April 2013

INFORMATION TECHNOLOGY

Consistently Applying Best Practices Could Help IRS Improve the Reliability of Reported Cost and Schedule Information

GAO

Accountability ★ Integrity ★ Reliability

GAO-13-401

INFORMATION TECHNOLOGY

Consistently Applying Best Practices Could Help IRS Improve the Reliability of Reported Cost and Schedule Information

Why GAO Did This Study

IRS relies extensively on IT systems to annually collect more than $2 trillion in taxes, distribute more than $300 billion in refunds, and carry out its mission of providing service to America's taxpayers in meeting their tax obligations. In fiscal year 2012, the agency spent about $2.5 billion for IT. Given the size and significance of IRS's IT investments, and the challenges inherent in successfully delivering these complex IT systems, it is important that Congress be provided ongoing, accurate, and objective information on the progress toward completion and the risks facing these projects.

Accordingly, GAO's objectives, among other things, were to (1) summarize the reported cost and schedule performance for IRS's major IT investments, and (2) for selected investments, determine the reliability of reported cost and schedule variances.

What GAO Recommends

GAO recommends that IRS improve the reliability of reported cost and schedule information by addressing weaknesses in future updates of cost and schedule estimates. GAO also recommends that IRS ensure projects consistently follow guidance for updating performance information 60 days after completion of an activity and develop and implement guidance that specifies best practices to consider when determining projected amounts. IRS agreed with three of GAO's four recommendations and partially disagreed with the fourth recommendation related to guidance on projecting cost and schedule amounts. GAO continues to believe this recommendation is warranted.

What GAO Found

According to the Internal Revenue Service (IRS), 12 of its 20 major information technology (IT) investments were within 10 percent of cost and schedule estimates or significantly below cost between October 2011 and October 2012. For the remaining 8 investments, 3 were reported as being significantly over cost and 5 were reported as being significantly behind schedule. Reported reasons for these significant variances include unplanned work activities, procurement delays, and additional costs associated with terminating an investment that was being replaced.

The reliability of the reported variance information—which is dependent upon having (1) a reliable cost estimate and a well-constructed and controlled schedule estimate, and (2) a process for determining variances using estimates and comparing them to actual or projected amounts—varied for the seven investments reviewed. Specifically, the cost estimates for CADE 2 and IRDM, and the schedule estimates for CADE 2, IRDM, and ACA were more favorable than for the remaining investments because they were more consistent with best practices (see table for assessments of cost estimates).

Status of Whether Each Key Investment Met the Four Characteristics of a Cost Estimate

	Comprehensive	Well-documented	Accurate	Credible
Affordable Care Act	◑	◑	◑	◔
Customer Account Data Engine 2	◕	◕	◑	◕
e-Services	◑	◔	◑	n/aª
Information Reporting and Document Matching	◕	◕	◕	◑
IRS.gov	◕	◑	◑	◔
Modernized e-File	◔	○	◔	◔
Return Review Program	◕	◑	◕	◔

Source: GAO analysis of IRS documentation.

ª Because the cost estimate for e-Services is only for maintenance costs, GAO determined that the credible criterion is not appropriate in the assessment.

Key: ●=Fully met—fully implemented the practices; ◕=Substantially met—implemented a large portion of the practices; ◑=Partially met—implemented about half of the practices; ◔=Minimally met—implemented a small portion of the practices; and ○=Not met—did not implement the practices or provided evidence that it only minimally implemented.

In addition, regarding IRS's process for determining variances, the agency generally determined investment cost and schedule variances for completed activities with actual amounts—although in about 25 percent of the cases it did not do so within the 60-day time frame specified in the Department of Treasury's guidance. Further, while IRS determined variances using projected cost and schedule for in-process activities, the guidance for doing so did not specify how projected amounts should be determined. This introduces the risk that projected amounts will not consistently reflect best practices and these amounts will therefore remain questionable.

Contents

Figures

Abbreviations

ACA	IRS Affordable Care Act
BSM	Business Systems Modernization
CADE	Customer Account Data Engine
CIO	Chief Information Officer
CMMI	Capability Maturity Model Integration
IMF	Individual Master File
IRDM	Information Reporting and Document Matching
IRS	Internal Revenue Service
IT	information technology
MeF	Modernized e-File
OMB	Office of Management and Budget
RRP	Return Review Program
Treasury	Department of Treasury

G A O

Accountability * Integrity * Reliability

United States Government Accountability Office
Washington, DC 20548

April 17, 2013

Congressional Committees

The Internal Revenue Service (IRS) relies extensively on information technology (IT) systems to annually collect more than $2 trillion in taxes, distribute more than $300 billion in refunds, and carry out its mission of providing service to America's taxpayers in meeting their tax obligations. For fiscal year 2013, the agency's budget request is $2.5 billion for IT, about 20 percent of IRS's total budget request for that fiscal year.

Given the size and significance of IRS's IT investments, and the challenges inherent in successfully delivering these complex IT systems, it is important that Congress be provided ongoing, accurate, and objective information on the progress toward completion and the risks facing these projects. Accordingly, we were directed to review the cost, schedule, and risks of IRS's major IT investments by the Conference Report accompanying the Consolidated Appropriations Act, 2012.[1] Specifically, our objectives were to (1) summarize the reported cost and schedule performance for all IRS's major IT investments; (2) for selected investments—IRS Affordable Care Act (ACA),[2] Customer Account Data Engine (CADE) 2, e-Services, Information Reporting and Document Matching (IRDM), IRS.gov,[3] Modernized e-File (MeF), and Return Review Program (RRP)—determine the reliability of reported cost and schedule

---

[1]H.R. Rep. No. 112-331, at 901 (2011) (Conf. Rep.); Consolidated Appropriations Act, 2012, Pub. L. No. 112-74, 125 Stat. 786, 884, div. C, title I (2011).

