



MEMORANDUM

January 19, 2011

Subject: Obama Administration Agenda for Government Performance: Evolution and Related Issues for Congress

From: Clinton T. Brass
Analyst in Government Organization and Management
7-4536
cbrass@crs.loc.gov

This memorandum was prepared to enable distribution to more than one congressional office.

This memorandum responds to congressional interest in the Barack Obama Administration's agenda for government performance, how the agenda has evolved, and related issues for Congress. Specifically, the memorandum discusses milestones in the agenda's development, including the Administration's statements of plans and commitments regarding how it would work with Congress and agencies.

Two tools of policy analysis—program evaluation and performance measurement—often are used to inform assessments of government performance. That is, these tools generate information for multiple audiences, including Congress, agencies, the President, and the public, to help inform assessments of how agencies, programs, and policies are performing or might be improved. Consequently, program evaluation and performance measurement may inform related oversight, substantive lawmaking, and resource allocation by Congress. Congress also may intend for the tools to inform related policy making, budgeting, implementation, and management in agencies and the Executive Office of the President.

However, perfunctory reliance on, and use of, such tools also can present challenges. For example, evaluations and measurements may be flawed or may be presented in a biased or selective fashion. Without careful use, they also may drive unintended behaviors by entities that are being evaluated.

The Obama Administration's approach to government performance may be of ongoing relevance to Congress in at least two ways. First, the Administration's approach may be of interest when Members and committees of Congress assess the credibility and relevance of evaluations or measurements that are presented in attempts to influence and inform Congress's thinking, lawmaking, and oversight. Second, and more broadly, the Administration's approach may be of interest when Members and committees consider establishing or modifying federal performance-related policies and processes such as evaluation, planning, and budgeting. Processes like these often are intended to generate useful information for Congress, agencies, the President, and the public, and thereby to inform both policy making and the process of faithfully and effectively executing the laws.

The memorandum concludes by analyzing several issues in each of these two areas.

Contents

Introduction	3
Congress and Government Performance	3
Announcements from Presidential Transition.....	4
New CPO Duties	5
Already Established DDM Duties.....	5
Statements at Confirmation Hearings for OMB Officials	6
Post-Inauguration Agenda and Activities	7
Inauguration and Killefer Withdrawal.....	7
Obama Administration Budget Overview for FY2010.....	8
Announcement of Zients Nomination	9
Detailed Budget Proposals for FY2010 and Related Policies.....	9
Proposals for Program Terminations and Reductions	9
Themes for a “New Management and Performance Agenda”	10
Announcement of a “Performance Improvement and Analysis Framework”.....	10
Pursuit of Performance Improvement: Three Related Efforts	14
Zients Statements About Plans at Confirmation Hearing	15
Planning for High-Priority Performance Goals and FY2011 Budget.....	16
Development of a “Management and Performance Framework”	16
High-Priority Performance Goals (HPPGs).....	16
Planning for the President’s FY2011 Budget Proposals.....	18
Developing a New Performance Management System	18
Fall 2009 Congressional Hearings	18
OMB Emphasis on Impact Evaluations.....	20
Statements on Implementation Challenges and OMB’s Role	22
FY2011 President’s Budget and Announcement of HPPGs	23
High-Priority Performance Goals and “Performance Management Strategies”	23
Impact Evaluations and Characterizations of “Evidence-Based Policy”.....	24
Line-by-Line Review and Proposals for Termination or Reduction	26
February 2010 Framework: Six “Performance Strategies”.....	26
Looking Ahead to the President’s FY2012 Budget Proposals and 112 th Congress	27
Potential Issues for Congress	28
Potential Issues When Evaluations and Performance Measures Are Presented to Congress.....	29
Scrutinizing Program Evaluations and Performance Measurements	29
Scrutinizing How Agencies or OMB Define “Success” for a Program.....	31
Identifying Possibilities of Perverse Incentives	32
Making a Connection Between Evidence and Policy Making.....	33
Potential Issues When Congress Focuses on Performance-Related Policy	36
Opportunities for Congressional Oversight: Lessons Learned from Past Efforts?.....	36
Oversight of the Administration’s Performance Agenda: Coverage and Potential Gaps.....	37
Generic Framework for Assessing Influences on Government Performance	37
Example of Using the Framework: Federal Housing Policies	38
Assessing Administration Initiatives for Coverage and Potential Gaps	40
Examining GPRA and E.O. 13450.....	41

Introduction

On January 7, 2009, President-elect Barack Obama announced part of his agenda for improving the performance of the federal government. He said he would establish a new, non-statutory position in the White House with the title Chief Performance Officer (CPO). The CPO would report directly to the President and would seek to make the federal government more efficient, effective, transparent, and accountable. The President-elect also announced his intention to appoint Nancy Killefer, a consulting firm partner, to the CPO position, and to nominate her also to be Deputy Director for Management (DDM) at the Office of Management and Budget (OMB). The DDM position was established by law and is subject to Senate-confirmation.¹ OMB's DDM is responsible for establishing general management policies for executive agencies across a variety of mission-support functions. In occupying White House and OMB positions, Ms. Killefer would wear "dual hats." On February 3, 2009, however, Ms. Killefer withdrew from her expected appointment as CPO and nomination for DDM to avoid potential controversy related to a personal tax issue. Two months later, President Obama named Jeffrey Zients, the founder of a private equity firm, to be CPO and DDM. The Senate confirmed Mr. Zients as DDM on June 19, 2009. A succession of related developments followed in 2009 and 2010.

This memorandum focuses on the Obama Administration's agenda related to "government performance," much of which has been led by the CPO-DDM (hereafter CPO). In particular, it focuses on issues related to program evaluation and performance measurement.² Program evaluation and performance measurement may help inform assessments of how agencies, programs, and policies are performing. Consequently, they may help inform related policy making, budgeting, oversight, implementation, and management. However, the nature of the CPO position and the Administration's agenda have emerged gradually and are still evolving. For example, after Ms. Killefer's withdrawal, the Administration no longer referred to the CPO as a White House position. In addition, although the CPO's portfolio was articulated in some detail during the 2008 presidential campaign, the Administration later modified aspects of these plans. Over time, the Administration has announced a series of "frameworks"—that is, strategies intended to guide actions—and commitments for how the Administration would proceed, both administratively with agencies and in working with Congress. Because the Administration's agenda continues to evolve, the memorandum discusses selected developments chronologically.

Congress and Government Performance

The Obama Administration's agenda for government performance arrives as the most recent in a line of presidential initiatives. To help frame this subject, however, the memorandum first presents a brief discussion of potential *congressional* roles and activities related to government performance. Many writings on the subject of government performance treat the subject from the perspective of agencies or the President.³ Fewer seem to approach the subject from the perspective of Congress. To conclude on this basis that Congress's roles are secondary to those of agencies and the President, however, may be a

¹ Chief Financial Officers Act of 1990 (CFO Act), P.L. 101-576; 104 Stat. 2839; codified at 31 U.S.C. § 502(c).

² Other statutorily established, presidentially appointed, and mostly Senate-confirmed positions at OMB and the Office of Personnel Management (OPM) more directly supervise other mission-support or "management" functions. These functions, which generally are not the focus of this memorandum, include financial management, procurement policy, information technology, information and regulatory policy, and human resources management.

³ For example, see Chris Wye, *Performance Management for Career Executives: A "Start Where You Are, Use What You Have" Guide* (Washington: IBM Center for the Business of Government, October 2004); and Barry White and Kathryn E. Newcomer, eds., *Getting Results: A Guide for Federal Leaders and Managers* (Vienna, VA: Management Concepts, Inc., 2005).

misinterpretation. Current law grants significant discretion in this area to agencies and OMB. However, in possessing the Constitution's Article I legislative powers, among others, Congress's potential roles are pervasive and difficult to overstate. Congress routinely legislates on the subject and conducts related oversight, variously cooperating and competing with the President for control of agencies and policies.

In particular, two tools of policy analysis—program evaluation and performance measurement—often are used to inform assessments of government performance.⁴ They also may be used along with other policy analysis tools to inform policy making. In either case, Congress may assume at least two major roles.⁵

First, Congress may use specific evaluations and measurements to help inform its thinking, policy making, and oversight activities. In the course of Congress's legislative work, actors inside and outside of government frequently cite evaluations and measures to justify policy proposals and recommendations. In these situations, consumers of performance information, including Congress, can face challenges of assessing (1) quality and depth of evaluation information, which can be uneven, and (2) the relevance of the information to a policy problem, which can vary. Should Congress want to scrutinize these representations in its legislative work, insight into how to assess the quality, coverage, and relevance of evaluations and measurements may be helpful.

Second, Congress often makes policy regarding how agencies are to conduct and use evaluations and measurements. Congress may, among other things, establish processes for the production and consideration of evaluation information, establish entities to evaluate programs and policies, set aside related funding, articulate questions to be researched, and specify methods to be used. When policy makers consider these options, many observers would likely view at least two perspectives as being important: the institutional arrangements under which agencies produce and use performance information, such as who does it, how they do it, and for which audiences' needs; and also the practical capabilities and limitations of various evaluation and measurement approaches. Both of these perspectives might be used when assessing the Administration's agenda, formulating congressional responses, or acting independently of the agenda in areas of congressional interest.

Announcements from Presidential Transition

During the 2008 presidential campaign, then-Senator Obama outlined detailed plans for a CPO and other related initiatives.⁶ After the election, the website of the presidential transition referred to many of the plans, including establishing the CPO position, creating a special "SWAT" team in the White House, setting "performance targets" for which managers and programs would be held responsible, and

⁴ As discussed later in this memorandum (see **Box 3**), program evaluations may use one or more methods to assess how, and the extent to which, programs achieve intended objectives or cause unintended consequences. By contrast, performance measurement is narrower in focus and typically refers to periodic monitoring of data that are related to government programs.

⁵ See discussion in CRS Report RL33301, *Congress and Program Evaluation: An Overview of Randomized Controlled Trials (RCTs) and Related Issues*, by Clinton T. Brass, Erin D. Williams, and Blas Nuñez-Neto.

⁶ See Obama '08, *Blueprint for Change: Obama and Biden's Plan for America*, no date (83 pp.), p. 20, at <http://www.barackobama.com/pdf/ObamaBlueprintForChange.pdf>; and Obama-Biden, *The Change We Need in Washington: Stop Wasteful Spending and Curb Influence of Special Interests So Government Can Tackle Our Great Challenges*, no date [September 22, 2008], pp. 5-7, http://obama.3cdn.net/0080cc578614b42284_2a0mvyxpz.pdf. See also CQ Transcripts Wire, "Sen. Barack Obama Speaks at Campaign Event in Green Bay, Wisconsin," September 22, 2008, at <http://www.washingtonpost.com/wp-dyn/content/article/2008/09/22/AR2008092201500.html>.

conducting a “line-by-line” review of the budget.⁷ Additional details soon emerged about the Administration’s performance agenda.

New CPO Duties

On January 7, 2009, President-elect Obama announced that he would appoint Nancy Killefer to serve as CPO.⁸ According to the presidential transition website, the CPO would be a “new White House position.”⁹ Consistent with campaign proposals, the website indicated that the CPO “will report directly to the president.”¹⁰ The President-elect also announced that he would nominate Ms. Killefer to be DDM at OMB. Her previous writings on government performance suggested strategies that she might pursue.¹¹ At the press conference, the President-elect tied the CPO appointment to his campaign plans to conduct a “line-by-line” review of the federal budget.¹² He also described other CPO duties. For example, he said he would instruct members of the cabinet and key members of their staffs to meet with the CPO soon after taking office—and on a regular basis thereafter—to discuss how they can run their agencies with greater efficiency, transparency and accountability. The President-elect also announced that by February 2009, when a “rough budget outline” would be released for FY2010, further details would be forthcoming about “how we’re going to approach eliminating waste in government, one of Nancy’s tasks.”¹³

Already Established DDM Duties

The President-elect’s announcement signaled a decision to combine the DDM and new CPO positions in one individual. By statute, the DDM reports to the Director of OMB.¹⁴ The DDM position has statutory responsibility—subject to the direction and approval of the Director of OMB—to establish general management policies for executive agencies across a large number of “management functions,” which also are known as “mission-support functions.” These include financial management, “managerial systems” (including performance measurement), procurement policy, grant management, information and statistical policy, property management, human resources management, regulatory affairs, organizational studies, long-range planning, program evaluation, productivity improvement, and experimentation and demonstration programs. The DDM also chairs or plays roles in interagency councils of “chief officers.”¹⁵ In practice, the DDM may exert considerable influence over agencies through these councils.

⁷ See Obama-Biden Transition Team, “Agenda: Ethics,” at http://change.gov/agenda/ethics_agenda/.

⁸ At the time of the announcement, Ms. Killefer was a senior partner and director of the management consulting firm McKinsey and Company. From 1997 to 2000, she was Assistant Secretary for Management and Chief Financial Officer at the Department of the Treasury. See Office of the President-elect (Obama), “President-elect Names Nancy Killefer as Chief Performance Officer,” press release, January 7, 2009, at http://change.gov/newsroom/entry/president-elect_obama_names_nancy_killefer_as_chief_performance_officer.

⁹ Office of the President-elect (Obama), “New and More Efficient Ways of Getting the Job Done,” press release, January 7, 2009, at http://change.gov/newsroom/entry/new_and_more_efficient_ways_of_getting_the_job_done.

¹⁰ Obama-Biden Transition Team, “Agenda: Ethics,” at http://change.gov/agenda/ethics_agenda/.

¹¹ For example, see Tony Danker, Thomas Dohrmann, Nancy Killefer, and Lenny Mendonca, “How Can American Government Meet its Productivity Challenge?,” McKinsey and Company, July 2006, p. 4, at http://www.mckinsey.com/aboutus/mckinseynews/americangovt_prodchallenge.asp.

¹² Office of the President-elect (Obama), “President-elect Names Nancy Killefer as Chief Performance Officer,” press release, January 7, 2009.

¹³ “Press Conference with President-elect Barack Obama; Announcement of Nancy Killefer for Chief Performance Officer,” transcript, *Federal News Service*, January 7, 2009, at <http://www.nexis.com> (subscription required).

¹⁴ The position was established by P.L. 101-576 (1990). Many of the DDM’s statutory functions are codified at 31 U.S.C. § 503.

¹⁵ For an analysis of some “chief officers” and related councils, see CRS Report RL32388, *General Management Laws: Major* (continued...)

The DDM also has statutory responsibility—subject to the direction and approval of the OMB Director—to “coordinate and supervise the general management functions of [OMB].”¹⁶ This responsibility refers to supervision of statutorily established offices in OMB that focus on mission-support functions. These so-called “statutory offices” include the Office of Information and Regulatory Affairs (OIRA), the Office of Federal Procurement Policy (OFPP), the Office of Federal Financial Management (OFFM), and the Office of Electronic Government (E-Gov Office).¹⁷ OMB has no statutory office that covers functions including program evaluation, performance measurement, or general management, leaving these functions to the DDM.

Statements at Confirmation Hearings for OMB Officials

Additional details about the Administration’s performance agenda emerged at early Senate confirmation hearings. On January 13 and 14, 2009, the Committees on the Budget and on Homeland Security and Governmental Affairs held hearings to consider the anticipated nominations of Peter Orszag to be OMB Director and Rob Nabors to be OMB Deputy Director. Director-designate Orszag’s prepared testimony before the Budget Committee shared some of the incoming Obama Administration’s “initial thinking about priorities” relating to government performance and transparency.¹⁸ He covered several mission-support functions, including performance, information technology (IT), human capital, and regulation. With regard to performance, he said that under the CPO’s leadership, the Administration

will create a set of performance metrics that are outcome-oriented and in line with public expectations, as well as a central repository of performance data that will be available to departments and agencies, Congress, and the general public.... We also plan to build a team of management experts within OMB who will work with individual agencies to improve the skills of their workforce. We will launch pilot programs with individual agencies to serve as demonstration projects through which we can test our approaches to improve program effectiveness and efficiency [and] share best practices....¹⁹

In subsequent questioning, Senator Robert Menendez asked Deputy Director-designate Nabors how he foresaw OMB and CPO-designate Killefer evaluating programs and working with Congress on appropriations. Mr. Nabors responded by emphasizing transparency and a commitment to consulting with Congress and stakeholders about how to evaluate programs.²⁰

(...continued)

Themes and Management Policy Options, by Clinton T. Brass.

¹⁶ 31 U.S.C. § 503(b)(1).

¹⁷ The CFO Act’s legislative history notes that the heads of OMB’s statutory offices, such as the administrators of OIRA and OFPP, “are statutorily required to report to the Director [of OMB],” but that the legislation establishing the DDM position would provide flexibility in this regard. A key committee report said the legislation suggests that the reporting relationship of the heads of the statutory offices to the OMB Director “be through the new [DDM position].” See U.S. Congress, House Committee on Government Operations, *Chief Financial Officer Act of 1990*, report to accompany H.R. 5687, 101st Cong., 2nd sess., October 6, 1990, H.Rept. 101-818 (Washington: GPO, 1990), p. 18.

¹⁸ U.S. Congress, Senate Committee on the Budget, *Nomination of Hon. Dr. Peter Orszag, of Massachusetts, to be Director of OMB and The Nomination of Robert Nabors, of New Jersey, to be the Deputy Director of OMB*, 111th Cong., 1st sess., January 13, 2009, S.Hrg. 111-1 (Washington: GPO, 2009).

¹⁹ *Ibid.*, p. 23.

²⁰ Nabors’ response suggested how the Administration might approach the issue. Specifically, he said that he saw the CPO as a partner with OMB officials and Congress in developing metrics of program effectiveness, and that there would be “more openness in terms of sitting down with the affected parties and with other stakeholders to determine exactly how programs should be measured.” Nabors contrasted this approach with his perception of how previous administrations evaluated programs, where “there really has not been enough interaction with other stakeholders, including the Congress, as to exactly how programs should (continued...)”

Messrs. Orszag and Nabors appeared the next day before the Senate Committee on Homeland Security and Governmental Affairs.²¹ Several topics in the hearing related to the Administration's performance agenda, including execution of economic stimulus legislation, assessments of program effectiveness, large IT projects, and Administration proposals to terminate programs. In response to a question about the incoming Administration's plans for gauging a program's effectiveness, Director-designate Orszag commented on the George W. Bush Administration's Program Assessment Rating Tool (PART). The Bush Administration previously had described the PART—a multi-attribute questionnaire that resulted in a score ranging from 0 to 100—as rating a program's "overall effectiveness."²² Mr. Orszag responded to the hearing question by saying the PART was "not particularly effective," because most federal officials do not know about it and, of those who are knowledgeable about it, "most do not use it."²³ The lack of use, he asserted, stemmed from the PART's development "without consultation with the Congress and with the agencies," and because the PART was "too focused on process rather than outcomes." Commenting on the PART, Deputy Director-designate Nabors said that as a former staff director of the House Appropriations Committee, he was "routinely asked what ... [he] thought about various PART scores."²⁴ He said that the PART was not helpful to appropriations or authorizing committees, because, among other things, congressional and other stakeholders were not involved in specifying the "appropriate measures by which a program's success or failure can be determined." In addition, he said a bad evaluation might be grounds for eliminating a program, but it also could be grounds for attempting to fix it.

