indicator variables for the presence or absence of each measure type as well as a variable for the total number of measures. The second column presents regression results with additional indicator variables for the type of program being evaluated as well as variables for the year assessment.

TABLE 5: Model of Section Two Scores and Performance Measure Characteristics

VARIABLES	(1) Section Two Score	(3) Section Two Score
Total Output Measures	-22.00 (20.52)	-19.15 (20.89)
Total Outcome Measures	-22.83 (20.43)	-19.98 (20.76)
Total Efficiency Measures	-24.53 (20.47)	-21.53 (20.83)
No Output Measures	-10.35** (4.883)	-8.892* (4.977)
No Outcome Measures	-15.99** (6.435)	-17.47*** (6.424)
No Efficiency Measures	-31.74*** (6.990)	-31.53*** (7.039)
Total Number of Program Measures	23.41 (20.48)	20.11 (20.82)
Program Type- Direct Federal		-0.105 (4.648)
Program Type- Credit Program		0 (0)
Program Type- Research & Development		5.853 (5.559)
Program Type- Block/Formula Grant		-5.933 (5.082)
Program Type- Competitive Grant		-6.075 (4.711)
Program Type- Capital Assets and Service Acquisition		6.718 (5.242)
Program Type- Regulatory-Based		4.073 (5.102)
Year Assessment Completed		2.076* (1.248)
Constant	76.46*** (4.439)	-4,083 (2,500)
Observations	167	167
R-squared	0.289	0.339

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

TABLE 6: Model of Section Four Scores and Performance Measure Characteristics

VARIABLES	(1) Section Four Score	(2) Section Four Score
Total Output Measures	-42.99** (21.24)	-39.74* (21.10)
Total Outcome Measures	-43.30** (21.14)	-39.82* (20.96)
Total Efficiency Measures	-43.17** (21.18)	-38.40* (21.03)
No Output Measures	-5.945 (5.053)	-6.089 (5.026)
No Outcome Measures	-10.52 (6.659)	-13.78** (6.488)
No Efficiency Measures	-13.41* (7.234)	-12.76* (7.109)
Total Number of Program Measures	44.26** (21.20)	40.08* (21.03)
Program Type- Direct Federal		2.460 (4.693)
Program Type- Credit Program		0 (0)
Program Type- Research & Development		14.68*** (5.613)
Program Type- Block/Formula Grant		-0.123 (5.132)
Program Type- Competitive Grant		-5.305 (4.758)
Program Type- Capital Assets and Service Acquisition		12.93** (5.294)
Program Type- Regulatory-Based		12.21** (5.153)
Year Assessment Completed		1.783 (1.260)
Constant	43.66*** (4.593)	-3,535 (2,525)
Observations	167	167
R-squared	0.147	0.245

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

In Table 5 the absence of measures for all three categories (outcome, output, or efficiency) was found to be negatively correlated with section two scores (at the 5% level, at the 10% level with program characteristics). This is a logical result as section two grades the presence and quality of performance measures. A lack of measures would consequently result in a lower score. However, the same result is

not true for the section four scores. Table 6 illustrates that when including program characteristics only the lack of efficiency and outcome measures are statistically significant (at the 10% level) and only the lack of efficiency measures are statistically significant without the program type characteristics (at the 10% level). Instead, in this model, the number of measures of each type appears to be negatively correlated with section four scores (significant at the 5% level, at the 10% level with program characteristics). A possible interpretation of this result is that an increase in number of performance measures will not lead to an increase in information to demonstrate results. In addition, the second model in Table 6 shows that programs that are categorized as Regulatory, Research and Development, and Capital Assets & Service Acquisition are more likely to receive positive section four scores. This result may be due to the fact that these types of programs have a slightly higher number of total measures than other categories.