[2]IRS Affordable Care Act (ACA) refers to IRS's IT efforts to implement provisions of the Patient Protection and Affordable Care Act, Pub. L. No. 111-148, 124 Stat. 119 (2010), as amended by the Health Care and Education Reconciliation Act of 2010, Pub. L. No. 111-152, 124 Stat. 1029 (2010).

[3]IRS.gov is used interchangeably to represent the Public User Portal and the entire investment which is comprised of three portals—the Public User Portal, Registered User Portal, and Employee User Portal—also collectively known as the Integrated Enterprise Portal. For our review, we looked at the entire investment known as Integrated Enterprise Portal.

GAO-13-401 IRS Major IT Investments

variances;[4] and (3) identify IRS's reported risks and mitigation strategies for the selected investments.

To address our first objective, we tracked the cost and schedule performance for each of IRS's major IT investments from October 2011 to October 2012, as reported by IRS. We also obtained the reasons for significant variances from the Chief Information Officer (CIO) ratings on the Office of Management and Budget (OMB) IT Dashboard.[5]

To address our second objective, we analyzed the cost and schedule information of seven major investments using criteria from the GAO cost and schedule assessment guides. Specifically, for cost, we evaluated each investment's cost estimate relative to the four characteristics of a reliable estimate, as defined in the Cost Estimating and Assessment Guide.[6] For schedule, we determined whether the integrated master schedule for each investment was well-constructed and controlled, two of the four characteristics of a schedule defined in the GAO Schedule Assessment Guide that were most relevant to our review.[7] We also reviewed documentation including the four quarterly reports submitted to the appropriations committees between July 2012 and January 2013 and interviewed officials to determine whether IRS was reporting cost and schedule variance information in a timely manner and calculating it in accordance with best practices. For our third objective, we identified each investment's key risks and the mitigation strategies through reviews of documents such as risk logs, and interviews with relevant officials.

We conducted this performance audit from July 2012 to April 2013, in accordance with generally accepted government auditing standards. Those standards require that we plan and perform the audit to obtain

---

[4]Variances measure differences between estimated cost and schedule goals and actual or projected cost and schedule at one point in time. As such, variances are good indicators of performance or progress in meeting goals.

[5]The IT Dashboard is a public website established by OMB in June 2009 that provides detailed information on about 800 federal IT investments, including assessments of actual performance against cost and schedule targets. It is intended to improve the transparency and oversight of these investments.

[6]GAO, *GAO Cost Estimating and Assessment Guide Best Practices for Developing and Managing Capital Program Costs*, GAO-09-3SP (Washington, D.C.: March 2009).

[7]GAO, *GAO Schedule Assessment Guide Best Practices for Project Schedules (Exposure Draft)*, GAO-12-120G (Washington, D.C.: May 2012).

sufficient, appropriate evidence to provide a reasonable basis for our findings and conclusions based on our audit objectives. We believe that the evidence obtained provides a reasonable basis for our findings and conclusions based on our audit objectives. Details on our objectives, scope, and methodology can be found in appendix I.

Background

The mission of IRS, a bureau within the Department of the Treasury (Treasury), is to provide America's taxpayers top quality service by helping them understand and meet their tax responsibilities and by applying the federal tax laws with integrity and fairness to all. In carrying out its mission, IRS annually collects over $2 trillion in taxes from millions of individual taxpayers and numerous other types of taxpayers and manages the distribution of over $300 billion in refunds. To guide its future direction, the agency has two strategic goals: (1) improve taxpayer service to make voluntary compliance easier and (2) enforce the law to ensure everyone meets their obligations to pay taxes.

IRS is organized into four primary operating divisions to meet the needs of specific taxpayer segments.

- The Wage and Investment Division services individual taxpayers and provides the information support, and assistance these taxpayers need to fulfill their tax obligations.

- The Small Business and Self-Employed Division services all fully or partially self-employed individuals and corporations and partnerships with assets of $10 million or less.

- The Large Business and International Division services corporations and partnerships with assets greater than $10 million.

- The Tax Exempt and Government Entities Division services a large and unique economic sector of organizations, which include pension plans, exempt organizations, governmental entities, and tax-exempt bond issuers.

IRS's Information Technology organization[8] is responsible for delivering IT services and solutions to support tax administration as well as the operations of the broader organization. Information Technology is headed by the Chief Technology Officer who reports to the Deputy Commissioner for Operations Support.[9] It includes nine subordinate offices, among those are:

- The Associate CIO for Affordable Care Act Program Management Office is responsible for setting the overall direction, providing day-to-day management and oversight and delivery of IT contribution for implementing provisions of the Patient Protection and Affordable Care Act. This Associate CIO is responsible for the Affordable Care Act program.

- The Associate CIO for Modernization Program Management Office provides state-of-the-art individual taxpayer account processing and data-centric technologies to improve service to taxpayers and enhance IRS tax administration. This Associate CIO is responsible for the CADE 2 program.

- The Associate CIO for Applications Development builds, tests, delivers, and maintains integrated information applications systems that will support modernized systems and the production environment. This Associate CIO is responsible for the e-Services, MeF, IRDM, and RRP programs, which we focused on in our review.