Post-Inauguration Agenda and Activities

Inauguration and Killefer Withdrawal

On January 20, 2009, President Obama gave his inaugural address. A section of the speech emphasized assessing whether a government activity "works," ending activities that do not "work," managing in an accountable way, correcting deficiencies, and increasing transparency.²⁵ More detail about the CPO's roles and Administration's performance agenda was expected in President Obama's first set of budget proposals for FY2010. Ms. Killefer's withdrawal on February 3, 2009, however, may have affected the Administration's plans. In a letter that the White House posted online, Ms. Killefer asked the President to "withdraw my name from consideration," citing a "personal tax issue" that might distract from her duties as CPO.²⁶ The Obama Administration would not announce another appointee for CPO and DDM until

(...continued)

be evaluated," and that "too often those decisions have been made behind closed doors." Quotations are from *ibid.*, p. 52.

²¹ U.S. Congress, Senate Committee on Homeland Security and Governmental Affairs, *Nomination of Peter R. Orszag*, 111th Cong., 1st sess., January 14, 2009, S.Hrg. 111-549 (Washington: GPO, 2010); and *Nomination of Robert L. Nabors II*, 111th Cong., 1st sess., January 14, 2009, S.Hrg. 111-440 (Washington: GPO, 2010).

²² See **Box 2** for more on the PART. For information about the PART's initial design and implementation, see CRS Report RL32663, *The Bush Administration's Program Assessment Rating Tool (PART)*, by Clinton T. Brass.

²³ U.S. Congress, Senate Committee on Homeland Security and Governmental Affairs, *Nomination of Peter R. Orszag*, 111th Cong., 1st sess., January 14, 2009, S.Hrg. 111-549 (Washington: GPO, 2010), p. 24.

²⁴ U.S. Congress, Senate Committee on Homeland Security and Governmental Affairs, *Nomination of Robert L. Nabors II*, 111th Cong., 1st sess., January 14, 2009, S.Hrg. 111-440 (Washington: GPO, 2010), p. 8.

²⁵ U.S. President (Obama), *The White House Blog*, "President Barack Obama's Inaugural Address," [January 20, 2009], posted January 21, 2009, at <http://www.whitehouse.gov/blog/inaugural-address/>.

²⁶ The White House, "Letter from Nancy Killefer to President Obama," February 3, 2009, at http://www.whitehouse.gov/the_press_office/LetterfromNancyKilleferToPresidentObama/.

April 2009. Nevertheless, more details about the CPO's role and Administration's agenda emerged shortly, as the Administration pursued efforts that reportedly were intended to set the stage for the CPO.

Obama Administration Budget Overview for FY2010

On February 26, 2009, the Administration released an initial budget overview for FY2010.²⁷ A more detailed submission of proposals was expected later in the spring.²⁸ In the meantime, the overview described some of the new Administration's priorities and plans regarding management and performance, including topics related to the CPO's duties. Among other things, the Administration proposed to "eliminate, cut, or place under intensive review" programs in the U.S. Department of Education that "are not helping to improve educational outcomes for students."²⁹ The Administration said these programs "lack strong evidence to justify taxpayer support and... in many cases, could be funded in competitive funding streams that require evidence of results." The Administration also said it is "critical to invest in learning which programs are effective and in growing the ones that are." Consequently, it proposed to "[increase] funding for rigorous evaluation ... for education research." These education proposals continued a focus on evaluation policy that inspired controversy during the previous Administration, regarding how to define and use evidence when making policy, and what the term *rigorous* should mean in that context.³⁰

The overview also argued for "investing taxpayer dollars in efforts and programs with proven records of success and reallocating or cutting programs that do not work or whose benefits are not worth their cost."³¹ In support of that objective, the Administration said it had begun an "exhaustive line-by-line review of the Federal Budget." The Administration would release related proposals in the full FY2010 Budget and in subsequent years. Meanwhile, the document identified several cuts and savings proposals that would be articulated in more detail in the Administration's full budget submission.

Under the heading "Making Government More Effective," the Administration identified other strategies it would pursue.³² Some were consistent with, or identical to, more detailed statements from presidential transition and campaign documents.³³ The document said the President "will meet regularly with cabinet officers to review the progress their agencies are making toward meeting performance improvement targets." However, the document no longer said the CPO would report directly to the President. The heading also covered other management-related topics, including procurement and technology.

²⁷ U.S. Executive Office of the President, Office of Management and Budget (hereafter OMB), *A New Era of Responsibility: Renewing America's Promise* (Washington: GPO, February 26, 2009).

²⁸ Recent Presidents have delayed submitting their full budget proposals until some time after they assumed office. For discussion, see CRS Report RS20752, *Submission of the President's Budget in Transition Years*, by Robert Keith.

²⁹ OMB, *A New Era of Responsibility: Renewing America's Promise*, p. 25.

³⁰ For discussion, see CRS Report RL33301, *Congress and Program Evaluation: An Overview of Randomized Controlled Trials (RCTs) and Related Issues*, by Clinton T. Brass, Erin D. Williams, and Blas Nuñez-Neto; and Stewart I. Donaldson, Christina A. Christie, and Melvin M. Mark, eds., *What Counts as Credible Evidence in Applied Research and Evaluation Practice?* (Los Angeles: Sage Publications, 2009).

³¹ OMB, *A New Era of Responsibility: Renewing America's Promise*, p. 34.

³² *Ibid.*, pp. 38-41.

³³ Among these were plans to address redundancy in administrative functions; create a team in the White House to work with agency leaders and OMB; and establish a CPO position to work with agencies to set performance targets. With regard to the Bush Administration's PART, the Obama Administration said it would, among other things, "fundamentally reconfigure" the PART; "open up the insular performance measurement process" to Congress and the public; and assess individual programs in the context of other programs that are "serving the same population or meeting the same goals."

Announcement of Zients Nomination

On April 18, 2009, President Obama announced his intention to appoint and nominate individuals for CPO and another new position, Chief Technology Officer (CTO).³⁴ For CPO, the President selected Jeffrey D. Zients, who also would be nominated to serve as DDM at OMB.³⁵ The announcement did not go into detail, but said Mr. Zients “will work to streamline processes, cut costs, and find best practices throughout our government.”³⁶ On the same day, OMB Director Orszag provided more information about the CPO’s duties. Mr. Zients would “lead the President’s efforts on contracting and procurement reform, improve government productivity by helping to root out error and waste, build a performance agenda across government, and enhance the transparency of the government’s finances.” In addition, Mr. Zients would work closely with the CTO, Mr. Aneesh Chopra, and the Federal Chief Information Officer [CIO], Vivek Kundra.³⁷ The Administration’s inclusion of contracting, procurement, and financial management within the CPO’s portfolio was new. The inclusion likely reflected the DDM’s statutory role to coordinate and supervise OMB’s “general management functions,” and, therefore, the activities of OMB’s management-related statutory offices.³⁸

Detailed Budget Proposals for FY2010 and Related Policies

On May 7 and 11, 2009, the Obama Administration submitted to Congress more detailed components of the President’s budget request for FY2010. The contents of some documents related closely to the previous announcements of the Administration’s performance agenda and CPO’s responsibilities.

Proposals for Program Terminations and Reductions

On May 7, 2009, the Administration issued a document that proposed program “terminations, reductions, and savings” (hereafter *FY2010 TRS Document*).³⁹ The Administration characterized the *FY2010 TRS Document* as identifying “programs that do not accomplish their intended objectives, are not efficient, or that replicate efforts being completed by another initiative and recommends these programs for either termination or reduction.”⁴⁰ By the Administration’s count, the document proposed 121 cuts or

³⁴ The White House, Office of the Press Secretary, “Weekly Address: President Obama Discusses Efforts to Reform Spending, Government Waste; Names Chief Performance Officer and Chief Technology Officer,” press release, April 18, 2009, at http://www.whitehouse.gov/the_press_office/Weekly-Address-President-Obama-Discusses-Efforts-to-Reform-Spending/. For an overview of the CTO position, see CRS Report R40150, *A Federal Chief Technology Officer in the Obama Administration: Options and Issues for Consideration*, by John F. Sargent Jr.

³⁵ At the time of the announcement, Mr. Zients was founder and managing partner of Portfolio Logic, a Washington, DC, private equity firm. Previously, he served in executive positions in two business research and executive education firms, the Advisory Board Company and the Corporate Executive Board, and as a management consultant.

³⁶ The White House, Office of the Press Secretary, “Weekly Address: President Obama Discusses Efforts to Reform Spending, Government Waste; Names Chief Performance Officer and Chief Technology Officer.”

³⁷ OMB, OMB Director Peter R. Orszag, “Move Over R2, CPO is Here,” April 18, 2009, at <http://www.whitehouse.gov/omb/blog/09/04/18/MoveOverR2CPOisHere/>. The CIO position refers to the Obama Administration’s non-statutory title for the Administrator of OMB’s E-Gov Office, a presidentially appointed position.

³⁸ 31 U.S.C. § 503(b)(1). This report does not focus in detail on the core activities of OMB’s statutory offices, including regulatory review (OIRA), information policy (OIRA), financial management (OFFM), procurement policy (OFPP), and information technology investment (shared between OIRA and the E-Gov Office).

³⁹ OMB, *Terminations, Reductions, and Savings—Budget of the U.S. Government, Fiscal Year 2010* (Washington: 2009), at <http://www.gpoaccess.gov/usbudget/fy10/browse.html>.

⁴⁰ OMB Web page, “Terminations, Reductions and Savings,” at <http://www.whitehouse.gov/omb/budget/TRS/>. The George W. Bush Administration released similar documents for President Bush’s FY2006, FY2007, FY2008, and FY2009 budget proposals.

restructurings totaling “approximately \$17 billion.” Each item in the volume included a brief justification. Notably, the document made no reference to the Bush Administration’s PART. According to one media outlet, OMB Director Orszag was asked in connection to the *FY2010 TRS Document* whether the Administration used the Bush Administration’s PART or something similar to identify inefficient programs.⁴¹ He reportedly responded that OMB is “in the process of overhauling the performance metrics system” and, in addition, that “overhauling the way [OMB] evaluates program performance will be a top priority for Jeffrey Zients.” Furthermore, he reportedly said that OMB employees already were working to ensure the process could move along quickly once Mr. Zients took office. Indications of what the process might involve became public only days later.

Themes for a “New Management and Performance Agenda”

On May 11, 2009, OMB released the *Analytical Perspectives* volume of the President’s budget.⁴² The volume did not mention the CPO, but included more detail about the Administration’s performance-related plans. In a section titled “Building a High-Performing Government,” the document outlined the first in a succession of evolving frameworks that the Administration would release. In this case, the Administration put forward six “themes” for a “new management and performance agenda” (see **Box 1**).⁴³ At the time, it appeared that all six themes might fall under the statutory responsibilities of OMB’s DDM, depending on how the OMB Director assigned duties to subordinates. At a later hearing, OMB Director Orszag said the CPO position’s role would embrace the entire agenda.⁴⁴ Prior announcements and the *Analytical Perspectives* volume suggested, nevertheless, that day-to-day leadership in pursuit of the last four or five themes would come from other officials.

Box 1: “Themes” for a “New Management and Performance Agenda”

- I. Putting Performance First: Replacing PART with a New Performance Improvement and Analysis Framework
- II. Ensuring Responsible Spending of Recovery Act Funds
- III. Transforming the Federal Workforce
- IV. Managing Across Sectors
- V. Reforming Federal Contracting and Acquisition
- VI. Transparency, Technology, and Participatory Democracy

Announcement of a “Performance Improvement and Analysis Framework”

Of the six themes, the first, relating to a “new performance improvement and analysis framework,” received the most extensive discussion.⁴⁵ This theme focused primarily on issues related to program evaluation and performance measurement. In the context of past announcements, these emphases suggested the CPO would play a primary role in establishing and implementing the new framework. To set the stage for the framework, the Administration discussed prior developments during the William Clinton and George W. Bush Administrations and cited three in particular: enactment of the Government

⁴¹ Elizabeth Newell, “Administration Short on Details About Program Review Process,” *GovExec.com*, May 7, 2009, at <http://www.govexec.com/dailyfed/0509/050709e1.htm>.

⁴² OMB, *Analytical Perspectives, Budget of the U.S. Government, FY2010* (Washington: GPO, 2009).

⁴³ *Ibid.*, pp. 9-12.

⁴⁴ U.S. Congress, House Committee on Appropriations, Subcommittee on Financial Services and General Government, *Financial Services and General Government Appropriations for 2010*, hearings, 111th Cong., 1st sess., part 8 [May 20, 2009] (Washington: GPO, 2009), p. 74.

⁴⁵ OMB, *Analytical Perspectives, Budget of the U.S. Government, FY2010*, pp. 9-10. The other five themes appeared to be more closely supervised by other presidential appointees who focus on IT, financial management, contracting, and human resources (i.e., the CTO, heads of OMB’s statutory offices, and the Director of OPM), and are not discussed in this memorandum.

Performance and Results Act of 1993 (GPRA), implementation the Bush Administration's PART, and issuance of the Bush Administration's Executive Order (E.O.) 13450. (See **Box 2** for highlights.)

Box 2: Recent Developments Related to Government Performance and Evaluation

Enactment and implementation of GPRA (active from 1993-present).⁴⁶ GPRA requires executive agencies to develop five-year strategic plans, annual performance plans, and annual program performance reports; to consult with Congress and stakeholders when developing strategic plans, but not when developing annual plans or indicators; and to specify in these plans and reports, among other things, mission statements, general goals and objectives, performance goals, performance indicators, and some descriptions of program evaluations. As drafted and implemented, GPRA has emphasized using performance information during the budget formulation process, but not necessarily when managing agencies or programs. The law has been tightly bound with the process of formulating presidential budget proposals, which occurs largely outside of public view. Annual plans, goals, performance indicators, and program evaluation priorities are formulated within discretion available to agencies and the President and are not required by GPRA to be revised to reflect congressional budget decisions. The law also has emphasized *performance measurement* considerably more than *program evaluation*. Program evaluation and performance measurement are different but often are used for complementary purposes; see **Box 3** (p. 12) for explanation of these terms.

Implementation of the Bush Administration's non-statutory PART initiative (active 2003-2009). For approximately 1,000 Bush Administration-defined "programs," the PART used a questionnaire to generate an index score. The score ranged from 0-100 and resulted in a corresponding rating: "effective" for a score between 85-100; "moderately effective", 70-84; "adequate", 50-69; and "ineffective", 0-49. When OMB and agencies disagreed on how to assess program performance (e.g., how to define "success") or when OMB believed a program lacked performance data, OMB gave another designation: "results not demonstrated." OMB said the PART was intended to "formalize" the Administration's efforts to achieve "budget and performance integration," which the Administration said "stresses making budget decisions based on results."⁴⁷

Issuance of Bush Administration E.O. 13450 (active November 13, 2007-present).⁴⁸ This E.O. adopts and codifies some aspects of the PART. It also centralizes decision making in OMB and the White House about many aspects of agencies' GPRA-related efforts. The E.O. requires an agency head to designate a career or noncareer member of the Senior Executive Service (SES) as "performance improvement officer" (PIO). A PIO is to work under the direction of the agency head and the OMB DDM. The E.O. also establishes a Performance Improvement Council (PIC) within OMB, which is to be directed by the DDM and composed of OMB-selected PIOs. Referring to the PART and its process, a Bush Administration OMB official reportedly said the E.O. was issued as "an effort to sustain what we think is valuable beyond this administration."⁴⁹

The Obama Administration briefly described its perceptions of these three developments, especially focusing on the PART. Congress and previous Administrations had made "some meaningful progress, but much more could be done." For example, the Obama Administration said the PART helped to establish performance measures, "[b]ut it has been less successful in encouraging the actual use of performance measurement as a performance improvement tool."⁵⁰ In support of this statement, the Administration cited a 2007 Government Accountability Office (GAO) survey. The survey results, which were published in 2008, indicated that among federal managers familiar with the PART, only "a minority—26 percent—indicated that PART results are used in decision making, and 14 percent viewed PART as improving performance."⁵¹ The survey finding may not be surprising, however. OMB guidance explicitly told agencies that a program's PART questionnaire should include performance goals that "provide

⁴⁶ P.L. 103-62, 107 Stat. 285 (1993), and later amended. Substantially amended in the 111th Congress, the new version of the law has not yet been implemented. See this memorandum's final section for discussion.

⁴⁷ Quotations are from OMB, *Budget of the United States Government, Fiscal Year 2004, Performance and Management Assessments* (Washington: GPO, 2003), p. 9. Some of the Bush Administration's rhetorical emphasis behind the PART changed in July 2007, when the name of the umbrella initiative was changed from "budget and performance integration" to "performance improvement." See The White House, Results.gov (archived website), "The President's Management Agenda," link entitled "What's in a name?" (July 30, 2007), at <http://georgewbush-whitehouse.archives.gov/results/agenda/budgetperformance.html>.

⁴⁸ E.O. 13450, "Improving Government Program Performance," 72 *Federal Register* 64519, November 13, 2007.

⁴⁹ Attributed to Robert Shea in Robert Brodsky, "Commanding Performance," *National Journal*, April 19, 2008, p. 65.

⁵⁰ OMB, *Analytical Perspectives, Budget of the U.S. Government, FY2010*, p. 9.

⁵¹ U.S. Government Accountability Office (hereafter GAO), *Government Performance: Lessons Learned for the Next Administration on Using Performance Information to Improve Results*, GAO-08-1026T, July 24, 2008, p. 9.

information that helps make budget decisions,” but need not include “performance goals to improve the management of the program.”⁵²

Box 3: Program Evaluation, Performance Measurement, and Logic Models

Like other subjects of federal policy, performance-related policy often involves special terms. Many terms that are used in lawmaking, policy implementation, and oversight may sound similar but often have distinct meanings.

Program evaluation is broader in scope than performance measurement. See **Figure 1** (p. 13), which illustrates that performance measurement activities may be a subset of some kinds of program evaluation. Program evaluation uses one or more systematic, analytic methods to answer questions about how well programs are working, whether they are achieving their objectives or causing unintended consequences, why they are performing as they are, and how they can be improved.⁵³ Evaluation may encompass one or more studies, or may be pursued as an ongoing activity. Program evaluations also may be conducted to assess nonfederal activities, such as when an industry self-regulates. Program evaluation “types” may be categorized in many ways.⁵⁴ In areas of complexity, there is increasing consensus that a single study of any type or method is rarely sufficient to reliably support decision making.⁵⁵ Rather, program evaluation often is viewed as a cumulative process of using multiple studies, methods, and analyses to better inform learning and decision making.⁵⁶ *Systematic reviews*, which sometimes are called *evaluation syntheses*, draw conclusions about the current state of knowledge from many evaluations.