TABLE 7: Model of Final Assessment Rating and Program and Agency Characteristics

VARIABLES	(1) Final Assessment Rating	(2) Final Assessment Rating	(3) Final Assessment Rating	(4) Final Assessment Rating
Question 4.1 Score	0.125** (0.0510)	0.191*** (0.0681)	0.302*** (0.0826)	0.311*** (0.0865)
Question 4.2 Score	0.231*** (0.0552)	0.199*** (0.0624)	0.270*** (0.0539)	0.288*** (0.0553)
Question 4.3 Score	0.243*** (0.0428)	0.232*** (0.0633)	0.294*** (0.0689)	0.320*** (0.0726)
Question 4.4 Score	0.144*** (0.0393)	0.237*** (0.0593)	0.372*** (0.0965)	0.386*** (0.101)
Question 4.5 Score	0.195*** (0.0483)	0.273*** (0.0553)	0.362*** (0.0682)	0.421*** (0.0753)
Question 2.1 Score	0.576*** (0.152)	0.606*** (0.138)	0.982*** (0.236)	0.865*** (0.180)
Question 2.2 Score	0.0962 (0.0681)			
Question 2.3 Score	-0.218 (0.161)			
Question 2.4 Score	-0.0388 (0.0584)			
Question 2.5 Score	0.00343 (0.0518)			
Question 2.6 Score	0.0966 (0.0714)			
Question 2.7 Score	0.0616 (0.0449)			
Question 2.8 Score	-0.0309			

	(0.0533)			
Congressional Saliency		-0.197 (0.296)		
Presidential Saliency		1.556** (0.658)	0.874** (0.365)	0.848** (0.361)
Media Saliency		-0.564 (0.557)		
Natural Log of Agency Size		-1.497*** (0.484)	-1.446*** (0.419)	-1.351*** (0.445)
Program Type- Direct Federal			-1.856** (0.845)	-2.035*** (0.690)
Program Type- Research & Development			1.616 (1.211)	
Program Type- Block/Formula Grant			-1.049 (0.814)	
Program Type- Competitive Grant			-1.504* (0.778)	-0.970 (0.726)
Program Type- Capital Assets and Service Acquisition			3.261*** (0.957)	3.147*** (0.946)
Program Type- Regulatory-Based			3.612*** (0.879)	3.338*** (0.749)
Year Assessment Completed			0.172 (0.212)	
Total Number of Program Measures				-0.00117 (0.0662)
cut1				
Constant	6.184*** (1.436)	-5.616 (4.104)	344.6 (426.3)	0.966 (4.003)
cut2				
Constant	7.114*** (1.500)	-5.006 (4.117)	345.7 (426.3)	1.893 (4.130)
cut3				
Constant	14.35*** (1.772)	3.558 (4.119)	358.0 (426.4)	13.81*** (4.660)
cut4				
Constant	20.36*** (2.184)	10.74** (4.231)	367.3 (426.1)	23.85*** (5.453)
Observations	1,154	1,114	1,114	1,114

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

In Table 7, the first model simply presents the influence of the score on individual section two and four questions on the final assessment rating. For the purposes of this regression analysis, the final PART

ratings were converted to a scale of 0 to 4 with 0 representing a 'Results not Demonstrated' rating and 4 representing 'Effective' (Table 7). The ordered logit specification is used because the outcome variable in this model is converted to this likert scale. This specification allows for the outcome variable to have unequal distance between possible values. The model shows that, as expected and previously discussed, section four scores have highly significant and positive influence on assessment rating (Table 7). Of the Section 2 scores only question 2.1 has a significant and positive effect on final assessment rating (Table 7).

Model two in Table 7 again uses the ordered logit specification. In this model, the insignificant Section two variables are removed and variables for salience are added, in addition to a variable for the natural logarithm of agency size. The salience variables measure the importance of an agency's mission to various bodies, in this case the President, Congress, and media. Presidential salience has a positive and significant effect on final PART rating while Congressional and Media salience do not have a significant effect (Table 7). Interestingly, the variable for the natural logarithm of agency size shows a highly significant and negative relationship with final PART assessment rating (Table 7). This indicates that programs housed in larger agencies are likely to receive lower PART ratings. This result is surprising because one would expect larger agencies to display higher PART ratings due to a greater capacity to undertake the PART process. GAO report 06-28 reports anecdotally that programs in smaller agencies did in fact face challenges in completing the PART process.

Model three keeps the significant variables from model two and adds in dummy variables for program type (Table 7). The model shows that being a direct federal program has a significant and negative effect on overall PART rating while being a capital asset or regulatory program has a positive and highly significant effect on PART rating (Table 7). Being classified as a competitive grant program has a negative effect on overall PART rating with significance at the 10% level and the year in which the program was assessed is also insignificant in determining the final PART rating (Table 7).

Model four keeps the significant variables from model three and adds a variable for the number of program measures given in the PART report in support of a program (Table 7). This variable's odds ratio has a very small magnitude and is not significant (Table 7). This result is counterintuitive because of the emphasis on proving performance through measures. One might assume that inclusion of more measures would improve PART ratings but this result does not support that assumption.