- The Associate CIO for Strategy and Planning facilitates the alignment of IT and business through strategic planning and financial management practices that offer transparency of overall IT demand, supply, and the value of IT investments. This office includes a suboffice that is responsible for, among other things, serving as the primary interface with Treasury's capital planning and investment control organizations to coordinate actions including baseline change requests, budget formulation documents, and IT Dashboard reporting. It also includes an Estimation Program Office, which is made up of a group of cost estimation experts that assists project teams by

---

[8]This organization was formerly referred to as the Modernization and Information Technology Services organization. Its name was officially changed on July 1, 2012.

[9]IRS's Chief Technology Officer also acts as the agency's Chief Information Officer.

developing and updating cost estimates using a standard documented process.

Figure 1: IRS's Information Technology Organization
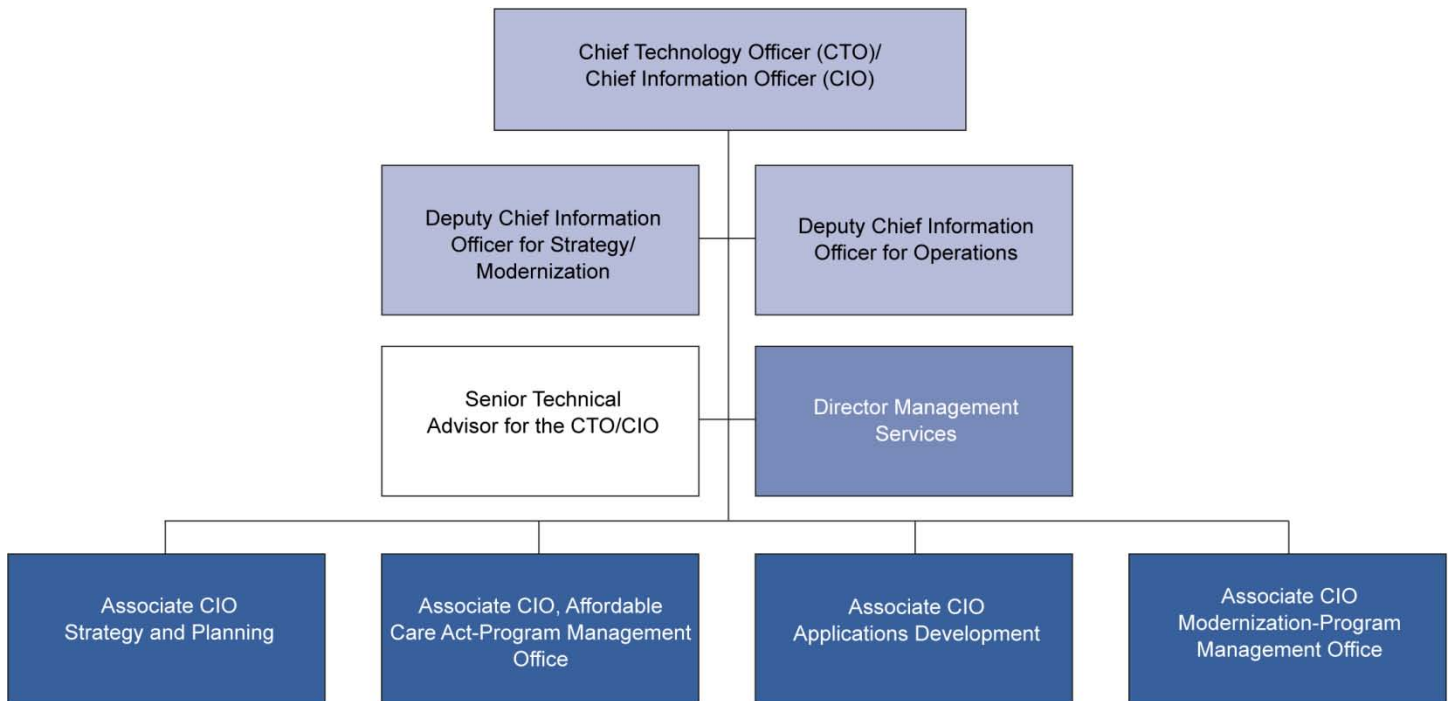
Source: IRS

IRS Relies on IT to Carry Out Its Mission

IT plays a critical role in enabling IRS to carry out its mission and responsibilities. For example, the agency relies on information systems to process tax returns, account for tax revenues collected, send bills for taxes owed, issue refunds, assist in the selection of tax returns for audit, and provide telecommunications services for all business activities, including the public's toll-free access to tax information.

For fiscal year 2012, IRS spent about $2.1 billion on IT investments. This included about $1.4 billion, or 67 percent, of the total amount, for 20 major systems, and about $700 million, or 33 percent, of the total amount for 125 nonmajor systems.[10] For IRS, a major investment is one that

---

[10]IRS's IT budget request for fiscal year 2013 is about $2.5 billion.

costs $10 million in either current year or budget year, or $50 million over the 5-year period extending from the prior year through the budget year +2. The following table provides a list of the 20 major investments and their descriptions.