Performance measurement is much narrower in focus and typically refers to periodic monitoring and reporting of data that are related to government policies and programs. In isolation, performance measurement data often are viewed as managerially useful. However, performance measures, which sometimes are called *indicators*, do not necessarily reflect how well a government program or policy is working, because they may be significantly driven by “external factors” other than the program or policy. Types of performance measures have been categorized in many ways. They may correspond to a program’s *activities*, such as actions taken in the process of completing a good or service; *outputs*, such as completed goods or services; *intermediate outcomes*, such as modified behaviors or new knowledge; and *end outcomes*, such as an improved condition or altered status for an inherently valued topic like poverty or economic well-being.⁵⁷

A *logic model* may be used to integrate program evaluation and performance measurement. A logic model shows how multiple inputs, activities, and outputs are expected to influence outcomes in a causal chain. See **Figure 2** (p. 13) for the diagram of a simple logic model. In a job training program, for example, an agency may offer classes to clients. The classes might be considered to be an *activity*. The classes culminate with clients who complete a course. These course completions might be called an *output*. If the clients learn new skills—an *intermediate outcome*—they may have a better chance of finding employment or increasing earnings, which might be viewed as *end outcomes*. Logic models may assist with understanding the relationships among activities and policy outcomes and thereby help determine how well programs work. These relationships sometimes are called *program theory*—that is, the theory behind how a program is intended or understood to achieve policy goals. Like other kinds of policy analysis, program evaluations, performance measurement, and logic models may raise more refined questions about how well a program is working or about the conditions that may be necessary for achieving success in certain circumstances. Therefore, evaluations, measurement, and logic models often complement each other when they inform policy making, oversight, or study.

⁵² OMB, “Performance Measurement Challenges and Strategies,” June 18, 2003, p. 4, at http://www.whitehouse.gov/sites/default/files/omb/part/challenges_strategies.pdf.

⁵³ In this context, the term *program* may be a policy, project, initiative, law, tax provision, regulation, or mix thereof.

⁵⁴ For two different presentations of types, see “Evaluation Dialogue Between OMB Staff and Federal Evaluators: Digging a Bit Deeper into Evaluation Science,” July 2006, Table 1 (slide 14 of PDF), http://www.whitehouse.gov/omb/assets/omb/performance/supp_materials/fed_eval_2006_slides.pdf; and GAO, *Performance Measurement and Evaluation: Definitions and Relationships*, GAO-05-739SP, May 2005.

⁵⁵ For discussion of analogous developments in medicine, see Committee on Reviewing Evidence to Identify Highly Effective Clinical Services, Institute of Medicine, Jill Eden, et al., eds., *Knowing What Works in Healthcare: A Roadmap for the Nation* (Washington: National Academies Press, 2008), p. 83.

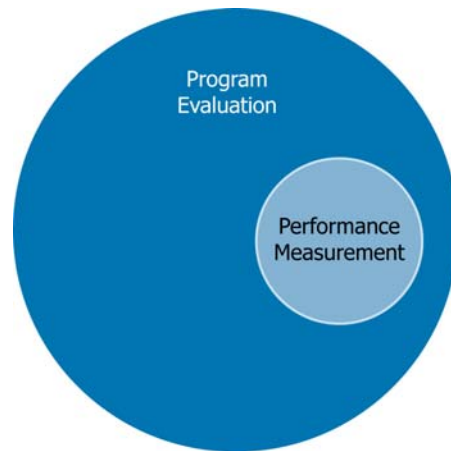
⁵⁶ CRS Report RL33301, *Congress and Program Evaluation: An Overview of Randomized Controlled Trials (RCTs) and Related Issues*, by Clinton T. Brass, Erin D. Williams, and Blas Nuñez-Neto.

⁵⁷ For discussion, see Harry P. Hatry, *Performance Measurement: Getting Results*, 2nd ed. (Washington: Urban Institute Press, 2006), pp. 26-27; and Lawrence B. Mohr, *Impact Analysis for Program Evaluation*, 2nd ed. (Thousand Oaks, CA: Sage Publications, 1995), pp. 15-18.

After setting the stage with discussion of these three developments, the Administration outlined how it would develop a new framework and related plans. Working along with the Performance Improvement Council (PIC; established by E.O. 13450), the Administration said it eventually would “fundamentally reconfigure how the Federal Government assesses program performance.” The PART would be replaced with a “performance improvement and analysis framework,” which would

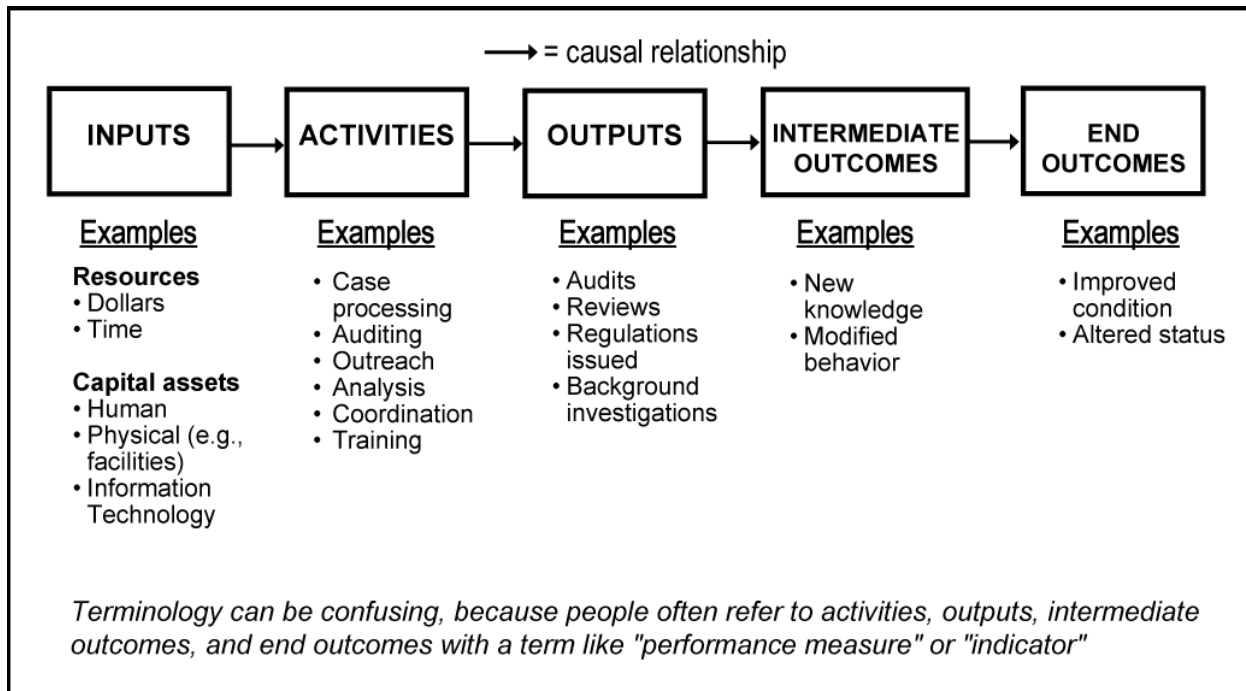
switch the focus from grading programs as successful or unsuccessful to requiring agency leaders to set priority goals, demonstrate progress in achieving goals, and explain performance trends. In order to break down silos, cross-program and cross-agency goals would receive as much or more focus as program-specific ones. In developing this new approach, the Administration will engage the public, Congress, and outside experts to develop a better and more open performance measurement process...⁵⁸

Figure 1. Program Evaluation vs. Performance Measurement



Source: CRS.

Figure 2. Basic Framework for A Logic Model: Evaluating How a Policy May Work



Source: CRS, adapted and modified from Hatry, p. 26 (2006) and United Way of America, *Measuring Program Outcomes: A Practical Approach* (1996), at <http://americallivesunited.com/Outcomes/Resources/MPO/model.cfm>.

⁵⁸ OMB, *Analytical Perspectives, Budget of the U.S. Government, FY2010*, p. 9.

Pursuit of Performance Improvement: Three Related Efforts

The Administration announced three efforts that would begin soon in pursuit of the performance improvement and analysis framework.⁵⁹ First, the Administration planned to identify “high priority goals.” As a first step in that process, the Administration said OMB would ask each major agency to “identify a limited set of high priority goals, supported by meaningful measures and quantitative targets, that will serve as the basis for the President’s meetings with cabinet officers to review their progress toward meeting performance improvement targets.” The Administration also said it would “identify ongoing opportunities to engage the public, stakeholders, and Congress in this effort.”

Second, the Administration said the framework also would “emphasize program evaluation.” The Administration said it would “conduct quality research evaluating the effectiveness of government spending,” making an analogy to what has been called “comparative effectiveness research” in health care.⁶⁰ The *Analytical Perspectives* document did not elaborate further. However, subsequent statements from OMB about “evidence-based policy” and “rigorous evidence” suggested a possible emphasis on using particular evaluation methodologies, potentially including randomized controlled trials (RCTs),⁶¹ to ascertain the impact of government policy interventions on policy outcomes.⁶² RCTs were the subject of some controversy during the George W. Bush Administration.⁶³ Later, OMB Director Orszag characterized this effort as “evidence-based policy” and related it to the President’s budget proposals for program terminations and reductions, saying the targeted programs “have been determined to be ineffective, duplicative, or obsolete.”⁶⁴

⁵⁹ Unless otherwise noted, quotations are taken from *ibid.*, pp. 9-10.

⁶⁰ For background on this topic in health care, see CRS Report RL34208, *Comparative Clinical Effectiveness and Cost-Effectiveness Research: Background, History, and Overview*, by Gretchen A. Jacobson.

⁶¹ OMB, OMB Director Peter R. Orszag, “Building Rigorous Evidence to Drive Policy,” June 8, 2009, at <http://www.whitehouse.gov/omb/blog/09/06/08/BuildingRigorousEvidenceToDrivePolicy/>. Director Orszag’s blog entry linked to an advocacy group, the Coalition for Evidence-Based Policy, that primarily has emphasized its views on the merits of RCTs and disadvantages of other evaluation methods. See Coalition for Evidence-Based Policy, “Mission & Activities,” at http://coalition4evidence.org/wordpress/?page_id=6.

⁶² An RCT is a form of “impact evaluation” that quantitatively estimates the extent to which a policy causes a change in an outcome of interest, compared to what would have happened without the policy. For a discussion of impact evaluations, see **Figure 3** and related discussion later in this memorandum. An “outcome of interest” is a policy goal that one or more stakeholders care about (e.g., unemployment rate). There can be many outcomes of interest related to a program, but there may not be consensus on which are most important.

⁶³ Topics that were subjects of debate included how to define “success” for a program or policy; how the term “effectiveness” is used and defined; what kinds of information and analysis constitute “evidence”; what constitutes “rigorous” evaluation; what methodology or methodologies are capable of producing “evidence”; how evidence should be presented and used; and what kinds of information and analysis should inform learning, management, and policy making. For discussion, see CRS Report RL33301, *Congress and Program Evaluation: An Overview of Randomized Controlled Trials (RCTs) and Related Issues*, by Clinton T. Brass, Erin D. Williams, and Blas Nuñez-Neto.

⁶⁴ Written testimony of OMB Director Orszag, in U.S. Congress, House Committee on Appropriations, Subcommittee on Financial Services and General Government, *Financial Services and General Government Appropriations for 2010*, hearings, 111th Cong., 1st sess., part 8 [May 20, 2009] (Washington: GPO, 2009), p. 111. He also mentioned *cost-benefit analysis* at the hearing as a methodology capable of “demonstrat[ing] results.” Under cost-benefit analysis, a person seeks to estimate the costs and benefits of a policy option for different actors, denominating both costs and benefits in dollar terms, although some costs and benefits may not be possible to quantify. Cost-benefit analysis then shows the distribution of benefits and costs among these actors and provides a perspective on whether the total benefits for society justify the total costs. As an input to that assessment, the analysis may produce a computation of “net benefits” or “net costs.” Many economists are uncomfortable with cost-benefit analysis, if it computes overall net benefits or costs for an option but does not take into account distributional concerns among winners and losers, or if the analysis is used to justify decisions without making side payments to compensate any losers. See Edward M. Gramlich, *A Guide to Benefit-Cost Analysis*, 2nd ed. (Englewood Cliffs, NJ: Prentice Hall, 1990), pp. 30-33.

Third, the Administration would develop further options for the framework. In the coming months, the Administration said it would work with agencies and the PIC to develop options for five evaluation-related topics. The *Analytical Perspectives* volume did not indicate whether Congress and stakeholders would be involved. The topics included (1) establishing a “comprehensive program and performance measurement system that shows how Federal programs link to agency and Government-wide goals”; (2) changing “program assessment and performance measurement processes to emphasize the reporting of performance trends, explanations for the trends, mitigation of implementation risks, and plans for improvement with accountable leads”; (3) “streamlining” GPRA- and PART-related reporting to reduce workload on agencies and OMB; (4) changing how evaluation-related information is presented to Congress, the public, and stakeholders, both in agency reports and on OMB’s ExpectMore.gov website;⁶⁵ and (5) beginning a “comprehensive research program to study the comparative effectiveness of different program strategies to ensure that programs achieve their ultimate desired outcomes.”

Zients Statements About Plans at Confirmation Hearing

On June 10, 2009, the Senate Committee on Homeland Security and Governmental Affairs held Mr. Zients’ confirmation hearing for the DDM position.⁶⁶ Mr. Zients said that, if confirmed, he would “focus on the five areas the President has emphasized,” including (1) creating an outcome-oriented measurement system; (2) helping to transform the Federal workforce; (3) improving acquisition practices and having the right mix of public and private labor; (4) ensuring that funding from the American Recovery and Reinvestment Act (ARRA; P.L. 111-5) would be spent quickly and wisely; and (5) achieving “unprecedented transparency and accountability” throughout government operations.⁶⁷ On the subject of IT investments, Mr. Zients alluded to an effort to “create a dashboard which will allow us to see where projects are, spot problems early and get them ... back on track.”⁶⁸ With regard to program evaluation and performance measurement, Mr. Zients said the Administration would transition away from PART to a

collaborative approach, working with the stakeholders at the senior-most level to understand what matters, what are the overall goals, how are they being translated to ... operating plans? And then, what are the best handful of outcomes-based metrics to track progress across time? And that you have the flexibility in the system ... to look at problems ... not just by program, but across agency. So, I’m wary of anything that’s one size fits all. At the same time, I think the ultimate test is ... is the system being used by senior managers, senior leaders, senior stakeholders to make decisions?⁶⁹

This approach appeared to view “senior-most” stakeholders as the primary audiences for collaborating on definitions of “success,” and as the primary audiences for encouraging use of the resulting information. It

⁶⁵ The website was established the Bush Administration in early 2006 to house PART-related information and is available at <http://www.whitehouse.gov/omb/expectmore/index.html>.

⁶⁶ U.S. Congress, Senate Committee on Homeland Security and Governmental Affairs, *Nominations*, 111th Cong., 1st sess., June 10, 2009 (hereafter, *Zients Nomination Hearing*), at http://hsgac.senate.gov/public/index.cfm?FuseAction=Hearings.Hearing&Hearing_ID=19476ae-7f0c-403e-a1fe-f06160b2f7e7; transcript at <http://www.cq.com> (subscription required).

⁶⁷ The fourth and fifth themes that appeared in the President’s FY2010 budget appeared to be combined here. Written testimony of Jeffrey Zients, in *Zients Nomination Hearing*, at http://hsgac.senate.gov/public/_files/061009Zients.pdf.

⁶⁸ Testimony in *Zients Nomination Hearing*. This referred to creation of a website that would show information about projects’ status with regard to cost, schedule, and performance using “earned value management” (EVM). The website was established as “IT Dashboard” on June 30, 2009, at <http://it.usaspending.gov/>. For more on EVM, see CRS Report RL34257, *Earned Value Management (EVM) as an Oversight Tool for Major Capital Investments*, by Clinton T. Brass.

⁶⁹ Oral testimony in *Zients Nomination Hearing*. Mr. Zients added that “[s]ome programs probably aren’t conducive to metric-based tracking across time, ... and ... therefore, require longer term studies around effectiveness and efficiency.”

was unclear, however, how the Administration might approach the subject of more numerous operational metrics and analyses that often are used at lower levels of an agency to assist with management, learning, and research. In addition, Mr. Zients' testimony did not elaborate on how the Administration and agencies would provide for public transparency and participation in the process of developing goals and metrics.

Planning for High-Priority Performance Goals and FY2011 Budget

On June 19, 2009, the Senate confirmed Mr. Zients to be DDM by unanimous consent.⁷⁰ His confirmation set the stage for executive branch-wide implementation of an OMB memorandum that was issued shortly after his confirmation hearing. On June 11, 2009, OMB Director Orszag issued a memorandum to heads of agencies that provided more details about how the Administration would establish the framework that it previously announced in the FY2010 President's budget.⁷¹ The memorandum also assigned several "deliverables" to agencies, including that they identify near-term "high-priority performance goals" and include certain performance information and termination proposals in FY2011 budget submissions.⁷²

Development of a "Management and Performance Framework"

The June 2009 memorandum described a forthcoming process to establish a "management and performance framework." The memorandum suggested the framework might encompass many of the presidential priorities that Mr. Zients outlined at his confirmation hearing.

Over the next several months, OMB also will work with Congress, interagency management councils, experts in Federal management policy, Federal employees, and other key stakeholders to craft a broad management and performance framework that will achieve near-term priorities and overcome long standing management challenges. This effort will include addressing the high-priority performance goals ... and will help inform budget decisions. The framework also will further promote reforms to Federal procurement and financial management processes, improve the openness and transparency of government, and maintain a focus on implementation of [ARRA].⁷³

Publication in August 2009 of OMB's updated version of *Circular No. A-11*, which guides agencies annually on how to develop budget submissions and comply with GPRA, shed little additional light on the Administration's framework other than removing references to the Bush Administration's PART.⁷⁴

High-Priority Performance Goals (HPPGs)

The June OMB memorandum provided detailed instructions to agencies on how to identify "high-priority performance goals" (HPPGs) that could be achieved in 12-24 months. Agencies would have six weeks,

⁷⁰ "Executive Calendar," *Congressional Record*, daily edition, vol. 155, part 93 (June 19, 2009), pp. S6840-S6841. On July 1, 2009, the White House released a list of White House Office employees. The list did not include Mr. Zients, which indicated that the CPO is not a White House position, but instead is a non-statutory, additional title for OMB's DDM position. The White House, "Annual Report to Congress on White House Staff," July 1, 2009; at <http://www.whitehouse.gov/blog/Annual-Report-to-Congress-on-White-House-Staff-2009/>.

⁷¹ OMB, "Planning for the President's Fiscal Year 2011 Budget and Performance Plans," memorandum, M-09-20, June 11, 2009 (hereafter *OMB June 2009 Memo*), at http://www.whitehouse.gov/omb/memoranda_2009/.

⁷² Other deliverables focused on the federal hiring process, employee survey findings, and employee wellness.

⁷³ *OMB June 2009 Memo*, pp. 1-2.

⁷⁴ OMB, *Circular No. A-11, Preparation, Submission, and Execution of the Budget*, August 2009, Part 6, at http://www.whitehouse.gov/sites/default/files/omb/assets/a11_current_year/a_11_2009.pdf.

until July 31, 2009, to submit candidate goals to OMB. When developing the goals, agencies would be required to “engage with and consider input” from OMB, policy councils in the Executive Office of the President (EOP), congressional authorizers and appropriators, and “agency program and management leadership, front-line employees, external partners, the public, and other stakeholders.” The memorandum did not explain how input would be solicited or whether it would be made publicly available.