TABLE 8: Model of Final Assessment Rating and Program and Agency Characteristics

VARIABLES	(1) Final Assessment Rating
Question 4.1 Score	0.300*** (0.0839)
Question 4.2 Score	0.312*** (0.0615)
Question 4.3 Score	0.329*** (0.0760)
Question 4.4 Score	0.388*** (0.106)
Question 4.5 Score	0.435*** (0.0784)
Question 2.1 Score	0.859*** (0.182)
Presidential Saliency	0.986*** (0.361)
Natural Log of Agency Size	-1.454*** (0.467)
Program Type- Direct Federal	-2.200*** (0.740)
Program Type- Competitive Grant	-0.939 (0.747)
Program Type- Capital Assets and Service Acquisition	2.802*** (0.940)
Program Type- Regulatory-Based	3.513*** (0.819)
Outcome Measure	0.0446 (0.338)
Efficiency Measure	0.0692 (0.235)
Long-term Measure	1.429** (0.684)
Annual Measure	1.386* (0.741)
Targets Identified for Measure	-1.272 (0.972)
Actual Measures Reported for Measure	1.144* (0.610)
Number of Actual Measures Reported	-0.0332 (0.103)
Number of Targets Reported	0.0559 (0.0809)
cut1	
Constant	1.471

cut2	(4.086)
Constant	2.458 (4.162)
cut3	
Constant	14.76*** (4.650)
cut4	
Constant	25.23*** (5.530)
Observations	1,093

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

In order to further examine the influence of measures on overall PART rating the model above was estimated based on the results from the previous models with the addition of characteristics of the measures provided in the PART reports (Table 8). This model incorporates variables that were significant in the previous models. Measures were characterized as outcomes, outputs, or efficiencies in the PART reports. The output variable is dropped in this model due to multicollinearity with the outcome and efficiency variables. In this model, having the characteristic of an outcome or efficiency measure does not significantly affect the final PART rating (Table 8). If a measure is a long term measure, it does have a significant and positive effect on the final PART rating (Table 8). If a measure is tracked on an annual basis and has an actual value reported it also positively affects the final PART rating, although this finding is only significant at the 10% level (Table 8).

Research Theme: PART as a Performance Budgeting Tool

In addition to evaluating the impact of performance rating as a product of the evidence and agency/program characteristics we also sought to show whether or not PART ratings had a significant, measureable effect on budgetary decisions. Using budget figures compiled from agency budget summaries, we have assembled a panel dataset of budget requests and enactments for the period 2000-2010 for a subset of the programs in our dataset. This subset is not random but rather created based on the budget data that was available for the programs in our full sample.

We then constructed a model of impact of PART rating on the proposed and enacted budget following PART review. The general model specification for this is as follows:

$$\text{Budget}_{it} = \beta_i + \beta_{it}\text{trend} + \gamma_G\text{Performing} + \gamma_B\text{Not_Performing} + \varepsilon$$

In this model, trend is a variable equal to the budget year minus 1999. The variables Performing and Not_Performing are indicators of the PART ratings. The null hypothesis is that performance ratings had no effect on the final budget decisions. The model assumes that a program's normal funding path would change as a result of an assessment rating. The effect of the Performing or Not_Performing ratings was applied to the budget two years after the assessment occurred because fiscal year budgets are developed two years prior to their implementation. For example, the effect of a rating would be applied beginning in 2006 for an assessment occurring in 2004. Therefore, beginning two years after the assessment, those programs rated 'Effective,' 'Moderately Effective,' or 'Adequate' were assigned a value of 1 for the variable Performing. Those rated 'Ineffective' or 'Results Not Demonstrated' were assigned a value of 1 for the variable Not_Performing beginning two years after the assessment. Regressions were performed with the model for both the President's proposed budget and the Congressional enactments to reveal if either used performance information when making funding decisions.

A second model was created to determine if individual assessment ratings would illustrate the effect on budget more appropriately than lumping ratings as 'Performing' or 'Not Performing.'" The general model specification is as follows:

$$\text{Budget}_{it} = \beta_i + \beta_{it}\text{trend} + \gamma_E\text{Effective} + \gamma_{ME}\text{Moderately_Effective} + \gamma_A\text{Adequate} + \gamma_I\text{Ineffective} + \gamma_R\text{Results_Not_Demonstrated} + \varepsilon$$

The null hypothesis is that performance ratings had no effect on the final budget decisions. Values of 1 were applied to the programs two years following the assessment rating for the appropriate assessment ratings. Again, regressions were performed with the model for both the President's proposed budget and the Congressional enactments to show if either used performance information when making funding decisions.