Table 1: IRS's Major Investments and Their Descriptions

Investment	Description
Account Management Services	Enhances customer support by providing applications that enable IRS employees to access, validate, and update individual taxpayer accounts on demand.
Affordable Care Act	Allows the IRS to continue the development of new systems and modification of existing systems required to implement provisions of the Affordable Care Act.
Current Customer Account Data Engine (CADE)	Initiated in 1999, was intended to replace the Individual Master File (IMF) processing system and house tax information for more than 200 million taxpayers while providing faster return processing and refunds. Starting in 2005, CADE processed and recorded tax return and tax account information for millions of individual taxpayers with simple returns. Because the development of the system took longer than anticipated, IRS decided to review its modernization approach and develop a new program (CADE 2). IRS retired Current CADE at the end of December 2011.
CADE 2	Is IRS's new program for replacing IMF. In January 2012, IRS delivered the initial phase of the program which shifted the weekly processing and posting of individual taxpayer accounts to a daily process. This enhanced tax administration and improved service by enabling faster refunds for more taxpayers, allowing more timely account updates, and faster issuance of taxpayer notices. Through CADE 2, IRS is also establishing a database to house all individual taxpayer accounts. Additional phases are being planned to further improve customer service and compliance and address financial and security material weaknesses.
Electronic Fraud Detection System	Assists in detecting fraud at the time that tax returns are filed in order to eliminate the issuance of fraudulent tax refunds.
e-Services	Comprises several web-based self-assisted services that are intended to allow authorized individuals to do business with the IRS electronically.
Foreign Account Tax Compliance Act	Intended to implement provisions of the Foreign Account Tax Compliance Act regarding financial institutions reporting to IRS information about financial accounts held by U.S. taxpayers, or foreign entities in which U.S. taxpayers hold a substantial ownership interest.
Implement Return Review Program (replaces Electronic Fraud Detection System)	Currently under development, is intended to maximize fraud detection at the time that tax returns are filed to eliminate issuance of questionable refunds.
Individual Master File	Represents the authoritative data source for individual tax account data. All other IRS information systems that process IMF data depend on output from this source. This investment is a critical component of IRS's ability to process tax returns.
Information Reporting and Document Matching	Intended to establish a new business information matching program in order to increase voluntary compliance and accurate income reporting.
Integrated Customer Communication Environment	Includes several projects that are intended to simplify voluntary compliance using voice response, Internet, and other computer technology such as the Modernized Internet Employee Identification Number, which allows third parties to act on the behalf of taxpayers.
Integrated Data Retrieval System	Intended to provide systemic review, improve consistency in case control, alleviate staffing

Investment	Description
	needs, issue notices to taxpayers, and allow taxpayers to see status of refunds. It is a mission-critical system used by 60,000 IRS employees.
Integrated Financial System/CORE Financial System	Used by IRS for budget, payroll, accounts payable/receivable, general ledger functions, and financial reporting; also used to report on the cost of operations and to manage budgets by fiscal year.
Integrated Submission and Remittance Processing System	Processes paper tax returns, and updates tax forms to comply with tax law changes.
IRS End User Systems and Services	Supports products and services necessary for daily functions for over 100,000 IRS employees at headquarters and field sites.
IRS Main Frames and Servers Services and Support	Intended to support the design, development, and deployment of server storage infrastructures, software, databases, and operating systems.
IRS Telecommunications Systems and Support	Supports IRS's broad and local network infrastructure such as servers, and switches for voice, data, and video servicing of about 1,000 IRS sites.
IRS.gov – Portal Environment	Provides web-based services such as tax filing and refund tracking, to internal and external users, such as IRS employees and other government agencies, taxpayers, and business partners.
Modernized e-File	Provides a secure web-based platform for electronic tax filing of individual and business tax and information returns by registered Electronic Return Originators.
Service Center Recognition/Image Processing System	Used as a data capture, management, and image storage system using high-speed scanning and digital imaging to convert data from the 940, 941, K-1, and paper returns from Information Returns Processing, into electronic format.

Source: GAO analysis of IRS data.

GAO Has Previously Reported on IRS's Key Major IT Investments

We have previously reported on IRS's ongoing effort to modernize its tax administration and internal management systems, the agency's investment management process and key major investments—CADE 2, IRDM, and ACA.

- We have reported on the agency's ongoing effort to modernize its tax administration and internal management systems, known as its Business Systems Modernization (BSM) program.[11] In particular, between 1999 and 2011, we focused on program management capabilities and controls that are critical to the effective management of this program, such as cost and schedule estimates, requirements

---

[11]IRS's BSM program involves the development and delivery of a number of modernized tax administration and internal management systems, as well as core infrastructure projects that are intended to replace the agency's aging business and tax processing systems. A long history of continuing delays and design difficulties and their impact on IRS's operations led us to designate the program as a high-risk area in 1995.

development and management, and post-implementation reviews of deployed projects. Accordingly, we made numerous recommendations aimed at strengthening these controls and capabilities. In February 2013, we removed the BSM program from the GAO high-risk list—after 18 years—because of IRS's progress in addressing the significant weaknesses that led to the high-risk designation, and its commitment to sustaining progress in the future.[12] Nonetheless, there are still many risks associated with the program and we will continue to closely monitor it. Two of the investments in our review—CADE 2 and MeF—are part of the BSM program.

- Regarding investment management, in July 2011 we found that IRS had established most of the foundational practices needed to manage its IT investments by executing 30 of the 38 key practices identified by the information technology investment management framework, including all of those needed for effective project oversight.[13]

- In March 2011, we provided an update on IRS's implementation of its CADE 2 strategy for managing individual taxpayer accounts, noting weaknesses in the agency's efforts to improve the credibility of cost estimates.[14] We recommended that IRS (1) improve the credibility of revised cost estimates by including all costs or provide a rational for excluding costs, and adjust costs for inflation; and (2) identify all of the second phase benefits, set the related targets, and identify how systems and business process might be affected. IRS agreed with our recommendations. The agency took steps to address the recommendations to improve the cost estimate and is currently working to address those related to the second phase of CADE 2.