OMB outlined numerous guidelines for the HPPGs (see **Box 4**). In combination, the guidelines appeared to be difficult to satisfy or reconcile. For example, OMB required that the HPPGs be “highly valued by the public,” which suggested that many HPPGs might be end outcomes (see **Figure 2**, p. 13). However, factors that are separate from a government program, which sometimes are called “external factors,” such as overall economic conditions, also may influence whether end outcomes are achieved.⁷⁵ Therefore, simple performance measurement of these end outcomes might leave unclear whether the program is contributing toward achievement of the goal, or if achievement or non-achievement instead is due to external factors that are separate from the program.⁷⁶ In addition, goals that are achievable and measurable in the short-term often tend to be more operational in character—for example, activities, outputs, or intermediate outcomes in a logic model, toward the left side of **Figure 2**—which might make it more difficult to select goals that are of high value to the public.

Box 4. OMB Guidance for HPPGs

OMB said the goals should be “near-term” (achievable in 12-24 months); highly valued by the public or reflecting achievement of “key agency missions”; authorized by law; and sufficiently funded by Congress in order to be achieved. OMB also said, however, that the goals should not be focused on internal agency management issues. The goals also would need to be quantifiable, “measureable in a timely fashion,” and chosen such that “outcomes ... can be clearly evaluated.” Acceptable goals would need to be difficult to achieve unless multiple challenges were overcome. Challenges could include resolution of “coordination, operational, or other implementation challenges including across multiple agencies,” or the need for a “concerted focus of agency resources.”

For each goal, the memorandum said an agency should define the problem being addressed and the goal to be accomplished. A goal would comprise one or more performance measures with “targets and timeframes.” In addition, the memorandum said an agency should identify programs inside and outside of the agency that contribute to achieving a goal; identify persons within and outside the agency who are responsible for policy direction, program management, and implementation; identify “who will be held accountable for whether the goal is achieved”; and outline the agency’s strategy to achieve the goal.

After the July 31 deadline, OMB reportedly was “working closely with agencies to tweak performance goals” that had been sent to OMB.⁷⁷ Later, Mr. Zients reportedly said the goals would be posted online “after they are finalized as part of the fiscal 2011 budget process.”⁷⁸ It is not clear how agencies solicited stakeholder input when developing the goals, including from Congress.

Aspects of this initiative appeared to be similar in some respects to previous presidential initiatives. For example, the Richard M. Nixon and Gerald R. Ford Administrations pursued “Management by Objectives” (MBO), under which agencies were instructed in 1973 to “select objectives which are of Presidential significance” and that are “capable of accomplishment in the coming year.”⁷⁹ In addition, the

⁷⁵ See **Figure 2** for an illustration of end outcomes and **Box 3** for more on external factors.

⁷⁶ See **Box 3** for discussion of how program evaluation is broader in scope and often more appropriate for addressing such questions than performance measurement.

⁷⁷ Elizabeth Newell, “Agencies Submit ‘Fresh, Innovative’ Performance Goals to OMB,” *GovExec.com*, August 4, 2009, at <http://www.govexec.com/dailyfed/0809/080409e1.htm>.

⁷⁸ Jason Miller, “OMB Wants to Change the Tone of Management,” *Federalnewsradio.com*, October 21, 2009, at <http://www.federalnewsradio.com/?nid=35&sid=1791266>.

⁷⁹ Chester A. Newland, “Policy/Program Objectives and Federal Management: The Search for Government Effectiveness,” *Public Administration Review*, vol. 36 (January/February 1976), p. 21.

Clinton Administration's National Performance Review recommended in 1993 that the "President should develop written performance agreements with department and agency heads."⁸⁰ The HPPG effort also resembles an initiative in the United Kingdom under former Prime Minister Tony Blair. In this initiative, which was led by the Prime Minister's Delivery Unit (PMDU), departments were assigned targets for certain performance measures and received high level attention from the Prime Minister.⁸¹

Planning for the President's FY2011 Budget Proposals

The June 2009 OMB memorandum also began the process of developing the President's FY2011 budget proposals. According to previous Administration statements, several of the memorandum's topics were supposed to have been influenced by Mr. Zients. Among other things, the memorandum directed each agency to identify at least five potential "significant" terminations or reductions in its budget submission to OMB. Each potential termination or reduction was to include "analysis and evidence showing why the reductions are warranted."⁸² OMB also directed agencies to submit three alternative budget requests, reflecting different funding scenarios, with corresponding estimates of the "performance" associated with each request. For these requests, OMB directed agencies to highlight "methodologies used to allocate base funding" and specifically cited the use of cost-benefit analysis. Furthermore, the memorandum said agency requests for funding for a specific project or location "should provide a clear merit basis for inclusion in the budget" but did not explain what would constitute "merit."⁸³

Developing a New Performance Management System

Fall 2009 Congressional Hearings

From September 2009 through the calendar year's end, several developments suggested how the Obama Administration's framework might operate. For example, Mr. Zients testified at several congressional hearings about the Administration's plans for developing a new "performance management" framework. At a hearing on September 24, 2009, for example, he announced the appointment of Shelley H. Metzenbaum as his deputy with responsibility for performance-related topics.⁸⁴ After the 2008 presidential election, Ms. Metzenbaum wrote a 66-page report that recommended what the Obama Administration should do regarding "performance management" (see **Box 5** for highlights).⁸⁵

⁸⁰ Office of the Vice President (Gore), *From Red Tape to Results: Creating a Government That Works Better & Costs Less, Report of the National Performance Review* (Washington: GPO, 1993), pp. 75-77, and *From Red Tape to Results: Creating a Government That Works Better & Costs Less, Mission-Driven, Results-Oriented Budgeting, Accompanying Report of the National Performance Review* (Washington: GPO, 1993), pp. 9-14.

⁸¹ The PMDU effort received mixed reviews. See Michael Barber, *Instruction to Deliver: Tony Blair, Public Services, and the Challenge of Achieving Targets* (London: Politico's, 2007); and Christopher Hood, "Gaming in Targetworld: The Targets Approach to Managing British Public Services," *Public Administration Review*, vol. 66 (July/August 2006), pp. 515-521.

⁸² *OMB June 2009 Memo*, p. 4.

⁸³ Some observers have called presidential budget requests that specify funds for a specific project or location a form of executive branch earmarking. For discussion, see CRS Report RL34648, *Bush Administration Policy Regarding Congressionally Originated Earmarks: An Overview*, by Clinton T. Brass, Garrett Hatch, and R. Eric Petersen.

⁸⁴ U.S. Congress, Senate Committee on Homeland Security and Governmental Affairs, Subcommittee on Federal Financial Management, Government Information, Federal Services, and International Security, *Getting to Better Government: Focusing on Performance*, 111th Cong., 1st sess., September 24, 2009, at http://hsgac.senate.gov/public/index.cfm?FuseAction=Hearings.Hearing&Hearing_ID=4a27d8ec-2e50-4e6e-aaba-72b62ca9a0ee.

⁸⁵ In **Box 5**, page citations refer to Shelley H. Metzenbaum, *Performance Management Recommendations for the New Administration* (Washington: IBM Center for the Business of Government, [January] 2009), at (continued...)

Box 5: Highlights of Recommendations from Metzenbaum Report

Ms. Metzenbaum's report recommended retaining a "revised" PART to be conducted by OMB, because it "introduced a useful program management discipline" (p. 52). A revised PART would emphasize using analysis to improve, rather than attaining targets or scoring well on overall program ratings (pp. 43, 52-54). It also would decrease emphasis on holding agencies accountable for using effective performance management practices, such as planning and management (p. 54). Separately, there would be four levels of performance measures or "targets":

- presidential,
- cross-agency,
- agency, and
- program (pp. 38, 41).

White House "policy councils" and OMB would identify cross-agency measures and targets (pp. 43, 53). Agencies would "identify a broader set of agency priority targets in their areas of responsibility" (p. 38). Echoing the Obama campaign's plans for a White House performance team, the report recommended that the CPO use practices somewhat analogous to the PMDU of the United Kingdom. A group of "highly regarded public and private sector managers" would provide feedback to members of the Cabinet about how to meet the President's targets and priorities (p. 42). The report focused primarily on performance measurement and analysis of these measures (pp. 41-42), but did not explicitly include the subject of program evaluation in its scope or how program evaluations could address the recommended analyses.

The report identified some ways in which agencies might interact with Congress and the public, but was largely silent about how the President and OMB should work or consult with Congress. For example, the report recommended that OMB "strongly encourage agencies to invite outside expertise and multiple perspectives to inform the selection of targets, performance measures, and strategies to improve performance" and particularly emphasized engagement with "outside experts" (p. 45). Whether Congress and the public would be invited by agencies to participate in specifying overall goals, however, seemed unclear. The report did not address the topic of presidential or OMB consultations with Congress about goal-setting. Non-executive branch involvement in goal-setting focused on allowing comments on already completed reviews under a revised PART (pp. 45, 53-54). The report emphasized reaching out to "key audiences" to address their information needs (pp. 44-45, 50-51). With respect to Congress, for example, the report recommended that agencies consult with "key committees to understand their performance information needs," including how to format information (pp. 50-51).

At the same time, the report described GPRA as a potential tool to "advance [the President's] agenda" (p. 10). The report's recommendations seemed to be premised on the President and agencies using discretion, where available, to engage in goal-setting to "advance [the President's] priorities" and agenda (pp. 10, 38, 49, 53). The report summarized a view under which Congress and the public would have a reactive role: "[g]oals, reinforced by measurement, ... [clarify] agency priorities to the public, allowing Congress and the public to take action if they disagree with the choices that have been made" (p. 10).

At the September 2009 hearing, Mr. Zients described how the performance framework would be developed. In addition to finalizing agencies' HPPGs, he and Ms. Metzenbaum also would develop an "improved Federal performance management framework that aligns these high priority performance goals, GPRA performance reporting, and many of the program-level performance measures developed for the PART."⁸⁶ He said the Administration's "government-wide performance measurement framework" would focus on outcomes, allow comparisons across programs and agencies, and show trends over time.

Mr. Zients also addressed how the success or failure of the Obama Administration's efforts might be assessed and which stakeholders' needs he was seeking to address. He said that the "ultimate test of our performance management efforts is whether or not the information is used." In so doing, he identified the targeted audiences for the effort's performance measures. This information, he said, ideally should be used "not just by government agencies, but also by Congress, the public, our service delivery partners,

(...continued)

<http://www.businessofgovernment.org/report/performance-management-recommendations-new-administration>.

⁸⁶ Prepared testimony of Jeffrey D. Zients, p. 3, at

http://hsgac.senate.gov/public/index.cfm?FuseAction=Hearings.Hearing&Hearing_ID=4a27d8ec-2e50-4e6e-aaba-72b62ca9a0ee.

and others.” Mr. Zients defined “useful” information in two categories, as that information which supports (1) “long-term, central office strategic decisions about targets and organizational strategies,” and (2) “key decision-makers in the service delivery chain including those in Federal field offices, State and local partners, non-profit grantees, and contractors.”⁸⁷ IT and “historic investments” in comparative effectiveness research and evaluations would make greater use possible, he said.

At a subsequent hearing on October 29, 2009, Mr. Zients identified five “key principles” that would guide the Administration’s efforts to develop a “performance management system.”⁸⁸ The principles included

- ownership of the process by senior agency leaders, who would set agency goals and be “held responsible for performance against those goals”;
- cascading goals, where a “clear line must link agency strategic goals and measurements to unit-level, program-level and ultimately individual targets”;
- development of outcome-oriented, cross-agency goals and measurements, with “clear delineation of lines of responsibility” and an emphasis on “outcomes [that require] contributions from multiple actors” inside and outside government;
- “relentless” reviews of performance against plans at all levels of government, including “program, unit and agency level reviews” and reviews of overall agency performance by OMB and other White House entities; and
- transparency, with “active engagement” of the public, Congress, and federal workers.

Where current practices did not meet the five principles, Mr. Zients said the Administration would seek to eliminate or streamline the practices. He also identified new plans. He said the Administration planned to implement “management dashboards,” in addition to the previously announced IT Dashboard.⁸⁹ The dashboards would focus on functions like procurement, financial management, and human resources.⁹⁰

OMB Emphasis on Impact Evaluations

On October 7, 2009, an OMB memorandum announced the Administration would put an emphasis on a specific type of program evaluation called *impact evaluation*.⁹¹ Previous OMB statements suggested the

⁸⁷ When Mr. Zients spoke about targeted audiences at his June 2009 confirmation hearing, he emphasized use of performance information by “senior-most” stakeholders and decision makers. It was not clear if Mr. Zients’ newer statements about audiences were intended to include middle and front-line managers and employees. The distinction may be significant, because service delivery personnel often have different needs compared to policy makers. See Charlie Bennett and Herb Hill, “Performance Management Systems: The Importance of Defining Their Purpose,” [2002], formerly at <http://www.dpb.state.va.us/>. The paper is available on request from the author of this CRS report.

⁸⁸ U.S. Congress, Senate Committee on the Budget, *Task Force Hearings*, hearings, 111th Cong., 1st sess. [October 29, 2009], S.Hrg. 111-381, pp. 1-63.

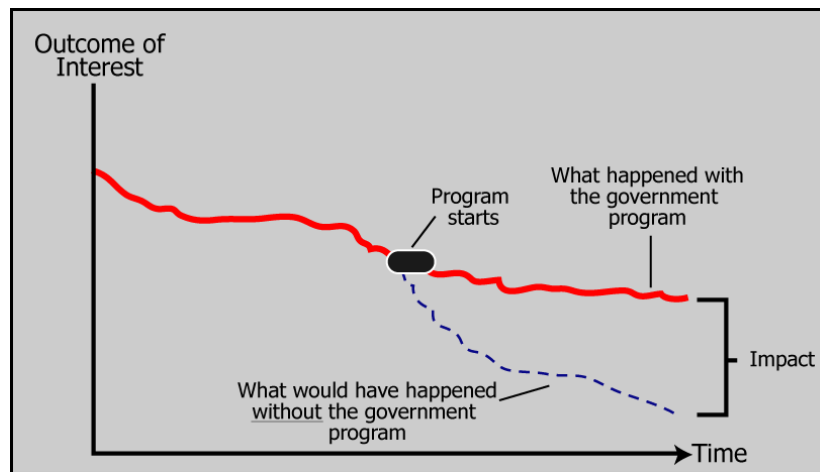
⁸⁹ For discussion of the IT Dashboard, see footnote 68 and accompanying text.

⁹⁰ A month later, the President issued E.O. 13520, which called for a dashboard of indicators on federal improper payments. See Section 2(b) of E.O. 13520, “Reducing Improper Payments,” 74 *Federal Register* 62201, November 20, 2009. OMB later established a related website, <http://www.PaymentAccuracy.gov>. See OMB Director Peter R. Orszag, “PaymentAccuracy.gov,” June 24, 2010, at <http://www.whitehouse.gov/omb/blog/10/06/24/PaymentAccuracygov/>. With regard to personnel management, on May 11, 2010, President Obama issued a memorandum that required OPM and OMB to develop a website to track “key human resource data,” including “progress on hiring reform implementation.” See The White House, Office of the Press Secretary, “Presidential Memorandum—Improving the Federal Recruitment and Hiring Process,” press release, May 11, 2010, at <http://www.whitehouse.gov/the-press-office/presidential-memorandum-improving-federal-recruitment-and-hiring-process>.

⁹¹ OMB, “Increased Emphasis on Program Evaluations,” memorandum, M-10-01, October 7, 2009, at (continued...)

memorandum specifically was referring to a particular meaning of that term: evaluations that quantitatively estimate the extent to which a policy causes a change in an outcome of interest, compared to what would have happened without the policy.⁹² To illustrate, if the U.S. were in a recession, observers might expect the number of jobs in the U.S. to decrease over time (see **Figure 3**). Further suppose that a policy was implemented during the recession with the objective of increasing employment, compared to what would have happened without any policy intervention. If an impact evaluation were conducted, and if the evaluation estimated that the policy caused an increase in the number of civilian jobs (e.g., 136 million jobs) compared to what would have happened otherwise (135 million), the estimated “impact” would be the difference (one million).⁹³

Figure 3. Illustration of Impact Estimate



Source: CRS.

Notes: In the figure, the dotted line is an estimate. The red line shows that performance continues to decrease after the start of a policy, but by less than it would have decreased without the government program.

The OMB memorandum cited several issues as problems. These included “many important programs [that] have never been evaluated”; insufficient influence of some evaluations on budget priorities and

(...continued)

http://www.whitehouse.gov/omb/assets/memoranda_2010/m10-01.pdf. See also OMB Director Peter R. Orszag, “Valuing Evaluation,” October 7, 2009, at <http://www.whitehouse.gov/omb/blog/09/10/07/ValuingEvaluation/>.

⁹² For background, see CRS Report RL33301, *Congress and Program Evaluation: An Overview of Randomized Controlled Trials (RCTs) and Related Issues*, by Clinton T. Brass, Erin D. Williams, and Blas Nuñez-Neto.

⁹³ That is, under this meaning of the term *impact evaluation*, an *impact* is defined as the difference between two numbers: (a) a measurement of what happened to an outcome after a policy was implemented, and (b) an estimate of what would have happened to the outcome, if the policy had not been implemented. The concept of (b)—what would have happened without the program—is not observed and must be estimated. Using the terminology of evaluation, (b) is called the *counterfactual* or *counterfactual estimate*. Under a second meaning of the term *impact evaluation*, an evaluation may be used to infer a cause-and-effect relationship, but without estimating quantitatively what would have happened without the program. In other words, estimation of a counterfactual is not the only way to demonstrate that a causal impact likely occurred. (For discussion, see Lawrence B. Mohr, *Impact Analysis for Program Evaluation*, 2nd ed. (Thousand Oaks, CA: Sage Publications, 1995), pp. 65, 248-273; and Stephen Gorard and Thomas Cook, “Where Does Good Evidence Come From?”, *International Journal of Research & Method in Education*, vol. 30 (November 2007), p. 311.) Some of the methods associated with the second meaning of impact evaluation may be used to understand how and why impacts occur. Hence, although the term *impact evaluation* may have two distinct meanings, both focus on whether a government policy causes a change and may be used to make causal inferences. A study’s focus on causation may be important when factors external to a program are known to influence outcomes. Otherwise, a study may leave unclear whether a change in outcomes is due to the program, external factors, or both.

management practices; a lack of evaluation offices and staffing in many agencies; studies of “insufficient rigor or policy relevance”; and a lack of evaluations that identify which of “multiple approaches to the same problem ... are most effective.” In response to these perceptions, OMB announced a three-part initiative to focus on impact evaluations of social and economic programs.

First, OMB would work with agencies to put information online about all federal impact evaluations that are planned or underway. Second, OMB and several White House policy councils would “re-constitute” an inter-agency evaluation working group under the PIC, which was established by E.O. 13450 (see **Box 2**, p. 11). The working group would build agencies’ evaluation capacities and develop government-wide guidance on program evaluation practices, with a goal of helping agencies “determine the most rigorous study designs appropriate for different programs given their size, stage of development, and other factors.”⁹⁴ Third, OMB invited agencies to propose funding for “high-priority” evaluation activities in their FY2011 budget submissions to OMB. OMB would include funding in the President’s budget proposal for up to 20 “rigorous” evaluations or efforts to improve evaluation capacity. Agencies wishing to be considered for such funding would need to assess the “adequacy of evidence” supporting their priorities. New initiatives and program expansions would need to provide “credible evidence of the program’s effectiveness.” In addition, OMB asked submitting agencies to identify statutes that “[limit] their ability to carry out rigorous evaluations, allocate evaluation resources to key priorities, or otherwise impede efforts to build evidence of what works.” OMB asked agencies to suggest strategies for working with Congress to address these topics.