Categorical regressions for neither model produced results with statistical significance. The regression evidence ultimately supports there was no correlation between PART ratings and budget decisions. However, due to the lack of full budget information for every program, there is a low number of observations included in each of these models. We acknowledge that this reduces the strength of our findings. However, both statistical and anecdotal evidence leads us to believe that further data would not change our findings, but would increase confidence in our results.

Discussion:

The Obama Administration elected not to continue the PART program, so consequently PART ended in 2008 with the conclusion of the Bush Administration. In light of this, we aim to apply our results and findings in order to provide broad recommendations for government performance evaluation tools, rather than specific recommendations for PART improvement. Our results indicated that the PART program did not always function as intended, particularly in regards to influencing budget decisions as a driver for improvement. These conclusions are supported both through government reviews and anecdotal evidence and will be used to develop our recommendations.

Recommendations

1. **Collect and Measure Performance Information for Program Improvement.**

A major critique of PART was that agencies frequently felt that they were meeting a checklist rather than completing an assessment that would encourage improvement; the PART was much less useful for the agencies than for OMB (GAO 2005). This is understandable when only a little over half of the questionnaire (sections 2 and 4) collected performance information.

Furthermore, agencies frequently cited frustration with the required third-party evaluations on the PART (GAO 2005). This was partially because of financial constraints, but particularly because evaluations designed for internal and external audiences have different focuses (GAO 2005). Agencies felt the required evaluation would not produce internally valuable results (GAO 2005). In addition, recommendations from OMB on program improvements were not clearly related to problems and most recommendations were disconnected from creating observable short term performance improvements (GAO 2005b). For a performance evaluation system to be effective, it is imperative that it measure performance in a way that encourages learning and improvement.

2. Educate Agencies on Performance Evaluation

Our analysis suggests that agencies might not be clear on how to select appropriate measures. For example, there were a number of instances where a program would list the same measure twice, once as an output and once as an outcome. This might suggest a lack of knowledge on the difference of these two types of measures.

There also needs to be an emphasis on tracking outcomes (which measure results) over outputs (which measure process). As mentioned earlier, a 2000 GAO study found that only 12% of EPA's measures were based on outcomes. To best demonstrate and improve performance, an increased collection of outcomes should be encouraged.

Throughout this analysis, we were surprised to find the vast spread in the number of measures identified for each program. The 2007 PART Guidance specifically emphasizes that performance measures should be few in number, which would be taken into consideration in the grading of the evaluation. It is important to emphasize quality versus quantity. Increased agency knowledge is likely to lead to better measures, collection of information, and utilization of results to encourage improvement.

3. Shift Emphasis to the "Management" in the Office of Management and Budget

Our results suggested that there was not a strong correlation between budget decisions and performance ratings. However, perhaps this is the way the system should work. As Gilmour and Lewis question, if a program is performing poorly should its budget be cut because it is wasting money or increased so that it can improve (Gilmour and Lewis 2006)? Furthermore, if a performance system is inherently linked to the fate of a program this has the potential to discourage honest evaluation and evaluations at all. Agencies may respond perversely by actively trying to game the system and establish goals and data collection that are easy or which hide deficiencies of their programs (Moynihan 2009). Instead, the government should focus on a performance evaluation system that encourages continual improvement through better management rather than budget decisions.

A 2008 GAO report highlights several ways to increase the management aspect of performance evaluation including a recommendation to demonstrate leadership commitment to results-

oriented management and adopt a more strategic and crosscutting approach to overseeing government performance. The report noted a decreased reporting by managers that their organizations used performance information when adopting new program approaches or changing work processes from 1997-2007. The benefit of collecting performance information to inform improvement cannot be realized unless federal managers use the performance data to make management decisions (GAO 2008).

Furthermore, a National Academy of Public Administration 2007 report highlighted one of the major failings of PART in that it inherently focused on the performance of individual programs rather than across goals of many programs. Such cross-cutting information is necessary for making smart choices about the nation's environment and can improve the use of performance information.