- Further, we reported in October 2011 that CADE 2 was one of seven investments considered to be successfully acquired in that it best achieved its respective cost, schedule, scope, and performance goals. Officials cited active engagement with program stakeholders as a critical factor to the success of those investments and, overall nine

---

[12]GAO, *High-Risk Series: An Overview*, GAO-13-283 (Washington, D.C.: Feb. 14, 2013).

[13]GAO, *Investment Management: IRS Has a Strong oversight Process but Needs to Improve How It Continues Funding Ongoing Investments*, GAO-11-587 (Washington, D.C.: July 20, 2011).

[14]GAO, *Taxpayer Account Strategy: IRS Should Finish Defining Benefits and Improve Cost Estimates*, GAO-11-168 (Washington, D.C.: Mar. 24, 2011).

GAO-13-401 IRS Major IT Investments

common factors that were critical to the success of major IT investments were identified.

- We reported in January 2012 that the cost estimate used to justify the IRDM program's projected budget of $115 million for fiscal years 2012 through 2016 generally did not meet all best practices for reliability.[15] For example, the cost estimate only minimally met best practices for a well-documented estimate because IRS did not provide detailed support for staff resources, and the cost estimate documentation only justified about 6 out of the 86 requested full-time equivalent staff for IRDM. We recommended, among other things, that IRS ensure that IRDM have a reliable cost estimate. In July 2012, IRS developed a revised estimate for IRDM. We reviewed this estimate as part of this review.

In our June 2011 report on ACA, we noted that the program's cost estimate did not meet best practices identified in the GAO Cost Guide.[16] We later found that IRS planned on revising its cost estimate by September 2012.[17] ACA program officials stated that they updated the cost estimate by the September deadline and briefed us on the methodology used in the process. However, we did not receive the updated estimate because, according to IRS, it is still being approved.

---

[15]GAO, *IRS Management: Cost Estimate for New Information Reporting System Needs to be Made More Reliable*, GAO-12-59 (Washington, D.C.: Jan. 31, 2012).

[16]GAO, *Patient Protection and Affordable Care Act: IRS Should Expand Its Strategic Approach to Implementation,* GAO-11-719 (Washington, D.C.: June 29, 2011).

[17]GAO, *IRS 2013 Budget: Continuing to Improve Information on Program Costs and Results Could Aid in Resource Decision Making*, GAO-12-603 (Washington, D.C.: June 8, 2012).

Majority of IRS's Major Investments Were Reportedly Within 10 Percent of Cost and Schedule Estimates

According to IRS, 12 of 20 major IT investments were within 10 percent of cost and schedule estimates or significantly below cost between October 2011 and October 2012. For the remaining 8 investments, 3 were reported as being significantly over cost and 5 were reported as being significantly behind schedule.[18] In addition, 5 of these 8 investments had significant variances for only 1 month (for the other 3—End User Systems and Services, Main Frames and Servers Services and Support, and MeF—significant variances were reported for at least 3 consecutive months). Specifically:

- ACA,[19] Current CADE,[20] and MeF were significantly over cost; and

- Foreign Account Tax Compliance Act,[21] Integrated Customer Communication Environment, Integrated Submission and Remittance Processing System, and Main Frames and Servers Services, End User Systems and Services, and Support were significantly behind schedule.

Account Management Services, ACA, e-Services, IMF, and Integrated Customer Communication Environment also reported that they were significantly below cost during the reporting period. We are highlighting these because while being significantly below cost may indicate good performance, it can also be an indicator of poor planning or functionality being deferred and therefore warrants attention.

Figures 2 and 3 illustrate the investments with significant cost and schedule variances between October 2011 and October 2012. Appendix II provides further detail on each investment's variances between October 2011 and October 2012.

---

[18]A significant variance is defined as 10 percent over or under the planned estimate.

[19]ACA was significantly under cost for the first 3 quarters of the year because of delayed funding, then significantly over cost for the fourth quarter—after it received funding. According to IRS, although the program experienced such variances, it completed the year below the overall planned costs for the year.

[20]According to IRS, Current CADE ended on December 31, 2011.

[21]Foreign Account Tax Compliance Act refers to IRS's IT efforts to implement provisions of the Foreign Account Tax Compliance Act, Pub. L. No. 111-147, title V, subtitle A, 124 Stat. 71, 97-117 (2010).