The OMB memorandum generally characterized impact evaluations as supplying “evidence” regarding “what works” and how best to allocate taxpayer dollars. As discussed later in this memorandum, several kinds of evidence typically are relevant when making policy or allocating resources. Impact evaluations may provide information about the extent of a change caused by a policy intervention, but they do not necessarily indicate whether a policy would have a similar impact at other times or settings. Impact evaluations also oftentimes do not assess how and why impacts occur, how and why subpopulations are affected, how a program might be modified to improve, or a program’s cost-effectiveness. Other kinds of evaluations may provide complementary or more useful information to make such judgments.

Statements on Implementation Challenges and OMB’s Role

In two October 2009 speeches, Mr. Zients reportedly made commitments for how OMB would work with Congress and agencies in the future. After a “listening tour” during his first 100 days on the job, he identified four barriers to getting things done.⁹⁵ These included (1) a focus by political appointees on

⁹⁴ OMB’s focus on “appropriate” study designs suggested a change from previous OMB guidance on impact evaluations. In 2004, OMB almost exclusively emphasized RCTs for assessing impacts, and, less precisely, for assessing “effectiveness.” See OMB, “What Constitutes Strong Evidence of a Program’s Effectiveness?,” [2004], at http://www.whitehouse.gov/omb/assets/omb/performance/2004_program_eval.pdf. OMB’s 2004 guidance apparently was written with assistance from the Coalition for Evidence-Based Policy. Some federal agencies’ evaluation personnel and a task force of the American Evaluation Association (AEA) critiqued OMB’s guidance for over-emphasizing RCTs and not addressing other methods. The federal evaluators gave a briefing that OMB posted on its website: “Evaluation Dialogue Between OMB Staff and Federal Evaluators: Digging a Bit Deeper into Evaluation Science,” July 2006, at http://www.whitehouse.gov/omb/assets/omb/performance/supp_materials/fed_eval_2006_slides.pdf. The AEA task force sent a letter to OMB: Letter from William M.K. Trochim, President, American Evaluation Association, to Robert Shea, Associate Director for Administration and Government Performance, OMB, March 7, 2008, at <http://www.eval.org/aea08.omb.guidance.responseF.pdf>.

⁹⁵ Jason Miller, “OMB Wants to Change the Tone of Management,” *Federalnewsradio.com*, October 21, 2009, at <http://www.federalnewsradio.com/?nid=35&sid=1791266>; and Elizabeth Newell, “Performance Chief Promises to Help Managers Overcome Challenges,” *GovExec.com*, October 5, 2009, at <http://www.govexec.com/dailyfed/1009/100509e2.htm>.

policy rather than management and implementation; (2) insufficient exploitation of IT, with gaps in quality and productivity compared to the private sector; (3) federal contracting that does not take advantage of the government's purchasing leverage; and (4) slow processes for hiring federal employees. Mr. Zients reportedly emphasized a need to prioritize, saying "[t]o focus on everything is to focus on nothing." Commenting on OMB's approach, he reportedly committed to regular communication with Congress and a collegial instead of command-and-control role with agencies.⁹⁶

FY2011 President's Budget and Announcement of HPPGs

The President submitted his FY2011 budget proposals to Congress on February 1, 2010. The budget documents charted further evolution of the Administration's performance-related agenda, but also announced that substantial parts of the agenda were still to be developed.

High-Priority Performance Goals and "Performance Management Strategies"

Following up on OMB's June 2009 memorandum and guidance (see **Box 4**, p. 17), the Administration released HPPGs for 24 federal agencies and one executive order.⁹⁷ The Administration did not provide a precise count of the goals, but identified them using 128 bullets. Each bullet, in turn, identified one or more metrics, milestones, or objectives. Release of the finalized HPPGs provided an opportunity to see how agencies attempted to implement the Administration's previous guidance. In OMB's June 2009 memorandum, OMB said the goals should be of "high direct value to the public" and not focus on "internal agency management or other administrative priorities." Although some of the finalized HPPGs focused on end outcomes, most of the goals and supporting metrics focused on inputs, activities, outputs, and intermediate outcomes, as means through which to accomplish end outcomes that might be of comparatively more direct value to the public (see **Box 3** and **Figure 2**, pp. 12-13).⁹⁸

The FY2011 budget also announced a new framework. In the coming fiscal year and thereafter, the Administration would pursue three "performance management strategies." The strategies focused predominately on goal-setting and performance measurement. Highlights of the strategies are summarized in **Table 1**.

⁹⁶ Mr. Zients reportedly went into some detail on how OMB would work with agencies and Congress, saying "[w]e're going to move from OMB sets priorities to the agency sets priorities; from oversight to partnership; from shipping reams of guidance to two-way dialogue about how we achieve the desired outcome; from transparency not just for accountability, but for idea flow to find and share the best practices; from ad hoc engagement for stakeholders such as Congress, to regular communication. ... I'm positive we'll make mistakes, we'll slip into some old bad habits, but I commit to you that we will serve you differently than we have in the past." Ibid.

⁹⁷ OMB, *Analytical Perspectives, Budget of the U.S. Government, FY2011* (Washington: GPO, 2010), p. 75.

⁹⁸ The agencies' emphasis on process-related and other intermediate metrics may not be surprising, given the short-term nature of the goals, the Administration's emphasis that performance information should be useful during implementation, and the nature of how managers typically set short- and medium-term agendas. A broad literature suggests managers tend to focus on intermediate milestones and accomplishments that are necessary in order to achieve more long-term, strategic objectives. See John P. Kotter, "What Effective Managers Really Do," *Harvard Business Review*, vol. 77 (March-April 1999), pp. 145-159; and Henry Mintzberg, *The Nature of Managerial Work* (New York: HarperCollins, 1973), p. 153.

Table I. “Performance Management Strategies” in President’s FY2011 Budget

Strategy	Administration Plans
“Use performance information to lead, learn, and improve outcomes”	In addition to pursuing HPPGs at the agency level, the Administration said it would ask agency leaders to carry out a similar goal-setting exercise at the bureau level in the coming year. Both agency leaders and OMB would put in place quarterly feedback and review sessions that are modeled after so-called “Stat” efforts in state and local governments. ⁹⁹
“Communicate performance coherently and concisely for better results and transparency”	The Administration said it would “eliminate performance measurements and documents that are not useful” and convey information about agency-, cross-agency-, and program-level measures. In combined performance plans and reports, the Administration would explain “why goals were chosen, the size and characteristics of problems Government is tackling, factors affecting outcomes that Government hopes to influence, lessons learned from experience, and future actions planned.” In addition, agencies would “take greater ownership” in communicating GPRA-required performance plans and reports to “key audiences.” In an effort to make performance data useful to “all audiences—congressional, public, and agency leaders,” a new federal performance portal would provide “a clear, concise picture of Federal goals and measures by theme, by agency, by program, and by program type.” The portal also would link to “mission-support management dashboards.” The dashboards would include the IT Dashboard and “similar dashboards planned for other ... functions including procurement, improper payments, and hiring.” Information about all federal impact evaluations would be available through the portal as well.
“Strengthen problem-solving networks”	The Administration planned to use “existing and new practitioner networks,” both inside and outside government. To address shared problems, the Administration would create cross-agency teams. The Performance Improvement Council (PIC), established by E.O. 13450 (see Box 2 , p. 11) would “function as the hub” of the overall network. OMB would work with the PIC to “advance a new set of Federal performance management principles, refine a Government-wide performance management implementation plan, and identify and tackle specific problems as they arise.” The PIC, in turn, would establish communities of practice organized by program type, problem, and methods.

Source: OMB, *Analytical Perspectives, Budget of the U.S. Government, FY2011*, pp. 73-75.

Impact Evaluations and Characterizations of “Evidence-Based Policy”

The FY2011 President’s budget proposal also displayed some evolution of OMB’s three-part initiative to emphasize impact evaluations, which had been announced in October 2009.¹⁰⁰ After reiterating that impact evaluations that are “planned or already underway” would be put online (see **Table 1**), the Administration offered more detail about the operation of an “inter-agency working group” (hereafter IWG). The IWG would work with four entities in the Executive Office of the President: OMB, the Council of Economic Advisers (CEA), the Domestic Policy Council, and the Economic Policy Council. Among other things, the IWG and these EOP entities would help build evaluation capacity and networks, assist in selecting evaluation designs, devise strategies for “using data and evaluation to drive continuous improvement,” and develop executive-branch-wide guidance on program evaluation practices with “sufficient flexibility for agencies to adopt practices suited to their specific needs.” Nevertheless, the

⁹⁹ The “Stat” approach, at least under that name, was pioneered in New York City in the mid-1990s, when the police department there began a management system called “computerized statistics,” or “Compstat” for short. The Obama Administration has cited CompStat and similar approaches when describing its goal-setting, performance measurement, and management efforts. See OMB, *Analytical Perspectives, Budget of the U.S. Government, FY2011*, p. 73. In July 2010, Senate Budget Committee’s Task Force on Government Performance held a field hearing on the subject. See U.S. Congress, Senate Committee on the Budget, *Performance “Stat”: Measuring Priorities, Progress and Results*, hearing 111th Cong., 2nd sess., July 12, 2010, at <http://budget.senate.gov/democratic/hearingstate.html>.

¹⁰⁰ Unfootnoted citations in this section are to OMB, *Analytical Perspectives, Budget of the U.S. Government, FY2011* (Washington: GPO, 2010), pp. 91-92. In a succeeding chapter of the same volume, the Administration devoted substantial discussion to benefit-cost analysis (BCA). Much of the discussion focused on the use of BCA prospectively with regulations. However, the chapter also newly focused on the use of BCA for regular programs. The chapter may signal the possibility of an increased Administration emphasis on the use of BCA as form of program evaluation.

budget left unclear whether the IWG would act autonomously or, alternatively, become a means through which the EOP would direct agency practices. The relationship between the IWG, which primarily is oriented toward program evaluation, and the PIC (see **Box 2**, p. 11), which primarily is oriented toward performance measurement, also was unclear.

Following up on the third part of the initiative, the Administration listed 35 FY2011 budget proposals for 18 agencies, amounting to “approximately \$100 million,” to conduct evaluations or build evaluation capacity. Of the 35 proposals, 28 focused on evaluating specific programs or policies. These proposals would fund new evaluations with “strong study designs that address important, actionable questions.” The remaining seven proposals would “strengthen agency capacity to support such strong evaluations.” OMB and CEA would “continue to work with these agencies on implementing strong research designs that answer important questions.” What would constitute a “strong” study design was not clear. The Administration also did not indicate what kind of relationship OMB and CEA would have with agencies.

The Administration also elaborated its views on the role that evidence should play when making policy, as well as how to define what constitutes evidence. The way the President or an agency characterizes evidence for policy making may have implications for Congress in the legislative process; for example, when justifying budget or policy proposals. Similar implications for Congress may be evident for purposes of oversight and transparency, when agencies or the President use discretion to allocate resources and implement policies based on this evidence.

Specifically, the Administration’s budget proposal said the President “has made it very clear that policy decisions should be driven by evidence.”¹⁰¹ The Administration identified two types: (1) evidence about “what works and what does not”; and (2) evidence that “identifies the greatest needs and challenges.” The Administration said little regarding how to define the second type of evidence. By contrast, with regard to the first type, investments would be made in programs that are supported by “strong” or “rigorous” evidence. The budget did not define what would constitute strong or rigorous evidence. However, for new initiatives, the Administration indicated it would use a “three-tiered approach” to “evidence-based policy making.” In the first or “top” tier, more resources would be allocated for programs backed by “strong” evidence. In the second tier, programs with “supportive” evidence would receive additional resources, “on the condition that programs will be rigorously evaluated.” Third, agencies could pursue approaches “supported by preliminary research findings or reasonable hypotheses.” The Administration said the three-tier approach provides “objective” criteria to inform resource allocation. As an example, the Administration cited the Department of Education’s (ED’s) Invest in Innovation Fund (i3).¹⁰²

The i3 program, therefore, may indicate how the Administration plans to define and use evidence in proposals to Congress and in administrative policy making. Shortly after publication of the President’s budget proposals, ED defined “evidence” in a final rule almost solely as the estimate of a program’s impact, as produced by an impact evaluation in the sense illustrated by **Figure 3** (p. 21). According to ED, impact evaluations would facilitate making “causal conclusions” and generalizing findings for “scaling up to the State, regional, or national level.”¹⁰³ By contrast, ED said it did not see a need to include in the definition of evidence the information provided by other types of evaluations that do not quantitatively estimate an impact, including qualitative, process, and “mixed method” approaches.¹⁰⁴ These evaluations

¹⁰¹ Ibid., p. 92.

¹⁰² Ibid.

¹⁰³ U.S. Department of Education, “Investing in Innovation Fund; Final Rule and Notice,” *75 Federal Register* 12004-12071, at 12063, March 12, 2010.

¹⁰⁴ Ibid., pp. 12053 and 12055.

often are used to understand mechanisms and contexts of causation and, therefore, circumstances under which a program may have an impact in different settings, populations, and times.¹⁰⁵ Given this focus on impact evaluations and the omission of other kinds of evaluation information from the definition of “evidence,” past experience and the program evaluation literature suggest that if the Administration applied the i3 definition of “evidence” more widely, such a move may be controversial.¹⁰⁶ For example, similar controversy arose in 2010 in the context of a grant program established by the Patient Protection and Affordable Care Act of 2010 (PPACA; P.L. 111-148).¹⁰⁷

Line-by-Line Review and Proposals for Termination or Reduction

The Administration issued its second annual *Terminations, Reductions, and Savings* document (*FY2011 TRS Document*) based on a continued “line-by-line review” of the budget.¹⁰⁸ By its count, the Administration proposed 126 cuts or program restructurings amounting to \$23 billion for FY2011.¹⁰⁹ Fifteen of the proposals came from federal employee suggestions under the President’s SAVE Award initiative.¹¹⁰ Some of the proposals’ justifications referred explicitly to program evaluations or performance measures, but most made analytical arguments in support of Administration proposals. Some justifications, for example, argued for reducing “duplication” and favored administrative decision making processes for allocating resources instead of congressionally directed spending.

February 2010 Framework: Six “Performance Strategies”

On February 18, 2010, Mr. Zients followed up on the FY2011 budget with a speech that outlined a new framework for the Administration’s performance agenda. The framework comprised six “performance strategies” that “offer the greatest potential to improve [government] performance.”¹¹¹ In order of presentation, the strategies included (1) “eliminate waste”; (2) “drive top priorities”; (3) “leverage

¹⁰⁵ CRS Report RL33301, *Congress and Program Evaluation: An Overview of Randomized Controlled Trials (RCTs) and Related Issues*, by Clinton T. Brass, Erin D. Williams, and Blas Nuñez-Neto.

¹⁰⁶ See *ibid.* and William Easterly, “The Civil War in Development Economics,” (New York University Development Research Institute), December 3, 2009, at <http://aidwatchers.com/2009/12/the-civil-war-in-development-economics/>.

¹⁰⁷ The controversy arose in the context of a PPACA grant program (P.L. 111-148, Section 2951; 124 Stat. 335) and a related proposal by the Department of Health and Human Services (HHS) to establish “criteria for evidence of effectiveness of home visiting program models.” HHS’s proposed criteria generally would elevate RCTs above other forms of impact evaluation in terms of study “quality” and, in addition, would not provide a role for program evaluation methods other than impact evaluations to inform assessments of “evidence of effectiveness” for a program model. In the relevant PPACA provision, Congress required HHS to articulate its proposal through public notice and solicitation of comments. HHS’s notice prompted expressions of concern from groups including the American Evaluation Association (AEA) and the Center for Law and Social Policy (CLASP). The AEA is a professional association of evaluators, and CLASP is an advocacy group that focuses on policies that relate to low-income people. For the HHS proposal, see U.S. Department of Health and Human Services, “Maternal, Infant, and Early Childhood Home Visiting Program,” *75 Federal Register* 43172, July 23, 2010. For comments from the AEA, see letter from Leslie Cooksy, Debra J. Rog, and Jennifer Greene, AEA, to Mary K. Wakefield, Administrator, Health Resources and Services Administration (HRSA), HHS, August 17, 2010, at <http://www.eval.org/EPTF/aea10.eptf.visitation.comments.box.pdf>. Also see letter from Rutledge Q. Hutson and Tiffany Conway Perrin, CLASP, to Mary K. Wakefield, Administrator of HRSA, HHS, August 16, 2010, at http://www.clasp.org/federal_policy/pages?id=0015.

¹⁰⁸ OMB, *Terminations, Reductions, and Savings—Budget of the U.S. Government, Fiscal Year 2011* (hereafter *FY2011 TRS Document*) (Washington: 2010), at <http://www.whitehouse.gov/omb/budget/TRS/>.

¹⁰⁹ The Administration also said Congress enacted \$6.8 billion of the Administration’s *FY2010 TRS Document* proposals.

¹¹⁰ *Ibid.*, p. 1. According to the Administration, the 15 employee-suggested items did not require action by Congress to be implemented. For more on the SAVE initiative, see <http://www.whitehouse.gov/save-award>.

¹¹¹ Speech given at Center for American Progress, February 18, 2010. A transcript of the speech and a subsequent panel discussion is available at <http://www.americanprogress.org/events/2010/02/dww.html>.

purchasing scale”; (4) “close the IT performance gap”; (5) “open government to get results”; and (6) “attract and motivate top talent.” Mr. Zients placed many of the Administration’s previously announced initiatives within this framework. For example, he included within the first strategy the Administration’s *FY2011 TRS Document*, emphasis on impact evaluations, and focus on reducing improper payments. HPPGs were the primary focus of the second strategy. Publicly visible, online dashboards would support many strategies, especially including initiatives relating to improper payments, IT, and human resources. He said periodic “Stat” reviews would occur for HPPGs and IT projects.¹¹² Mr. Zients also announced that the Administration would establish a “government-wide performance portal” in the summer of 2010.

Looking Ahead to the President’s FY2012 Budget Proposals and 112th Congress

In July 2010, a White House document described an Accountable Government Initiative (AGI) that included many of the Administration’s performance-related initiatives from the framework that Mr. Zients outlined in his February 2010 speech.¹¹³ Several weeks later, the President and Mr. Zients issued two memoranda to the Senior Executive Service (SES) that more formally transmitted the AGI.¹¹⁴ It remains to be seen whether this framework will be presented in the President’s FY2012 budget, which will be submitted in early 2011 to the 112th Congress.