4. Improve Transparency

Greater communication of performance information is frequently associated with greater use (GAO 2008). The performance information reported for the PART was consistently difficult to locate and rarely updated with the most current information. It is important that for the performance information to be effectively used, it be made readily available and up to date. The PART results might not have been tied to funding decisions partly because the results were not effectively communicated to Congressional staff members. GAO reported Congressional opinion that PART assessments were disconnected from the Congressional budget process (GAO 2005b). Performance information should not only be shared with the agencies but the public and other government parties as well. Increased transparency can lead to increased accountability to stakeholders and use by leadership.

Implications for the Future of Performance Evaluation

The Government Performance and Results Act (GPRA) was recently updated with the GPRA Modernization Act (GPRAMA) signed into law on January 4, 2011. This represents a shift from the in depth evaluation of programs under PART back towards a focus on high level goals and moves government evaluation in the direction of our recommendations.

GPRAMA mandates each agency to create a few High Priority Performance Goals (HPPGs) that are achievable within 18 to 24 months and are reflective of the core functions of agencies and programs.

Each HPPG has a number of indicators identified to measure progress. EPA, for example, now has three HPPGs for which there are eight indicators of progress, whereas EPA had over 100 indicators of progress under the PART system. This new approach focuses performance efforts in order to achieve results. One concern, however, regards the short time frame of the goals. As mentioned previously, environmental programs face a challenge in identifying short-term outcomes, and consequently this has led environmental programs, such as those at the EPA, to select output measures over outcomes for their HPPGs. This ongoing challenge will need to be considered in the future. (Public Law 111-352).

GPRAMA also enhances performance evaluation by improving consistency and transparency. Because GPRAMA is law, it more effectively holds agencies accountable for their performance by requiring progress reviews every four months and disallowing the abandonment of goals between administrations. GPRAMA also encourages improved transparency between agencies, Congress and the Executive office. Transparency is to be accomplished through the creation of a central repository for performance data accessible to agency employees as well as decision makers. There is hope to make a similar database available to the public in the near future (called Performance.gov). In addition, within each agency a deputy officer is now vested with the position of Chief Operating Officer to lead the development of goals and to monitor performance. A high level career employee will also be assigned the role of Performance Improvement Officer with responsibility for assuring that goals are accomplished and that all stakeholders are involved in the process. (Public Law 111-352)

GPRAMA holds great promise to improve on the lessons learned from PART. Where PART was an in depth review of the performance of all federal programs, GPRAMA aims to better understand the core functions of each agency while generating a culture accepting of performance evaluation.

Literary Citations:

Bennear, Lori S.; Coglianese, Cary (2005) "Program Evaluation of Environmental Policies: Toward Evidence-Based Decision Making" in Brewer, Gary D. and Paul C. Stern, eds. *Decision Making for the Environment: Social and Behavioral Science Research Priorities*, (Washington, DC: National Academies Press).

Bennear, Lori S.; Katherine L. Dickinson. "The Role of Program Evaluation in Environmental Policy: A Critical Evaluation of Incentive Structures." *Association of Public Policy Analysis and Management (APPAM) Fall Conference*. Ed. Print.

Brodsky, Robert. 2009. "Obama's Performance Agenda a Mix of Old and New". GovernmentExecutive.com. January 10, 2011 2011. <<http://www.govexec.com/dailyfed/0509/052909rb1.htm>>.

Gilmour, John B.; David E. Lewis. 2006. "Does Performance Measurement Work? An Examination of OMB's PART Scores." *Public Administration Review* 66.5, 742-52.

Gilmour, John B.; David E. Lewis. 2005. Assessing Performance Budgeting at OMB: The Influence of Politics, Performance, and Program Size. *J Public Adm Res Theory*. 16(2): 169-186

Government Accountability Office. 2000. "Managing for Results: Epa Faces Challenges in Developing Results-Oriented Performance Goals and Measures." Washington, DC, 2000. Print. GAO/RCED-00-77.

Government Accountability Office. 2005. "Program Evaluation: OMB's PART Reviews Increased Agencies' Attention to Improving Evidence of Program Results." Washington, DC. GAO-06-67.

Government Accountability Office. 2005b. "Performance Budgeting: PART Focuses Attention on Program Performance, but More can be Done to Engage Congress." Washington DC. GAO-06-28.

Government Accountability Office. 2008. "Government Performance: Lessons Learned for the Next Administration on Using Performance Information to Improve Results." Washington, DC. GAO-08-1026T.

Heinrich, Carolyn J. 2010. "How Credible Is the Evidence, and Does It Matter? An Analysis of the Program Assessment Rating Tool." Robert M. La Follette School of Public Affairs at the University of Wisconsin-Madison. Working Paper No. 2010-019.