Figure 2: Investments with Significant Cost Variances between October 2011 and October 2012
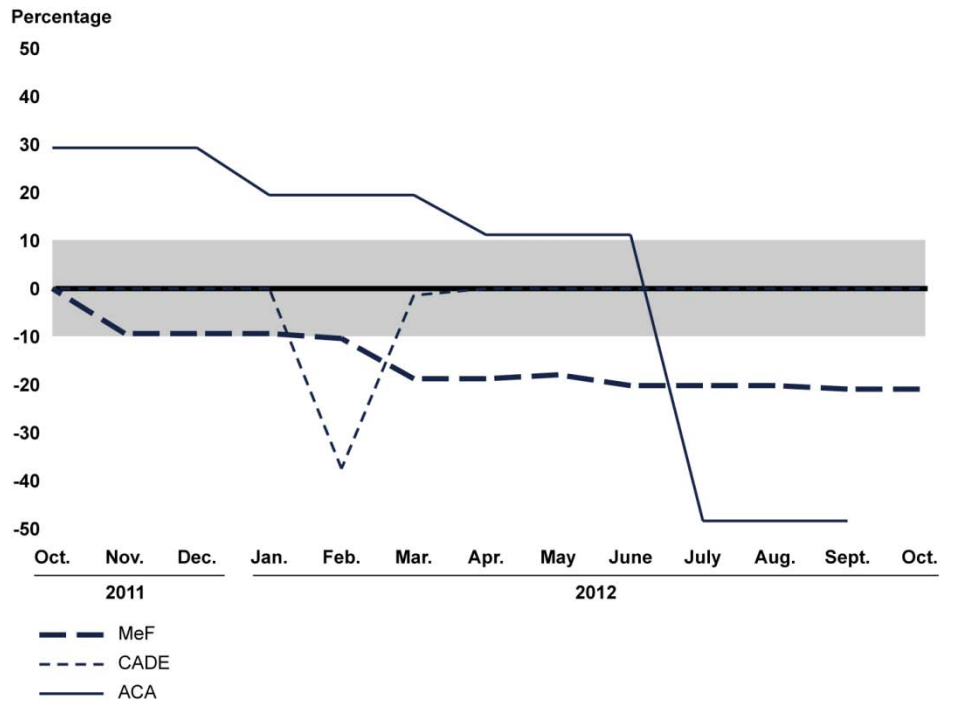
Percentage

Source: GAO analysis of IRS data.

GAO-13-401 IRS Major IT Investments

Figure 3: Investments with Significant Schedule Variances between October 2011 and October 2012

Percentage

Legend	
— —	End User Systems and Services
— ∙ —	Main Frames and Servers Services and Support
∙∙∙∙∙∙	Integrated Submission and Remittance Processing
– – –	Integrated Customer Communication Environment
——	Foreign Account Tax Compliance

Source: GAO analysis of IRS data.

IRS provided several reasons for why its investments reported such variances. For example, investments were significantly overrunning costs because of unplanned work activities, and additional costs associated with terminating an investment that was being replaced. In addition, investments that were significantly behind schedule had procurement delays, resource constraints, and the need to address unanticipated requirements. Finally, those that were significantly below costs cited funding delays, labor costs being less than planned, resources being shifted to other priorities, and work being shifted to another investment.

Reliability of Reported Cost and Schedule Variance Information Varies by Investment

The reliability of cost and schedule variances is dependent upon having (1) a reliable cost estimate and a well-constructed and controlled schedule estimate, and (2) a process for determining variances using estimates and comparing them to actual or projected amounts. The GAO cost guide[22] and schedule guide[23] identify best practices for developing cost and schedule estimates. In addition, according to the cost guide, variances can be determined by comparing estimated cost or schedule to actual or projected amounts. The cost estimates for CADE 2 and IRDM, and the schedule estimates for CADE 2, IRDM, and ACA were more favorable than for the remaining investments because they were more consistent with best practices. In addition, IRS generally determined investment cost and schedule variances for completed activities with actual amounts—although in a small number of cases it did not do so within the 60-day time frame specified in Treasury guidance. While IRS determined variances using projected information for in-process activities, the guidance for doing so did not specify how projected amounts should be determined, introducing the risk that these amounts will not consistently reflect best practices.

Cost Estimates for CADE 2 and IRDM Were More Reliable than for Other Investments

GAO's cost guide[24] identifies four characteristics of a reliable cost estimate that management can use for making informed program and budget decisions: a reliable cost estimate is comprehensive, well-documented, accurate, and credible. These four characteristics are explained in table 2.

---

[22]GAO-09-3SP.

[23]GAO-12-120G.

[24]GAO-09-3SP.

Table 2: Four Characteristics of a Reliable Cost Estimate

Characteristic	Explanation
Comprehensive	A comprehensive estimate should include both government and contractor costs of the program over its full life cycle, from inception of the program through design, development, deployment, and operation and maintenance, to retirement of the program. It should also completely define the program, reflect the current schedule, and be technically reasonable. Comprehensive cost estimates should be structured in sufficient detail (at least three levels of cost elements) to ensure that costs are neither omitted nor double-counted.[a] Specifically, the cost estimate should be based on a product-oriented work breakdown structure that allows a program to track cost and schedule by defined deliverables, such as hardware or software components. Finally, where information is limited and judgments must be made, the cost estimate should document all cost-influencing ground rules and assumptions.
Well-documented	A well-documented estimate—while taking the form of a single number—is supported by detailed documentation that describes how it was derived and how the expected funding will be spent in order to achieve a given objective. Therefore, the documentation should capture in writing such things as the source data used, the calculations performed and their results, and the estimating methodology used to derive each work breakdown structure element's cost. Moreover, this information should be captured in such a way that the data used to derive the estimate can be traced back to and verified against their sources so that the estimate can be easily replicated and updated. The documentation should also discuss the technical baseline description and how the data were normalized. Finally, the final cost estimate should be reviewed and accepted by management on the basis of confidence in the estimating process and the estimate produced by the process.
Accurate	To be accurate, an estimate should provide for results that are unbiased, and it should not be overly conservative or optimistic. An estimate is accurate when it is based on an assessment of most likely costs, adjusted properly for inflation, and contains few, if any, minor mistakes. In addition, the estimate should be grounded in a historical record of cost estimating and actual experiences on other comparable programs. Finally, a cost estimate should be updated regularly to reflect material changes in the program, such as when schedules or other assumptions change, and actual costs, so that it is always reflecting current status.
Credible	To be credible, an estimate should discuss any limitations of the analysis because of uncertainty or biases surrounding data or assumptions. Major assumptions should be varied, and other outcomes recomputed to determine how sensitive they are to changes in the assumptions (i.e., sensitivity analysis). Risk and uncertainty analysis should be performed to determine the level of risk associated with the estimate. For management to make good decisions, the program estimate must reflect the degree of uncertainty, so that a level of confidence can be given about the estimate. Having a range of costs around a point estimate is more useful to decision makers because it conveys the level of confidence in achieving the most likely cost and also informs them on cost, schedule, and technical risks.[b] Further, the estimate's results should be cross-checked, and an independent cost estimate conducted by a group outside the acquiring organization should be developed to determine whether other estimating methods produce similar results.