During the summer of 2010, OMB also issued documents in preparation for the FY2012 proposal and in anticipation of rolling out the Administration’s performance-related initiatives. With regard to the FY2012 budget, OMB and the White House issued two memoranda to the heads of agencies on June 8, 2010, that provided guidance for how agencies should submit their budget requests to OMB. In one memo, OMB Director Orszag and White House Chief of Staff Rahm Emanuel required agencies to identify discretionary programs that have the “lowest impact... on your agency’s mission and relevant Presidential initiatives.”¹¹⁵ In particular, they urged agencies to identify at least 5% of their enacted FY2010 discretionary appropriations, excluding emergency and supplemental funding, as having the lowest impact. In the second memo, Director Orszag directed agencies to include at least five termination and reduction proposals with supporting “analysis and evidence showing the effects of the reductions and why they are warranted.”¹¹⁶ OMB also issued a memorandum to emphasize impact evaluations with the President’s FY2012 budget proposals,¹¹⁷ continuing the emphasis that OMB began in October 2009 for the President’s FY2011 budget proposals.

¹¹² For more on TechStat, see CIO Vivek Kundra, “TechStat: Improving Government Performance,” February 24, 2010, at <http://www.whitehouse.gov/blog/2010/02/24/techstat-improving-government-performance>.

¹¹³ See The White House, Office of the Press Secretary, “Backgrounder: The Accountable Government Initiative: The Obama Administration’s Effort to Cut Waste and Modernize Government,” July 22, 2010, at http://www.whitehouse.gov/sites/default/files/Background_on_Accountable_Government_Initiative.pdf. The PDF link may be found at <http://www.whitehouse.gov/the-press-office/president-obama-sign-improper-payments-elimination-and-recovery-act>.

¹¹⁴ The White House, Office of the Press Secretary, “Presidential Memorandum—Accountable Government Initiative,” September 14, 2010; and OMB, “The Accountable Government Initiative—An Update on Our Performance Management Agenda,” September 14, 2010, at <http://www.whitehouse.gov/the-press-office/2010/09/14/presidential-memorandum-accountable-government-initiative>, and <http://www.whitehouse.gov/blog/2010/09/14/sending-out-ses>.

¹¹⁵ They urged agencies to focus on programs that have “an unclear or duplicative purpose, uncertain Federal role, completed mission, or lack of demonstrated effectiveness.” OMB, “Identifying Low-Priority Agency Programs,” memorandum, M-10-20, June 8, 2010, p. 2, at http://www.whitehouse.gov/sites/default/files/omb/assets/memoranda_2010/m10-20.pdf.

¹¹⁶ OMB, “Fiscal Year 2012 Budget Guidance,” memorandum, M-10-19, June 8, 2010, p. 3, at http://www.whitehouse.gov/sites/default/files/omb/assets/memoranda_2010/m10-19.pdf.

¹¹⁷ OMB, “Evaluating Programs for Efficacy and Cost-Efficiency,” memorandum, M-10-32, July 29, 2010, at (continued...)

In June and July of 2010, OMB also issued two documents that related to performance measurement and goal-setting. In a June memorandum, OMB Associate Director for Performance and Personnel Management Shelley Metzenbaum provided “transition” guidance to agencies on these matters. She said agencies “should consider this year a transition year during which OMB and the PIC will move to a more dynamic performance planning, management, improvement, and reporting framework.”¹¹⁸ The memorandum provided general guidance to agencies on three topics: achieving HPPGs that had been announced in the President’s FY2011 budget, establishing quarterly reviews of progress toward meeting HPPGs, and developing GRPA-related plans and reports. Just weeks later, OMB issued its annual guidance to agencies in *Circular No. A-11*, which included new instructions on how to comply with GPRA. The circular drew from the June memorandum and added that quarterly reviews would be held between agencies and OMB.¹¹⁹

Potential Issues for Congress

The Obama Administration’s performance agenda has evolved, with several announcements of frameworks to guide priorities and actions. Some of these frameworks closely relate to program evaluation and performance measurement, and also to the relationship of these tools to management and budgeting. These announcements have included

- identification of five “key principles” for a “performance management system” (October 2009, CPO testimony);
- three “performance management strategies” (February 2010, FY2011 budget proposal); and
- a “performance planning, management, improvement, and reporting framework” (June 2010, OMB memo).

Other announcements of frameworks appear to draw on the broader responsibilities of OMB’s DDM position to lead OMB’s statutory offices and establish general management policies for mission-support functions. The frameworks incorporated multiple efforts to use evaluation and measurement and included

- six themes for a “new management and performance agenda,” with further options to be developed (May 2009, FY2010 budget proposal);
- five areas of presidential emphasis (June 2009, OMB DDM confirmation hearing);
- an upcoming process to craft a “management and performance framework” (June 2009, OMB memo);
- six “performance strategies” (February 2010, CPO speech); and
- an Accountable Government Initiative (July-September 2010, White House documents).

(...continued)

<http://www.whitehouse.gov/sites/default/files/omb/memoranda/2010/m10-32.pdf>.

¹¹⁸ OMB, “Performance Improvement Guidance: Management Responsibilities and Government Performance and Results Act Documents,” memorandum, M-10-24, June 25, 2010, p. 1, at

http://www.whitehouse.gov/sites/default/files/omb/assets/memoranda_2010/m10-24.pdf.

¹¹⁹ OMB, *Circular No. A-11, Preparation, Submission, and Execution of the Budget*, July 2010, Part 6, Section 200, p. 2, at http://www.whitehouse.gov/sites/default/files/omb/assets/a11_current_year/a_11_2010.pdf.

As the Administration further develops and implements its agenda during the 112th Congress, complex policy, fiscal, and performance challenges will continue to face Congress. Agencies and OMB likely will present policy analysis tools such as program evaluation and performance measurement to Congress in efforts to inform and influence its deliberation, lawmaking, and oversight. At the same time, agencies will engage in performance-related activities while they formulate and implement federal government policies. OMB also may attempt to exert influence on agency activities, acting on the President's behalf. Any of these activities might be subjects of congressional oversight. Nevertheless, significant aspects of the Administration's performance-related agenda have yet to be articulated. In some cases, it remains to be seen if previous announcements of plans will come to fruition. Whether the Administration's agenda draws on previously announced plans or moves in new directions, the agenda suggests potential issues for Congress. The issues might be placed in two general categories.

First, if Congress acts to approach the agenda without attempting to modify it, and to consider the evaluation and performance information that agencies and OMB present to Congress, several issues may arise. During the legislative process, for example, appropriations and authorizing committees routinely receive proposals from agencies, OMB, and interest groups that cite evaluations and measurements to justify funding and policy changes.¹²⁰ Because evaluations and measurements may be flawed or may be presented in ways that some observers might view as biased, Congress might use several resources to scrutinize the credibility and relevance of the information. Congress also might conduct oversight over how well agencies are using evaluations and measurements when implementing policy, allocating funds through contracts or grants, creating incentives for good management, and conducting policy research.¹²¹

A second, broader category of issues may arise if, instead of taking the Administration's agenda on its own terms, Congress decides to scrutinize policies and practices that relate to government performance. In this second category, Congress might examine the extent to which the Administration's agenda seems appropriate for the federal government's current and future challenges. Congress also might consider whether federal performance-related statutes, like GPRA, and administratively pursued policies, such as E.O. 13450, merit reexamination. Furthermore, when agencies or OMB justify legislative proposals with evaluation and measurement information, issues that routinely face appropriations and authorizing committees, including issues of credibility and relevance, might be addressed more systematically through legislation or oversight.

Analysis of several issues in each category follows below.

Potential Issues When Evaluations and Performance Measures Are Presented to Congress

Scrutinizing Program Evaluations and Performance Measurements

Congress has numerous resources and opportunities to inform itself, make policy, and influence implementation. Consequently, participants in the policy process may, in an attempt to influence Congress, make representations about the supposed "performance" of a program or policy. Some of this

¹²⁰ For example, see CRS Report RL33071, *Even Start: Funding Controversy*, by Gail McCallion; and CRS Report RL34306, *Vulnerable Youth: Federal Mentoring Programs and Issues*, by Adrienne L. Fernandes-Alcantara.

¹²¹ For related discussion, see CRS Report RL33301, *Congress and Program Evaluation: An Overview of Randomized Controlled Trials (RCTs) and Related Issues*, by Clinton T. Brass, Erin D. Williams, and Blas Nuñez-Neto; and GAO, *A Variety of Rigorous Methods Can Help Identify Effective Interventions*, GAO-10-30, November 2009.

information may come in the form of program evaluations or performance measurement. As discussed earlier in **Box 3** and **Figure 1** (pp. 12-13), program evaluations and performance measurement focus on different kinds of information and may have different uses. Furthermore, they have distinct capabilities and limitations. The Administration's performance agenda has emphasized some aspects of performance measurement with a focus on HPPGs (**Box 4**, p. 17) and various dashboards, and one type of program evaluation with impact evaluations (**Figure 3**, p. 21). Other approaches to evaluation and measurement, such as logic modeling as described in **Box 3** and **Figure 2** (p.13), have not received explicit emphasis.

Issues that Congress may consider exploring, both generally and in specific policy areas, include whether the President, agencies, and interest groups are using evaluation and measurement methods consistently with their capabilities and limitations; whether findings are appropriately qualified and presented in an unbiased way; and whether the Administration's approaches to evaluation and measurement address the needs of key audiences. For example, for FY2010, the Administration proposed that Congress terminate ED's school-based "Student Mentoring Program" partly on the basis of an impact evaluation that found, "for *the full sample of students*, the program did not lead to statistically significant impacts" on outcomes relating to school engagement, academic achievement, and other topics.¹²² The information presented by the Administration was accurate, but did not provide a complete picture. For example, OMB and ED did not highlight that the study found statistically significant impacts in academic outcomes for a major *subpopulation*, girls;¹²³ did not provide cautions that the study's impact estimates may have been biased toward zero because of potential contamination of the experiment;¹²⁴ and did not address what the underlying study's lead author reportedly characterized as flaws in implementing the program and the failure to address those flaws, that, according to the author, "resulted in weak evaluation outcomes."¹²⁵

Members and committees of Congress might encounter issues like these in the context of annual appropriations, program and agency reauthorizations, and oversight. During the appropriations process, for example, appropriations subcommittees typically review the President's budget requests and justifications. In situations when proposals are justified by evaluation or measurement information, committees might use their own staff resources and other resources such as GAO and CRS to evaluate the credibility and relevance of the evaluations and measurements. Similar issues may arise in the context of substantive lawmaking, such as when a committee is considering changes to a program's authorizing statute. In an oversight context, Members and committees might examine whether agencies are using appropriate social science and other methods to support administrative policy making and allocation of funds. Finally, if Congress became more broadly interested in or concerned about how the Administration or agencies were using discretion to present or use performance information, Congress might address the topic on a government-wide basis as a matter of performance-related policy.

¹²² Italics added. See OMB, *FY2010 TRS Document*, p. 61. For more detailed discussion of the program, see CRS Report RL34306, *Vulnerable Youth: Federal Mentoring Programs and Issues*, by Adrienne L. Fernandes-Alcantara.

¹²³ Lawrence Bernstein, et al., *Impact Evaluation of the U.S. Department of Education's Student Mentoring Program: Final Report*, Institute of Education Sciences, U.S. Department of Education, March 2009, pp. xxii, 90-91, at <http://ies.ed.gov/ncee/pubs/20094047/>.

¹²⁴ Students not receiving mentoring services may have become aware that peers were receiving mentoring under the study, potentially leading non-served students and their families to seek mentoring elsewhere that they might not have sought otherwise. According to ED's What Works Clearinghouse (WWC, the department's "central and trusted source of scientific evidence for what works in education") in response to a query from CRS, if this contamination occurred, the study's impact estimates "could be biased toward zero"; however, WWC "does not have standards for evaluating this type of contamination" (e-mail communication from WWC to CRS, January 20, 2010). For discussion of this threat of contamination to a study's validity, see Thomas D. Cook and Donald T. Campbell, *Quasi-experimentation: Design & Analysis Issues for Field Settings* (Chicago: R. McNally, 1979), pp. 54-55.

¹²⁵ Jen Russell, "Evaluation Spotlight: School-Based Mentoring Needs a Friend," *Youth Today*, June 2009.

Scrutinizing How Agencies or OMB Define “Success” for a Program

When considering the performance of an agency, program, or policy in any of these legislative contexts, Congress also might look more broadly at how an agency or OMB is defining “success.” The definition of success that an agency or OMB uses in a particular circumstance, such as a budget proposal for Congress’s consideration, may dictate whether currently existing performance information portrays a program in a positive or negative light. Consequently, Congress might in specific circumstances consider whether it agrees with these representations about how to define success.

These situations frequently arise in the context of appropriations, substantive lawmaking, and oversight, because the definition of success for an existing program often is politically contested. That is, participants and stakeholders in the policy process may have differing views about the proper means and ends of an existing program, policy, or agency. A related issue may arise when a program, policy, or agency has multiple goals that trade off against each other. For example, the mission statement of the Internal Revenue Service (IRS) arguably encompasses three general goals: tax compliance, customer service, and fairness.¹²⁶ For IRS, there is a trade-off between resources allocated to tax compliance and customer service. Judgments about policy success or failure in the presence of goal trade-offs may not receive widespread acceptance, because the process of making trade-offs involves value judgments that may differ among stakeholders.

Past experience in the federal budget process suggests that a President may make representations about performance from the perspective of one definition of success, while omitting any mention of other perspectives. For example, federally supported vocational education could have been rated either effective or ineffective, depending on whether “success” was defined as increased earnings for education recipients, on one hand (effective), or increased incidence of seeking higher education, on the other (ineffective).¹²⁷ Consequently, when the President or his or her Administration makes representations to Congress about performance for a particular policy or program, Congress may consider whether the definition of success that is being used reflects congressional intent for the underlying program and major perspectives that the broader community of stakeholders may have, or reflects a more limited perspective.

In the context of presidential appropriations proposals, for example, an appropriations committee might examine whether a proposal for a budget increase or decrease is justified by program evaluations or performance measures. If OMB makes such a justification, a committee might examine whether claims of how to define good or bad performance conform with a program’s authorizing statute and legislative history. The committee also might examine whether available performance information addresses stakeholders’ conceptions of how to define success. Nongovernmental entities, reputable evaluators, and congressional support agencies like CRS and GAO might assist the committee with such analysis. Congress might undertake similar activities when considering changes to an authorizing statute on the basis of evaluations or measurements. In an oversight context, a Member or committee might examine how an agency is defining success when it uses discretion in how it identifies HPPGs (see **Box 4**, p. 17) or awards grant funding. If the agency is using a definition of success other than what Congress intended, or a definition that appears to reflect only part of what Congress intended, Congress might consider additional oversight activities or lawmaking.¹²⁸

¹²⁶ See IRS, “The Agency, its Mission and Statutory Authority,” at <http://www.irs.gov/irs/article/0,,id=98141,00.html>.

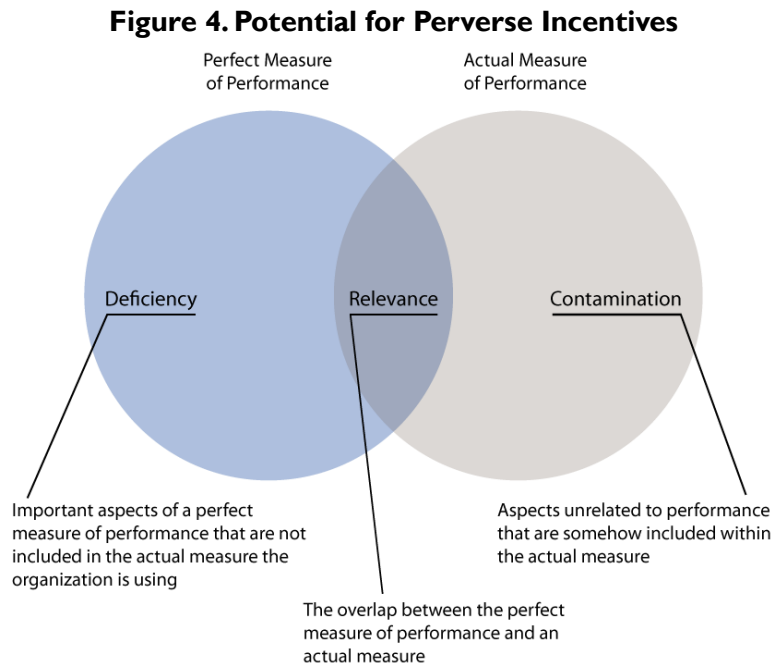
¹²⁷ See discussion of contrasting definitions of success for federally supported vocational education in CRS Report RL33301, *Congress and Program Evaluation: An Overview of Randomized Controlled Trials (RCTs) and Related Issues*, by Clinton T. Brass, Erin D. Williams, and Blas Nuñez-Neto.

¹²⁸ In some cases, Congress’s or a committee’s specific intent may not be clear from the statutory text or legislative history.

Identifying Possibilities of Perverse Incentives

When agencies identify and pursue goals or other definitions of success, measurements can be a strong motivator to action. In some situations, measurements may create incentives that, from the perspective of congressional intent, may not be viewed as desirable. In its regular legislative work, and particularly when agencies and the President are identifying what they consider to be priority goals and potentially relegating other potential goals to a relatively lower priority, Congress might encounter such issues and consider whether remedies are advisable.

The problem of perverse incentives has long been recognized. Industrial psychologist Mason Haire is frequently quoted: “What gets measured gets done. If you are looking for quick ways to change how an organization behaves, change the measurement system.”¹²⁹ At the same time, even if consensus were possible on how to define an agency’s or program’s major purposes, it may be difficult to find a “vital few” outcome measures that fully and accurately cover an agency’s or program’s mission. In addition, many outcomes may be influenced by factors that are not under a federal program’s control. In combination, these tendencies may create the potential for perverse incentives in any system of goal-setting and performance measurement. **Figure 4** helps illustrate these possibilities.



Source: Adapted from Cynthia D. Fisher, Lyle F. Schoenfeldt, and James B. Shaw, *Human Resource Management*, 6th ed. (New York: Houghton Mifflin, 2006), p. 431.

If the outcome measures that are used for a program (right-hand circle) do not fully cover its mission (left-hand circle), they are considered “deficient.” Significantly, if the deficient outcome measures are emphasized over other considerations that are not measured or that may not be easily measurable, the omitted aspects of the mission might be neglected. Another problem may arise if one or more outcome measures are influenced by factors external to the program. In that case, the outcome measures are

¹²⁹ Quoted in Richard L. Lynch and Kelvin F. Cross, *Measure Up! Yardsticks for Continuous Improvement* (Cambridge, MA: Blackwell, 1991), p. 144.