Lee, Jung Wook; Hal G. Rainey; Young Han Chun. 2009. "Of Politics and Purpose: Political Salience and Goal Ambiguity of US Federal Agencies." *Public Administration*. Vol. 87, No. 3, p. 457-484.

Metzenbaum, Shelley. 1998. *Making Measurement Matter: The Challenge and Promise of Building a Performance-Focused Environmental Protection System*.

Moynihan, Donald (2009) "The Politics Measurement Makes: Performance Management in the Obama Era," *The Forum*: Vol. 7 : Iss. 4, Article 7.

Office of Management and Budget. 2007. "Guide to the Program Assessment Rating Tool (PART)." No. 2007-02.

Public Law 111-352. "GPRA Modernization Act of 2010. Washington DC, January 4, 2011.

White House, Office of Management and Budget. "Expectmore.gov"
<http://www.whitehouse.gov/omb/expectmore/>

Appendix:

The list presented below shows the 167 programs for which PART reviews were analyzed and data were collected.

Agency	Program Name
Department of Agriculture	Conservation Operations
Department of Agriculture	Conservation Reserve Program
Department of Agriculture	Emergency Conservation Program
Department of Agriculture	Emergency Pest and Disease Management Programs
Department of Agriculture	Emergency Watershed Protection Program
Department of Agriculture	Environmental Quality Incentives Program
Department of Agriculture	Farmland Protection Program
Department of Agriculture	Forest Service: Energy Resource Needs
Department of Agriculture	Forest Service: Forest Legacy Program
Department of Agriculture	Forest Service: Invasive Species Program
Department of Agriculture	Forest Service: Land Acquisition
Department of Agriculture	Forest Service: Recreation
Department of Agriculture	Forest Service: Watershed

Department of Agriculture	In-House Research for Natural Resource Base and Environment
Department of Agriculture	National Forest Improvement and Maintenance
Department of Agriculture	Natural Resource Base & Environment (Grants)
Department of Agriculture	On-going Pest and Disease Management Program
Department of Agriculture	Resource Conservation and Development
Department of Agriculture	USDA Wildland Fire Management
Department of Agriculture	Watershed Protection and Flood Prevention
Department of Agriculture	Wetlands Reserve Program
Department of Agriculture	Wildlife Habitat Incentives Program
Department of Commerce	Coastal Zone Management Act Programs
Department of Commerce	National Marine Fisheries Service
Department of Commerce	National Oceanic & Atmospheric Administration: Tsunami Monitoring, Forecasting, and Warning Program
Department of Commerce	National Oceanic & Atmospheric Administration: Climate Program
Department of Commerce	National Oceanic & Atmospheric Administration: Ecosystem Research
Department of Commerce	National Oceanic & Atmospheric Administration: Marine and Aviation Operations
Department of Commerce	National Oceanic & Atmospheric Administration: Navigation Services
Department of Commerce	National Oceanic & Atmospheric Administration: Protected Areas
Department of Commerce	National Oceanic & Atmospheric Administration: Weather and Related Programs
Department of Commerce	Pacific Coastal Salmon Recovery Fund
Department of Defense-- Military	National Security Space Weather Programs
Department of Energy	Bonneville Power Administration
Department of Energy	Environmental Management
Department of Energy	Environmental and Post-Retirement Liabilities
Department of Homeland Security	Coast Guard Fisheries Enforcement
Department of Homeland Security	Coast Guard Marine Environmental Protection
Department of Homeland Security	Coast Guard: Polar Icebreaking Program
Department of the Interior	Bureau of Indian Affairs - Dam Safety and Dam Maintenance
Department of the Interior	Bureau of Indian Affairs - Forestry Maintenance
Department of the Interior	Bureau of Indian Affairs - Natural Resource Programs
Department of the Interior	Bureau of Indian Affairs - Operation and Maintenance of Irrigation Projects
Department of the Interior	Bureau of Land Management - Energy and Minerals Management
Department of the Interior	Bureau of Land Management - Mining Law Administration
Department of the Interior	Bureau of Land Management - Realty and Ownership Management
Department of the Interior	Bureau of Land Management - Recreation Management
Department of the Interior	Bureau of Land Management - Resource Management
Department of the Interior	Bureau of Land Management - Southern Nevada Land Sales
Department of the Interior	Bureau of Reclamation - Hydropower
Department of the Interior	Bureau of Reclamation - Recreation and Concessions
Department of the Interior	Bureau of Reclamation - Rural Water Supply Projects
Department of the Interior	Bureau of Reclamation - Safety of Dams Program
Department of the Interior	Bureau of Reclamation - Science and Technology Program
Department of the Interior	Bureau of Reclamation - Water Management - Project Planning and Construction
Department of the Interior	Bureau of Reclamation - Water Reuse and Recycling
Department of the Interior	Bureau of Reclamation - Water Management & Environmental Mitigation