Source: GAO-09-3SP.

[a]The appropriate number of levels for a work breakdown structure varies from program to program and depends on a program's complexity and risk. However, each work breakdown structure should, at the very least, include three levels. The first level represents the program as a whole and therefore contains only one element—the program's name. The second level contains the major program segments, and level three contains the lower-level components or subsystems for each segment.

[b]A point estimate is the most likely value for the cost estimate, given the underlying data. The level of confidence for the point estimate is the probability that the point estimate will actually be met. For example, if the confidence level for a point estimate is 80 percent, there is an 80 percent chance that the final cost will be at or below the point estimate and a 20 percent chance that costs will exceed the point estimate.

Our analysis found that six investments at least minimally met all four characteristics of a reliable cost estimate. The estimated costs for CADE 2 and IRDM were more reliable than those of the remaining five investments because they substantially met at least three of the four characteristics of a reliable cost estimate. Specifically, CADE 2 substantially met the practices for a comprehensive, well-documented, and credible cost estimate and partially met the practices for an accurate estimate; and IRDM substantially met the practices for a comprehensive, well-documented, and accurate cost estimate and partially met the practices for a credible estimate. Table 3 shows our assessment of each investment's cost estimate against the practices associated with the characteristics for a reliable estimate. (Summary-level and detailed assessments of the cost estimates can be found in the project profiles in app. III.)

Table 3: Status of Whether Each Key Investment Met the Four Characteristics of a Cost Estimate

	Comprehensive	Well-documented	Accurate	Credible
ACA[a]	◐	◐	◐	◔
CADE 2	◕	◕	◐	◕
e-Services	◐	◔	◐	n/a[b]
IRDM	◕	◕	◕	◐
IRS.gov	◕	◐	◐	◔
MeF	◔	○	◔	◔
RRP	◕	◐	◕	◔

Key: ●=Fully met—the program provided evidence that it fully implemented the cost-estimating practices for this characteristic.

◕=Substantially met—the program provided evidence that it implemented a large portion of the cost-estimating practices for this characteristic.

◐=Partially met—the program provided evidence that it implemented about half of the cost-estimating practices for this characteristic.

◔=Minimally met—the program provided evidence that it implemented a small portion of the cost-estimating practices for this characteristic.

○=Not met—the program did not provide evidence that it implemented the practices or provided evidence that it only minimally implemented

Source: GAO analysis of IRS documentation.

[a]We reported on ACA's initial cost estimate in June 2011 and made recommendations for IRS to address weaknesses we had identified. IRS updated its estimate in September 2012 but we have not yet received it because, according to IRS, it is still being approved. The assessment here reflects the ratings we reported in June 2012.

[b]Because the cost estimate for e-Services is only for maintenance costs, we have determined that the credible criterion is not appropriate for this assessment.

Officials attributed the greater reliability of the cost estimates for IRDM to the fact that they are newer estimates and were therefore prepared using practices that have recently been instituted through improvements in software development practices.[25] They also noted improvements in cost estimating practices that were made in response to recommendations made during GAO reviews of earlier estimates. However, the weaknesses in remaining investments' cost estimates put IRS at risk of having an unreliable baseline from which to measure variances.

Schedule Estimates for ACA, CADE 2, and IRDM Were Better Constructed and Well Controlled than for Other Investments

According to the GAO schedule estimating guide,[26] a schedule is reliable when it is (1) comprehensive, (2) well-constructed, (3) controlled, and (4) credible. The guide identifies best practices that must be implemented for each of these characteristics to be achieved. For purposes of our review, we determined that the well-constructed and controlled characteristics had the minimum set of practices needed to establish a good baseline from which to measure progress. A schedule is well-constructed when all its activities are logically sequenced with the most straightforward logic possible, unusual or complicated logic techniques are used judiciously and justified in the schedule documentation, and the schedule's critical path represents a true model of the activities that drive the project's earliest completion date and total float accurately depicts schedule flexibility.[27] In addition, a schedule is controlled if it is updated periodically by trained schedulers using actual progress and logic to realistically forecast dates for program activities, and regularly compared against a

---

[25]IRS's Applications Development organization received Carnegie Mellon University Software Engineering Institute's Capability Maturity Model Integration (CMMI) level 3 certification in September 2012. IRS has been improving its software development practices using CMMI. CMMI calls for disciplined software development and acquisition practices which are considered industry best practices. CMMI maturity level 3, also known as the "defined" level is a high achievement by industry standards. At this level, processes are well characterized and understood, and are described in standards, procedures, tools, and methods. The organization's set of standard processes, which is the basis for maturity level 3, is established and improved over time. A defined process clearly states the purpose, inputs, entry criteria, activities, roles, measures, verification steps, outputs, and exit criteria. In addition, processes are managed more proactively using an understanding of the interrelationships of process activities and detailed measures of the process, its work products, and its services.

[26]GAO-12-120G.