“contaminated” (portion of right-hand circle). An organization’s personnel tend to dislike being held accountable for factors outside their control, and they may sacrifice performance of the mission in order to avoid sanctions or blame for missing targets that may suffer from deficiency or contamination. In either case, measures may create perverse incentives, if program evaluations or a system of performance measurement focus on outcome measures that are deficient or contaminated. The IRS cited this problem in the late 1990s, for example, in the wake of high-profile hearings of the Senate Finance Committee.¹³⁰ More recently, some observers alleged that the New York City Police Department’s “CompStat” approach to performance measurement created pressure that caused personnel to manipulate crime statistics and establish quotas for issuing tickets.¹³¹

Any of these issues might arise when the Administration and agencies undertake performance measurement, such as with the HPPG effort, or impact evaluation. If the possibility of perverse incentives and behaviors were of concern, Congress might consider pursuing strategies to become aware of them and potentially acting to reduce the risks of unintended consequences. From an oversight perspective, for example, Congress might examine the HPPGs that agencies and the Administration have chosen and ascertain whether key dimensions of an agency’s mission have been omitted. If some dimensions of an agency’s mission are not receiving attention, the problem of deficiency may drive behaviors within the agency in ways that are not consistent with congressional intent. Alternatively, if achievement of some HPPGs were judged to be substantially outside the control of an agency, the challenge of contamination may be evident. Inquiries with agencies or oversight hearings might provide venues for these kinds of discussions. Alternatively, during the appropriations process, appropriations committees might examine the HPPGs and accompanying performance metrics that are submitted by agencies along with agencies’ GPRA-required annual performance plans. If agency goal-setting seemed to be at risk of deficiency or contamination, a committee might investigate further to assess whether there is a risk of perverse incentives. Ultimately, if appropriations or authorizing committees became more broadly concerned with the potential for perverse incentives, they might consider pursuing remedies through lawmaking. For example, many organizations have attempted to formally address the problem of perverse incentives by focusing on multiple aspects of the mission and multiple stakeholder perspectives.¹³²

Making a Connection Between Evidence and Policy Making

Another recurring issue for Congress, and particularly for appropriations and authorizing committees, concerns how “evidence” may be defined and used in policy making. A related issue might be to raise the question when and whether evidence should be formally defined, as a matter of policy. In recent years, agencies, OMB, and advocacy groups have made occasional efforts to formally define evidence and how it should be used to allocate resources during the budget process and in allocating grant awards. In addition, the Obama Administration has offered views on what it considers to constitute evidence for policy making, especially with respect to impact evaluations. These views have been reflected in Administration proposals to Congress. Consequently, many of these efforts were viewed as high stakes in

¹³⁰ IRS Customer Service Task Force, *Reinventing Service at the IRS* (Washington: GPO, 1998), pp. 75-81.

¹³¹ John A. Eterno and Eli B. Silverman, “The NYPD’s Compstat: Compare Statistics or Compose Statistics?”, *International Journal of Police Science & Management*, vol. 12, no. 3, September 2010, pp. 426-449. For related media coverage, see William K. Rashbaum, “Retired Officers Raise Questions on Crime Data,” *New York Times*, February 7, 2010, p. A1; and Al Baker and Ray Rivera, “Secret Tape Has Police Pressing Ticket Quotas,” *New York Times* (New York edition), September 10, 2011, p. A1.

¹³² In response to similar concerns, some organizations have attempted to create a “balanced scorecard.” See Robert S. Kaplan and David P. Norton, “The Balanced Scorecard—Measures that Drive Performance,” *Harvard Business Review*, January-February 1992, pp. 71-79. For more on perverse incentives, see Steve Kerr, “On the Folly of Rewarding A, While Hoping for B,” *Academy of Management Journal*, vol. 18, no. 4, 1975, pp. 769-783.

nature. At the same time, efforts to define and use evidence have generated some controversy in specific policy areas and among evaluation methodologists.

Some efforts to formally define and use evidence during the policy making process appear to have been motivated by frustrations with past experience. Reflecting on past performance-related initiatives, some observers have expressed concerns that policy makers and agency personnel were not “using” performance information to make budget and management decisions.¹³³ However, there does not appear to be consensus about the nature of underlying problems in fostering the use of evaluation and measurement information by agencies, legislative bodies, and the public. Some commentators have attributed a perceived lack of use to, among other things, politics and ideology.¹³⁴ Another explanation for lack of use relates to credibility, when legislative bodies or other stakeholders may not trust that agencies or the President are presenting performance and evaluation information without bias.¹³⁵ Another possible explanation is that performance information may not be used, because key audiences may not find the kind of information that is presented to also be useful.¹³⁶

In addition to the explanations summarized above, it also is possible that some frustrations may be the result of unrealistic expectations. The perceived usefulness of performance measures and impact evaluations, unaccompanied by other kinds of program evaluation, may be limited. For example, if an impact evaluation or a performance measurement shows that a program is not meeting an outcome objective, it could be argued that the implication for budget and management decisions would be ambiguous. A decision to cut, maintain, or increase spending might depend on assessments of whether the program could be modified or more effectively implemented. Conversely, for a program that is perceived as performing well, it is conceivable the program’s budget could be cut without adverse performance consequences. Possibilities such as these suggest that a definition of evidence that focuses primarily on impact evaluations or performance measures could be considered to be narrow if it did not have support from other evaluations, logic modeling, and analysis that are focused on questions of “how” and “why” a program is performing as it is. Issues like these may become controversial in the context of legislative proposals and lawmaking. For example, several scholars criticized the Administration’s FY2010 proposal

¹³³ For example, see discussion in Shawn Zeller, “Performance Anxiety for ‘New’ Federal Standards,” *CQ Weekly*, March 30, 2009, pp. 708-709. Separately, GAO has done several surveys of federal managers and their use of performance measures. For a 2007 survey, see GAO, *Government Performance: Lessons Learned for the Next Administration on Using Performance Information to Improve Results*, GAO-08-1026T, July 24, 2008.

¹³⁴ For example, see Jerry Ellig, “Ten Years of Results From the Results Act,” working paper no. 10-21, Mercatus Center at George Mason University, May 2010, pp. 23-25, at <http://mercatus.org/publication/ten-years-results-results-act>. For discussion of such sentiments, see Shawn Zeller, “Performance Anxiety for ‘New’ Federal Standards,” *CQ Weekly*, March 30, 2009, p. 708.

¹³⁵ For discussion, see Carolyn Bourdeaux, “Legislative Influences on Performance-Based Budgeting Reform,” FRC Report No. 128, (Atlanta, GA: Andrew Young School of Policy Studies, Georgia State University, May 2006), pp. 27-30, at <http://aysps.gsu.edu/frc/files/report128.pdf>; and Clinton T. Brass, “Performance Management for Networked Government: A Transformation in Waiting?,” paper presented at the 2007 annual conference of the American Society for Public Administration, March 25, 2007, at <http://www.aspanet.org/source/Communities/documentArchive.cfm?section=Communities&CmtyId=178&ParID=309>. See also a classic 1990 study of congressional oversight, which found that Congress much more frequently used program evaluations done by its support agencies than it did evaluations done by executive agencies. Joel D. Aberbach, *Keeping a Watchful Eye: The Politics of Congressional Oversight* (Washington: Brookings, 1990), pp. 130-132. The study occurred before widespread proliferation of Inspectors General (IGs).

¹³⁶ Historically, many previous efforts in the executive branch to produce performance and evaluation information have been driven by the perceived needs of OMB and agency senior leaders. (For example, see GAO, *Performance Budgeting: Observations on the Use of OMB’s Program Assessment Rating Tool for the Fiscal Year 2004 Budget*, GAO-04-174, January 2004, p. 27.) However, it is not clear that past efforts at collecting performance information were preceded by attempts to ascertain the kinds of information that may be useful to other key audiences such as Congress, agency program managers, and nonfederal stakeholders who might assist Congress with its oversight function.

to Congress to scale-up a nurse-home visitation program on the basis of impact evaluation findings through randomized controlled trials. In a high-profile letter to the President, the scholars said that “we know of no evaluation scholar who would conclude that randomized trials provide sufficient insight for making program replication decisions,” because they argued that “the core features of a well-done randomized trial—a highly specified intervention, consistent implementation, and a specific target population—limit the ability to generalize its findings to diverse populations and diverse contexts.”¹³⁷

In that light, the kinds of “evidence” that may inform decisions about budgets, management, and policy arguably extend beyond estimating a program’s impact in a particular time and setting, or assessing whether performance goals are met. From a broader program evaluation perspective, other kinds of evidence that likely would be important considerations in policy making include evaluation of whether a program or policy could be replicated or scaled up; conditions under which a government intervention would likely result in a similar impact or outcome in other settings or times, or with other subjects; whether measured outcomes are “valid,” such as whether a test score actually measures student achievement or something else; whether a policy has been implemented as designed and intended; and whether a policy’s impacts occurred through the expected causal mechanism.¹³⁸ Even more broadly, from the perspective of multi-disciplinary policy analysis, there are many other conceptions of what could be considered as evidence for policy making. These include basic and applied research; forecasting and scenario planning; risk assessment; professional judgment from individual and group experience; theory; logic; intuition; and finally, an amalgam of values, priorities, and ethics.¹³⁹

Because conceptions of what constitutes evidence for policy making may vary depending on circumstances and values, several issues related to the role of evidence in policy making may be salient for Congress and its committees. If OMB or an agency appears to have adopted a formal definition for what constitutes evidence, a committee might scrutinize the definition in light of relevant legislative history. In the context of appropriations, this approach might involve examining the justification for funding changes and examining whether other conceptions of evidence also should be considered. From an oversight perspective, committees might examine how agencies are applying the concept of evidence when allocating grant and other awards. If agency practices were found to be problematic, committees might use oversight tools to influence agency behaviors or pursue remedies through lawmaking.

¹³⁷ Letter from Deborah Daro, Ph.D., et al. to President Barack Obama, April 21, 2009, p. 2, at <http://www.preventchildabuseny.org/pdf/CommentsonHomeVisitingProposal.pdf>. For background on this policy area, see CRS Report R40705, *Home Visitation for Families with Young Children*, by Emilie Stoltzfus and Karen E. Lynch. If the topic of evidence and policy making were of interest to Congress, Congress might look to the field of health care. Some conceptions of evidence in health care appear to be trending toward the critics’ contextual view, where decisions about “external validity,” such as judgments when specific interventions would work well in other contexts, would be made in a decentralized way by clinical decision makers, based on interpretations of evidence as applied to particular circumstances. Compare Deborah Daro, Ph.D., et al., p. 3, with Institute of Medicine, *Knowing What Works in Health Care: A Roadmap for the Nation* (Washington: National Academies Press, 2008), p. 21. Nevertheless, the issue in health care of how to define and use evidence from broad studies in the context of individual patients remains an area of active debate. For similar discussion about “comparative effectiveness research” and PPACA’s authorization to establish the Patient-Centered Outcomes Research Institute (PCORI; P.L. 111-148; 124 Stat. 727), see Rebecca Adams, “Health Policy by Comparison,” *CQ Weekly*, August 16, 2010, pp. 1968-1981.

¹³⁸ Using the terminology of program evaluation, these questions address matters of external validity and construct validity.

¹³⁹ See CRS Report RL33301, *Congress and Program Evaluation: An Overview of Randomized Controlled Trials (RCTs) and Related Issues*, by Clinton T. Brass, Erin D. Williams, and Blas Nuñez-Neto. Intuition often assumes importance in decision making when information is incomplete, consensus interpretations are lacking, the future is uncertain, or synthesis is necessary.

Potential Issues When Congress Focuses on Performance-Related Policy

Opportunities for Congressional Oversight: Lessons Learned from Past Efforts?

The Obama Administration's performance agenda arrives as the most recent in a line of presidential initiatives that have focused on government performance and related issues of management and budgeting. Past government-wide efforts include the Lyndon B. Johnson Administration's Planning-Programming-Budgeting System (PPBS, 1965); the Nixon/Ford Administrations' Management by Objectives (MBO, 1973); and the Jimmy Carter Administration's Zero-Base Budgeting (ZBB, 1977). More recently, the Clinton Administration undertook a National Performance Review (NPR, 1993) and the George W. Bush Administration pursued the President's Management Agenda (PMA, 2001) and PART. All were generally abandoned due to changes in Administration, a widespread perception of unrealistic ambitions, or lack of congressional support. With regard to congressional support, two former officials who were involved in the NPR and PMA efforts observed that recent Presidents were largely unsuccessful at garnering support from Congress for their management initiatives. In the authors' view, "the lack of congressional support has been a chronic limitation to gaining full acceptance by the agencies or to maintaining continuity beyond a particular administration."¹⁴⁰

Some observers have characterized a lack of congressional support as "political" in nature, in contrast with more "rational" presidential approaches that, in their view, would better serve the national interest.¹⁴¹ Others have focused on underlying policy disagreements to help explain conflicts. Still others have noted that, if adopted, many of the proposals associated with presidential initiatives would have reduced the congressional role in federal government management and thereby redistributed power from Congress to the President or agencies, prompting congressional resistance.¹⁴² Conflict between Congress and the President need not be one-way, of course. Congress has established the statutory framework for how the federal government is managed and how it performs, sometimes in ways resisted by Presidents. In the past, Congress has acted independently of, in reaction to, and in cooperation with Presidents on management and performance issues.¹⁴³

If Congress wants to examine the federal government's current policies that relate to government performance, both in law and as pursued by the Obama Administration, a starting point might be to identify "lessons learned" from past efforts in the United States and even overseas.¹⁴⁴ For example, as noted in this memorandum's section entitled "High-Priority Performance Goals (HPPGs)," efforts such as the Nixon and Ford Administrations' MBO and the UK Blair Government's PMDU may be relevant to the

¹⁴⁰ Jonathan D. Breul and John M. Kamensky, "Federal Government Reform: Lessons from Clinton's 'Reinventing Government' and Bush's 'Management Agenda' Initiatives," *Public Administration Review*, vol. 68 (November/December 2008), p. 1023.

¹⁴¹ See Jerry Ellig, "Ten Years of Results From the Results Act," working paper no. 10-21, Mercatus Center, May 2010, pp. 23-25; and Shawn Zeller, "Performance Anxiety for 'New' Federal Standards," *CQ Weekly*, March 30, 2009, pp. 708-709.

¹⁴² See Donald F. Kettl, "Beyond the Rhetoric of Reinvention: Driving Themes of the Clinton Administration's Management Reforms," *Governance*, vol. 7 (July 1994), p. 309; and CRS Report 31409, *The President's Management Agenda*, by Henry B. Hogue and Ronald C. Moe (archived and available on request from the first author).

¹⁴³ CRS Report RL30795, *General Management Laws: A Compendium*, by Clinton T. Brass et al.

¹⁴⁴ The literature on this subject is large and offers diverse and oftentimes conflicting perspectives. For a sampling of analyses, see Donald P. Moynihan, *The Dynamics of Performance Management: Constructing Information and Reform* (Washington: Georgetown University Press, 2008); Clinton T. Brass, "Performance Management for Networked Government: A Transformation in Waiting?," paper presented at the 2007 annual conference of the American Society for Public Administration, March 25, 2007; Beryl A. Radin, *Challenging the Performance Movement: Accountability, Complexity, and Democratic Values* (Washington: Georgetown University Press, 2006); and GAO, *Results-Oriented Government: GPRA Has Established a Solid Foundation for Achieving Greater Results*, GAO-04-38, March 2004.

Administration's HPPG agenda, with their emphases on performance measurement and accountability to the President or Prime Minister. Other topics of congressional study might involve exploring reasons why past presidential efforts did not generate congressional support and whether any of these lessons might apply to the Obama Administration's agenda for government performance. If common themes were identified, Congress might work through non-lawmaking means to improve its working relationship with the President, OMB, and agencies, or, alternatively, to clarify points of contention. Committees that focus on government-wide performance issues also might conduct oversight in areas of the Administration's agenda where they perceive risks of the Obama Administration not applying lessons from the past. If some topics were of substantial concern, Congress might conduct more in-depth oversight or consider changes to performance-related statutes, as discussed in the next two sections.

Oversight of the Administration's Performance Agenda: Coverage and Potential Gaps

When addressing the federal government's performance challenges, Mr. Zients has emphasized prioritization. These choices may not be obvious. On one hand, federal agencies, programs, and personnel must focus on accomplishing their current missions, subject to the resources and capabilities they already have. They work in myriad policy areas, including housing, transportation, and national defense. At the same time, improvement of the government's capabilities and practices is desired. For example, evaluations may be conducted and IT systems may be improved. However, the budgetary and staff resources that are available for capacity-building are limited. Finding a balance between "getting the work done," on one hand, and capacity-building, on the other, may be difficult. Given this tension, another issue that Congress might consider is the question whether the Obama Administration's agenda is covering the capacity-building areas that Congress views as most important, or, conversely, if there are gaps. For example, if committees that focus on government-wide performance perceive gaps, they might address the gaps by engaging with the Administration through communications, oversight, or lawmaking.

Generic Framework for Assessing Influences on Government Performance

Figure 5 (see p. 39) provides a generic framework that might be used at any given time as a lens through which to assess where OMB and agencies are focusing their activities and capacity-building priorities, with respect to goal-setting, program evaluation, and performance measurement. Consequently, the figure also might be used to help identify the presence of gaps. The figure builds on the logic model framework shown in **Box 3** and **Figure 2** (pp. 12-13), which may be used to integrate evaluation and measurement. In this case, however, **Figure 5** looks beyond a single program or agency. Instead, it illustrates how multiple agencies, policies, and nonfederal actors may work together to achieve desired policy outcomes or, conversely, work at cross-purposes.

In the figure, federal agencies "A" and "B" each implement several programs or policies through a causal chain of inputs, activities, and outputs. These causal chains may affect intermediate and end outcomes. The agencies and their activities are supported by a variety of mission-support functions like human resources and IT. Separately, some federal laws and policies that operate on their own, without substantial implementation by federal agencies, also may contribute to influencing outcomes. Tax expenditures, for example, are special tax breaks that may be used instead of direct expenditures to influence the behaviors of nonfederal entities in a way that will contribute to the achievement of policy goals.¹⁴⁵ The federal government also relies more directly on nonfederal actors to help accomplish policy goals. Many external factors that are not directly influenced by an agency or program also may influence outcomes. These

¹⁴⁵ For related discussion, see CRS Report RL33641, *Tax Expenditures: Trends and Critiques*, by Thomas L. Hungerford.

external factors may include, for example, broad economic conditions, natural disasters, or lack of cooperation from nonfederal stakeholders.

On the left side of the figure, a variety of performance-related processes and activities may be conducted by Congress, the President, and agencies. These processes may include, among other things, goal-setting, planning, and coordinating across organizational boundaries. Using procedural and other tools, Congress, the President, and agencies may thereby interact and influence how policies are implemented.