Department of the Interior	Bureau of Reclamation Water Management - Operation and Maintenance
Department of the Interior	Bureau of Reclamation - California Federal Bay-Delta (CALFED)
Department of the Interior	Bureau of Reclamation - Central Valley Project Improvement Act
Department of the Interior	Department of the Interior - Central Utah Project
Department of the Interior	Department of the Interior - Land and Water Conservation Fund Land Acquisition
Department of the Interior	Department of the Interior - Wildland Fire Management
Department of the Interior	Fish and Wildlife Service - Endangered Species
Department of the Interior	Fish and Wildlife Service - Fisheries
Department of the Interior	Fish and Wildlife Service - Habitat Conservation
Department of the Interior	Fish and Wildlife Service - National Wildlife Refuge System
Department of the Interior	Fish and Wildlife Service - Wildlife and Sport Fish Restoration
Department of the Interior	Minerals Management Service - Minerals Revenue Management
Department of the Interior	Minerals Management Service - Outer Continental Shelf Environmental Studies
Department of the Interior	Minerals Management Service - Outer Continental Shelf Minerals Evaluation and Leasing
Department of the Interior	Minerals Management Service - Outer Continental Shelf Minerals Regulation and Compliance
Department of the Interior	National Park Service - Concessions Management
Department of the Interior	National Park Service - Cultural Resource Stewardship
Department of the Interior	National Park Service - Facility Maintenance
Department of the Interior	National Park Service - Heritage Partnership
Department of the Interior	National Park Service - Land and Water Conservation Fund State Grants
Department of the Interior	National Park Service - National Historic Preservation
Department of the Interior	National Park Service - Natural Resource Stewardship
Department of the Interior	National Park Service - Park Police
Department of the Interior	National Park Service - Technical Assistance
Department of the Interior	National Park Service - Visitor Services
Department of the Interior	Office of Surface Mining - Federal Managed Regulation of Surface Coal Mining and Abandoned Mine Land Reclamation
Department of the Interior	Office of Surface Mining - State Managed Abandoned Coal Mine Land Reclamation
Department of the Interior	Office of Surface Mining - State Managed Regulation of Surface Coal Mining
Department of the Interior	US Geological Survey - Biological Information Management and Delivery
Department of the Interior	US Geological Survey - Biological Research and Monitoring
Department of the Interior	US Geological Survey - Coastal and Marine Geology
Department of the Interior	US Geological Survey - Energy Resource Assessments
Department of the Interior	US Geological Survey - Geographic Research, Investigations, and Remote Sensing
Department of the Interior	US Geological Survey - Geologic Hazard Assessments
Department of the Interior	US Geological Survey - Mineral Resource Assessments
Department of the Interior	US Geological Survey - National Cooperative Geological Mapping
Department of the Interior	US Geological Survey - Water Information Collection and Dissemination
Department of the Interior	US Geological Survey - Water Resources Research
Department of State	Contributions to International Fisheries Commissions
Department of State	International Boundary and Water Commission
Department of Transportation	Federal Lands Highway Program
Department of Transportation	Maritime Administration Ship Disposal Program
Corps of Engineers- Civil Works	Aquatic Ecosystem Restoration
Corps of Engineers- Civil Works	Coastal Storm Damage Reduction