Example of Using the Framework: Federal Housing Policies

The framework might be applied in a specific policy context to help visualize these goal-setting, performance measurement, and evaluation activities, assess how they fit together, and identify whether they have gaps or raise further questions. For example, federal support for housing is a policy area that cuts across organizational and programmatic boundaries. The federal government also uses numerous policy tools to pursue housing policy objectives, including spending programs, tax expenditures, and regulations.¹⁴⁶ Using **Figure 5**, one might consider the efforts of two departments—the Departments of Housing and Urban Development (HUD) and Veterans Affairs (VA)—to reduce homelessness among veterans. The Obama Administration designated this combined effort as an HPPG for FY2011 and set a specific goal of reducing the number of homeless veterans to 59,000 over two years.¹⁴⁷

In **Figure 5**, HUD’s and VA’s activities might be viewed as programs “A” and “B,” respectively. The departments said they would work toward this goal by assisting 16,000 homeless veterans each fiscal year to move into permanent housing: 6,000 through HUD’s Continuum of Care grant programs, and 10,000 through their joint HUD-Veterans Affairs Supportive Housing (HUD-VASH) voucher program.¹⁴⁸ Therefore, from the perspective of the logic models in **Figure 5**, the departments identified *output* measures they would use: 6,000 and 10,000 veterans assisted per year, respectively. HUD and VA also estimated a combined, projected impact for an *end outcome*: a reduction of 135,000 homeless veterans over two years, compared to what would happen without the departments’ intervention.¹⁴⁹

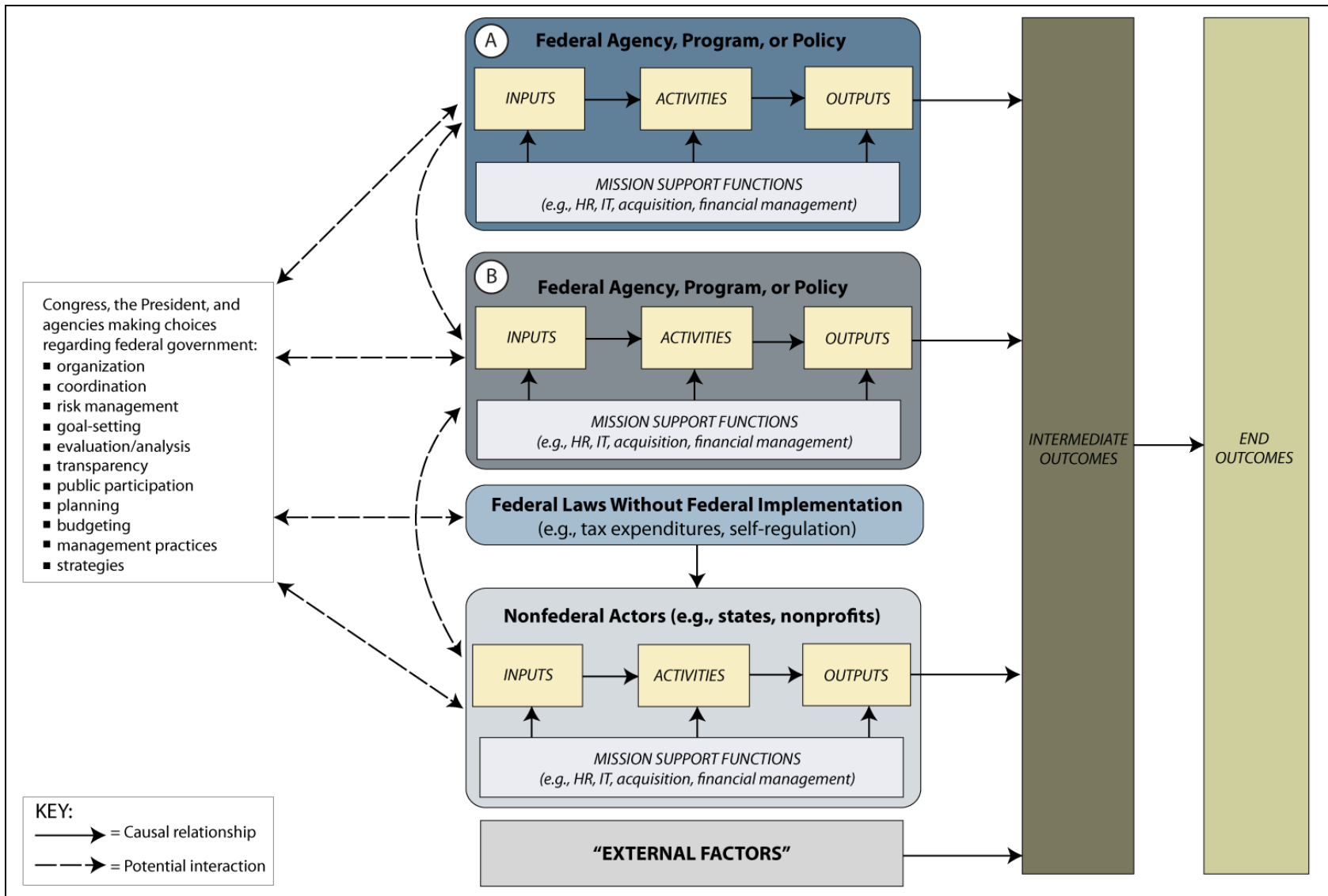
¹⁴⁶ All of these activities may contribute to accomplishing policy outcomes like supporting homeownership, making rental housing affordable for low-income families, and pursuing other, more intermediate outcomes. However, consensus may not always be apparent regarding which outcome goals are most important and how, or whether, the federal government should be involved in their pursuit. For background discussion, see CRS Report RL34591, *Overview of Federal Housing Assistance Programs and Policy*, by Maggie McCarty et al.; CRS Report R40210, *Preserving Homeownership: Foreclosure Prevention Initiatives*, by Katie Jones; U.S. Congressional Budget Office, *An Overview of Federal Support for Housing*, November 3, 2009; and OMB, *Analytical Perspectives, Budget of the U.S. Government, FY2011* (Washington: GPO, 2010), pp. 81-85, 91-92.

¹⁴⁷ HUD and VA said they would work together to reduce the number of homeless veterans from the then-current estimate of 131,000 to their goal of 59,000 by June 2012, in contrast with their estimate of 194,000 homeless veterans in June 2012 in the absence of the departments’ “intervention.” Thus, the departments estimated the projected impact of their efforts as 135,000: 194,000 without their intervention, minus 59,000 with their intervention. See OMB, *Analytical Perspectives, Budget of the U.S. Government, FY2011* (Washington: GPO, 2010), pp. 82, 85; and VA, *FY 2011 Funding and FY 2012 Advance Appropriations Request*, vol. II, pp. 1A-2 and 1-J12 - 1J-14, at http://www.va.gov/budget/docs/summary/Fy2011_Volume_2-Medical_Programs_and_Information_Technology.pdf.

¹⁴⁸ *Ibid.*, and see also supporting information in the HUD FY2011 budget summary (<http://hud.gov/budgetsummary2011/full-budget-2011.pdf>, p. 32) and VA FY2011 budget summary (http://www.va.gov/budget/docs/summary/Fy2011_Volume_1-Summary_Volume.pdf, p. 1A-3). For background information on the Continuum of Care and HUD-VASH programs, see <http://www.hud.gov/offices/cpd/homeless/programs/coc/>, and <http://www.hud.gov/offices/pih/programs/hcv/vash/>, respectively.

¹⁴⁹ Most of the Administration’s other HPPGs that concern housing are focused on activities and outputs, as opposed to outcomes. See OMB, *Analytical Perspectives, Budget of the U.S. Government, FY2011* (Washington: GPO, 2010), pp. 81-82.

Figure 5. Potential Framework for Assessing Influences on Government Performance



Source: CRS.

However, it is not clear how serving a total of 32,000 veterans over two years would reduce veteran homelessness by this overall number, unless other activities, agencies, and external factors were also instrumental in the effort. VA's FY2011 budget submission detailed other proposed activities that would contribute to addressing veteran homelessness,¹⁵⁰ suggesting that the HPPG output measures and outcome goal may provide only a partial picture of how HUD and VA planned to accomplish the Administration's stated outcome goal. Other considerations that remain unclear from HUD's and VA's descriptions include what role external factors might play in decreasing the number of homeless veterans over time, separate from federal activities, and how the departments would validly or credibly estimate the impact of their efforts, separate from the influence of other factors. Other kinds of evaluations and performance measures also might provide information about service quality and the appropriateness of the departments' models of service delivery for the needs of clients. In combination, questions such as these suggest that it may be difficult to assess whether accomplishment of the Administration's HPPG output goals, such as assisting veterans in the specified ways, is significantly contributing to accomplishment of the Administration's outcome goal, reducing veteran homelessness. Consequently, questions like these may present challenges for Congress when deliberating about agency budgets, programs' authorizing statutes, and oversight priorities, and also present similar challenges for agencies, OMB, and the President during policy development and implementation.

Assessing Administration Initiatives for Coverage and Potential Gaps

The figure's framework also may be used more generally, when Congress is overseeing or attempting to influence the performance-related policies that are being implemented by an Administration across the entire executive branch. Several observations might be made specifically about the Obama Administration's executive-branch-wide performance policies. For example, the Administration appears to be pursuing substantial capacity-building in the realm of performance measurement through its development of HPPGs and mission-support-related dashboards. At the same time, most of the HPPGs are focused on individual agencies, rather than across organizational boundaries. The Administration also proposed a number of impact evaluations that would focus on intermediate and end outcomes. However, other kinds of program evaluation, which focus on understanding how and why programs are working, generally have not received explicit emphasis from the Administration. For example, the Administration's HPPG and impact evaluation efforts do not formally appear to be integrated through the use of logic models. Logic models might be used to improve understanding of how achieving activity- and output-focused HPPGs, such as the number of clients served, may assist in achieving intermediate and end outcomes, like reducing the homelessness rate.¹⁵¹ The Administration's online performance portal eventually may provide such evaluations and analyses, but it remains to be seen how the Administration will use the planned website.

The Administration also appears to be emphasizing some performance-related processes, such as goal-setting. However, it is less clear whether the Administration is focusing prominently on other processes. For example, the Administration said it would involve Congress and the public in goal-setting for HPPGs. However, the HPPG effort appears to have operated substantially within the Administration's "deliberative process," outside of public view. If public and congressional participation in goal-setting

¹⁵⁰ See VA's more detailed budget proposal, at http://www.va.gov/budget/docs/summary/Fy2011_Volume_2-Medical_Programs_and_Information_Technology.pdf, pp. 1A-2, 1A-12 - 1A-13.

¹⁵¹ For background, see Kathryn E. Newcomer and Mary Ann Scheirer, *Using Evaluation to Support Performance Measurement: A Guide for Federal Executives* (Washington: IBM Center for the Business of Government, January 2001), at http://www.businessofgovernment.org/publications/grant_reports/details/index.asp?GID=25.

were of interest to certain congressional committees, they might inquire how, or if, the public and Congress were engaged when agencies developed HPPGs. In addition, it is not clear if the Administration is planning significant efforts to evaluate or analyze whether tax expenditures assist in the achievement of desired outcomes. In areas like these, Congress might conduct oversight to gain a more specific understanding of the Administration's plans and priorities. Several related processes also could be subjects of oversight or lawmaking, if they were perceived as gaps or needing to be addressed. These topics might include whether agencies have capacity to effectively undertake and use evaluations and measurement; how to have agencies produce evaluations and measurements that may be perceived by Congress as credible and not "politicized"; whether the needs of key audiences are being served; how to structure the production of evaluation information to serve the needs of key audiences; whether the potential exists in some cases for perverse incentives; and whether the Administration's efforts might be complementary or duplicative with "key national indicators" that, if funded, may be developed.¹⁵²

Examining GPRA and E.O. 13450

Much of the foregoing discussion may have relevance, if Congress wants to use the Administration's agenda as an opportunity to evaluate the design and implementation of GPRA and E.O. 13450. Specifically, Congress might evaluate how the law and executive order have been used in practice, how they could be used by the legislative and executive branches in the face of current and future challenges, and any implications for lawmaking, oversight, or study. In addition, GPRA was substantially amended in the 111th Congress (H.R. 2142, P.L. 111-352), and Congress may exercise oversight over the law's design and implementation.

Enacted in 1993, GPRA typically has been regarded as the first general framework for goal-setting and performance measurement to be enacted into law, as opposed to implemented by presidential directive.¹⁵³ Congress had multiple objectives in enacting GPRA, as noted in the Senate committee report that accompanied the legislation. These included addressing the needs especially of Congress and program managers, because "congressional policymaking, spending decisions, and oversight are all seriously handicapped by the lack both of sufficiently precise program goals and of adequate performance information," and "[f]ederal managers, too, are greatly disadvantaged in their own efforts to improve program efficiency and effectiveness by that same lack of clear goals and information on results."¹⁵⁴ The committee commented on these needs, saying "[t]he Committee realizes that, as valuable as the goal-setting and performance reporting under [GPRA] will be to Congress in its [policymaking] and oversight roles, its greatest value will likely be to the program managers themselves."¹⁵⁵ Furthermore, "performance measurement under this Act should focus on those measures and that information useful to and used by program managers," an "emphasis [that] very much relies on a bottom-up approach, rather than a top-down imposition."¹⁵⁶ In the face of top-down imposition that could occur with a new President, the committee said "[e]ven when a change in Administration brings about a shift in political philosophy, the program's missions and long-term general goals remain largely intact. The priorities and means of

¹⁵² Section 5605 of the Patient Protection and Affordable Care Act (P.L. 111-148) established a Commission on Key National Indicators and authorized \$70 million to be appropriated for a "Key National Indicator System." Congressional leaders appointed eight members of the commission; see "Congress Appoints Key National Indicators Commission," December 16, 2010, at <http://www.stateoftheusa.org/content/commission-on-key-national-ind.php>.

¹⁵³ As noted earlier, the law tended to emphasize performance measurement more than program evaluation.

¹⁵⁴ U.S. Congress, Senate Committee on Governmental Affairs, *Government Performance and Results Act of 1993*, report to accompany S. 20, 103rd Cong., 1st sess., June 16, 1993, S.Rept. 103-58 (Washington: GPO, 1993), p. 3.

¹⁵⁵ *Ibid.*, p. 9.

¹⁵⁶ *Ibid.*, p. 27.

achieving those goals, and the approach to problem solving, may vary significantly, but the long-term goals usually remain the same.”¹⁵⁷

Experience from the last three presidencies suggests that GPRA’s framework gave Presidents and OMB substantial influence over the law’s de facto emphases. Some might argue, for example, that aspects of the law primarily served the interests of the institutional presidency. For example, the law requires that annual performance plans, with corresponding goals and metrics, accompany presidential budget proposals to Congress. However, the law has not required revision of these annual plans to reflect annual congressional budget decisions, which might establish a baseline to which agencies and the President could be held accountable. Until GPRA was amended in the 111th Congress, the law was silent about using and reporting these goals, metrics, and evaluations during policy implementation. GPRA also required agencies to “consult” with Congress and solicit views from stakeholders about strategic plans, which were required to be revised at least every three years.¹⁵⁸ Notably, these consultations may have been infrequent and were not subject to transparency requirements, such as requirements to identify how, or from whom, views were solicited. Furthermore, the law did not require transparency or public participation in the selection of goals and metrics for inclusion in annual performance plans. The process of notice and comment rulemaking may be a point of contrast, under which agencies publish draft rules, receive comments from the public, and then address significant issues raised in the comments. Hence, these topics were left to the discretion of the President, OMB, and political appointees in agencies, with less opportunity for “fire alarm” oversight to assist Congress in identifying agency or Administration behaviors that may deviate from congressional expectations or intent.¹⁵⁹ Other topics, such as GPRA’s primary emphasis on performance measurement (as opposed to program evaluation) and goal-setting (as opposed to operational planning, risk management, and “institutionalizing imagination”) also might be considered,¹⁶⁰ in addition to topics discussed in the previous section on coverage and potential gaps in the Administration’s agenda.

E.O. 13450 appears to codify certain aspects of the PART and explicitly authorizes centralized decision making in OMB and the White House about many aspects of agencies’ GPRA-related efforts. In contrast with Senate-confirmed agency “chief officers” (e.g., chief financial officers), the PIC and its PIOs have little to no independent authority under the E.O. They are tasked with providing advice to agency heads, the Director of OMB, and the President in many aspects of OMB’s and presidential decision making. Given the challenges of congressional oversight of the White House’s deliberative process, congressional oversight of E.O. 13450 and the activities of PIOs and the PIC may present challenges. Some parts of the

¹⁵⁷ Ibid., p. 15.

¹⁵⁸ GAO has suggested that strategic plans be updated every four years instead of every three years, to coincide with presidential administrations and reduce agency workload. (See GAO, *Results-Oriented Government: GPRA Has Established a Solid Foundation for Achieving Greater Results*, GAO-04-38, March 2004, pp. 10, 110, and 129.) On the other hand, congressional intent for GPRA appeared to contemplate some agency independence from the President when producing strategic plans, raising the possibility that four-year updates might “politicize” GPRA to reflect the President’s policy views instead of agencies’ relatively stable statutory missions. GAO’s recommendation, if implemented, might have had the effect of decreasing the frequency of agency consultations with Congress about strategic plans from once every three years to once every four years.

¹⁵⁹ As formulated in a classic article, “fire-alarm” oversight refers to “rules, procedures, and informal practices that enable individual citizens and organized interest groups to examine administrative decisions (sometimes in prospect), to charge executive agencies with violating congressional goals, and to seek remedies from agencies, courts and Congress itself.” See Mathew D. McCubbins and Thomas Schwartz, “Congressional Oversight Overlooked: Police Patrols versus Fire Alarms,” *American Journal of Political Science*, vol. 28, February 1984, p. 186.

¹⁶⁰ The reference to “institutionalizing imagination” comes from the 9/11 Commission’s report and recommendations, which advocated routinizing risk identification, assessment, and management within and across organizations. (U.S. National Commission on Terrorist Attacks Upon the United States, *The 9-11 Commission Report* (Washington: GPO, 2004), p. 344.) If the topic were of interest to Congress, approaches that borrow from “enterprise risk management” (ERM) might be explored.

executive order, including those establishing PIOs and the PIC, were enacted into law with the most recent amendment to GPRA.

In the 111th Congress, some hearings focused on the Administration's agenda, as noted earlier. Separately, the Senate Committee on the Budget established a five-Senator Task Force on Government Performance. The task force held hearings and focused on several management and performance issues, including performance measurement, mission-support functions, and cross-agency coordination.¹⁶¹ In the House, the introduced version of H.R. 2142 would have enacted into law many aspects of the PART and E.O. 13450. The bill was referred to a subcommittee of the House Committee on Oversight and Government Reform. It passed the House on June 16, 2010, after significant changes in subcommittee and full committee markups.¹⁶² The House-passed version would have made significant modifications to GPRA by enacting into law aspects of E.O. 13450 and the Obama Administration's approach to government performance. On September 28, 2010, several Senators introduced S. 3853, which also would have made significant modifications to GPRA and adopted portions of E.O. 13450 and the Obama Administration's approach. The Senate Committee on Homeland Security and Governmental Affairs marked up H.R. 2142 on September 29, 2010, substituting the text of S. 3853 for the contents of the House-passed version, and ordered the measure reported out of committee. The committee reported the measure on December 7, 2010, and the Senate passed the reported measure by unanimous consent on December 16, 2010.¹⁶³ On December 17, 2010, the House voted on the Senate amendment under suspension of the rules, but the bill did not receive the two-thirds vote necessary for passage. Four days later, on December 21, 2010, the House voted on the Senate amendment under a rule and passed the measure with a vote of 216-139. President Obama signed the legislation, the GPRA Modernization Act of 2010, on January 4, 2011 (P.L. 111-352).

¹⁶¹ For more information, see <http://budget.senate.gov/democratic/TaskForce.html>.

¹⁶² For an analysis of the subcommittee-reported version of H.R. 2142, see CRS Congressional Distribution Memorandum, "Analysis of Subcommittee-Reported H.R. 2142 (111th Congress) and Related Issues," May 19, 2010, by Clinton T. Brass. The memorandum is available upon request from the author.

¹⁶³ The committee filed a written report on December 16, 2010. U.S. Congress, Senate Committee on Homeland Security and Governmental Affairs, *GPRA Modernization Act of 2010*, report to accompany H.R. 2142, 111th Cong., 2nd sess., December 16, 2010, S.Rept. 111-372 (Washington: GPO, 2010).