Corps of Engineers- Civil Works	Corps of Engineers: Coastal Ports and Harbors
Corps of Engineers- Civil Works	Corps of Engineers: Environmental Stewardship
Corps of Engineers- Civil Works	Corps of Engineers: Hydropower
Corps of Engineers- Civil Works	Corps of Engineers: Recreation Management
Corps of Engineers- Civil Works	Corps of Engineers: Water Storage for Water Supply
Corps of Engineers- Civil Works	Flood Damage Reduction
Corps of Engineers- Civil Works	Flood and Coastal Storm Damage Reduction (with Dam Safety)
Corps of Engineers- Civil Works	Formerly Utilized Sites Remedial Action Program
Environmental Protection Agency	Air Quality Grants and Permitting
Environmental Protection Agency	Alaska Native Village Water Infrastructure
Environmental Protection Agency	Brownfields Revitalization
Environmental Protection Agency	Chesapeake Bay Program
Environmental Protection Agency	Clean Water State Revolving Fund
Environmental Protection Agency	Drinking Water Protection Program
Environmental Protection Agency	Drinking Water Research
Environmental Protection Agency	Drinking Water State Revolving Fund
Environmental Protection Agency	EPA Acid Rain Program
Environmental Protection Agency	EPA Chemical Risk Review and Reduction
Environmental Protection Agency	EPA Climate Change Programs
Environmental Protection Agency	EPA Ecological Research
Environmental Protection Agency	EPA Enforcement of Environmental Laws (Civil)
Environmental Protection Agency	EPA Enforcement of Environmental Laws (Criminal)
Environmental Protection Agency	EPA Environmental Education
Environmental Protection Agency	EPA Great Lakes Program
Environmental Protection Agency	EPA Human Health Research
Environmental Protection Agency	EPA Indoor Air Quality
Environmental Protection Agency	EPA Lead-Based Paint Risk Reduction Program
Environmental Protection Agency	EPA Oil Spill Control
Environmental Protection Agency	EPA Pesticide Enforcement Grant Program
Environmental Protection Agency	EPA Pesticides and Toxics Research
Environmental Protection Agency	EPA Radiation Protection Program
Environmental Protection Agency	EPA Support for Cleanup of Federal Facilities
Environmental Protection Agency	EPA Tribal General Assistance Program
Environmental Protection Agency	EPA's Recycling, Waste Minimization, and Waste Management Program
Environmental Protection Agency	Endocrine Disruptors
Environmental Protection Agency	Global Change Research
Environmental Protection Agency	Human Health Risk Assessment Program
Environmental Protection Agency	Land Protection and Restoration Research
Environmental Protection Agency	Leaking Underground Storage Tank Cleanup Program
Environmental Protection Agency	Mobile Source Air Pollution Standards and Certification
Environmental Protection Agency	National Ambient Air Quality Standards Research
Environmental Protection Agency	National Ambient Air Quality Standards and Regional Haze Programs
Environmental Protection Agency	Nonpoint Source Pollution Control Grants
Environmental Protection Agency	Ocean, Coastal, and Estuary Protection
Environmental Protection Agency	Pesticide Field Programs
Environmental Protection Agency	Pesticide Registration
Environmental Protection Agency	Pesticide Reregistration
Environmental Protection Agency	Pollution Prevention Program
Environmental Protection Agency	Pollution Prevention and New Technologies Research
Environmental Protection Agency	Public Water System Supervision Grant Program

Environmental Protection Agency	Resource Conservation and Recovery Act Corrective Action
Environmental Protection Agency	Stratospheric Ozone Protection
Environmental Protection Agency	Superfund Remedial Action
Environmental Protection Agency	Superfund Removal
Environmental Protection Agency	Surface Water Protection
Environmental Protection Agency	Toxic Air Pollutants - Regulations and Federal Support
Environmental Protection Agency	U. S.-Mexico Border Water Infrastructure
Environmental Protection Agency	Underground Injection Control Grant Program
Environmental Protection Agency	Underground Storage Tank Program
Environmental Protection Agency	Water Pollution Control Grants
Environmental Protection Agency	Water Quality Research
International Assistance Programs	US Agency for International Development Climate Change Program
National Aeronautics & Space Administration	NASA Earth Science
National Aeronautics & Space Administration	NASA Earth-Sun System Research
Tennessee Valley Authority	Tennessee Valley Authority - Nox Emissions Reduction
Tennessee Valley Authority	Tennessee Valley Authority Resource Stewardship

Acknowledgements:

We would like to thank our Advisor, Dr. Lori Benneer, for her patience and guidance the past few months. Without her support and willingness to take on yet another MP this project would not have been possible. We would also like to thank Dr. Jeff Vincent and Dr. Mike Orbach for their assistance in the fall semester. We extend our gratitude to Dr. Hal Rainey at the University of Georgia for sharing his dataset. We also extend our appreciation to James Hurban, Michael Messinger, and John MacNeil at the Office of Management and Budget for sharing their knowledge of PART and government performance evaluation. Finally, we thank our family and friends for their long lasting humor, advice, and support throughout this process.