[27]Total float, the amount of time an activity can be delayed or extended before delay affects the program's finish date, can be positive, negative, or zero.

designated baseline schedule to measure, monitor, and report the project's progress.

Overall, the seven investments at least minimally met the criteria for a well-constructed and controlled schedule estimate. Specifically, the schedule estimates for ACA and IRDM substantially met one of the two characteristics and partially met the other, while the schedule estimate for CADE 2 partially met the practices for a well-constructed and controlled schedule, and the remaining four investments fared lower. Table 4 below shows our assessment of each investment's schedule against the practices associated with the characteristics for a well-constructed and controlled schedule. (Summary-level and detailed assessments of the schedule estimates can be found in the project profiles in app. III.)

Table 4: Status of Whether Key Investments' Schedules Were Well-constructed and Controlled

	Well-constructed	Controlled
ACA	◑	◕
CADE 2	◑	◑
e-Services	◑	◔
IRDM	◕	◑
IRS.gov	◔	◑
MeF	◔	◑
RRP	◔	◑

Key: ●=Fully met—IRS provided complete evidence that satisfies the entire criterion.

◕=Substantially met—IRS provided evidence that satisfies a large portion of the criterion.

◑=Partially met—IRS provided evidence that satisfies about half of the criterion.

◔=Minimally met—IRS provided evidence that satisfies a small portion of the criterion.

○=Not met—IRS provided no evidence that satisfies any of the criterion.

Source: GAO analysis of IRS documentation.

Similar to the cost estimates, the higher schedule estimate assessments for IRDM were due in part to the fact that they are newer estimates and were therefore prepared using practices that have recently been instituted through improvements in IRS's software development practices. Improving the schedule estimates for the remaining investments so they are better constructed and controlled will improve the reliability of reported variances.

Most Variance Calculations Using Actual Information Were Updated With Actual Information in a Timely Manner, but Guidance for Projected Information Is Not Specific

Consistent with best practices, Treasury's guidance requires IRS to determine variances for completed activities with actual information and variances for in-process activities with projected information. Further, Treasury's guidance states that projects have 60 days to report actuals from the time an activity has been completed. In addition, projected cost and schedule can be determined in a number of ways, including evaluating critical path (for projected schedule), using earned value management data, evaluating the performance of completed work and comparing it to the remaining budget, assessing commitment values for material needed to complete remaining work, and estimating future conditions.[28]

To IRS's credit, for the six investments which report cost and schedule variances to the Appropriations Committees,[29] the majority of completed activities[30] had associated actual cost and schedule information within the 60-day reporting window, as required by Treasury. Our review of the quarterly reports submitted to the appropriations committees between July 2012 and January 2013 for performance between January 2012 and December 2012, shows that IRS updated investment cost and schedule variance information with actual amounts within the 60-day time frame required by Treasury in about 75 percent of the cases. For example, two investments, e-Services and IRS.gov, updated all activities within the required time frame. In addition, the majority of RRP, MeF, and IRDM's investments were updated on time. However, one investment—CADE 2—only updated its investments 50 percent of the time. Table 5 shows the number of activities that were updated within the 60-day time frame required by Treasury.

---

[28]GAO-09-3SP.

[29]IRS has not included ACA cost and schedule performance information in the quarterly reports it provides to the Appropriations Committees.

[30]Internally, IRS tracks cost and schedule performance at the activity level. Activities are rolled up into projects and related projects make up the investments. Since activities are at different stages, at any given time, investments can have a combination of completed and in-process activities.

Table 5: Timeliness of Reporting Status of Activities (January 2012 to December 2012)

Investment and total number of activities[a]	Number of activities completed or intended to be completed	Number of activities completed or intended to be completed that were updated within 60 days (percentage)
CADE 2 (38)	6	3 (50%)
e-Services (33)	12	12 (100%)
IRDM (58)	17	11 (65 %)
IRS.gov (35)	6	6 (100%)
MeF (17)	7	5 (71%)
RRP (16)	5	4 (80%)
Total (197)	**53**	**41 (77%)**

Source: GAO analysis of IRS documentation.

[a]Every quarter, IRS reports on certain activities for each investment. We added the activities for the four quarters we reviewed to come up with the total number of activities for each investment.

According to IRS officials, actual information is sometimes not updated within the 60-day time frame because of delays associated with approval through the governance process. However, we believe that 60 days is a reasonable amount of time to provide information on the status of an activity that was expected to be completed. Meeting the required time frame is critical given that, according to Treasury's guidance, activities are expected be planned for 6 months or less. While the number of activities which are expected to be completed is relatively low and IRS updates the variance calculations for these activities in the majority of the cases, ensuring that updated actual information is consistently reported within the required 60-day time frame will strengthen the reliability of their variances and provide information that better reflects their performance.

IRS determined variances using projected cost and schedule for in-process activities—which comprised 75 percent of all its activities—however, Treasury's guidance does not specify how projected amounts should be determined. Specifically, Treasury guidance states that

projected amounts should be determined using the project manager's best, most current assessment, but does not specify what factors project managers should consider in doing so. Officials from IRDM use several best practices when determining project amounts. For example, they consider the commitment values for material such as hardware and software needed to complete remaining work and they estimate future conditions. However, the remaining five investments did not provide documentation supporting what was included in the project manager's best assessment, and therefore it was not clear they considered best practices. According to IRS officials, IRS's improvements in software development practices and the fact that project managers are certified or on their way to being certified through the Federal Acquisition Certification for Program and Project Managers should ensure consistency in their approach and consideration of the right factors. However, having guidance which specifies best practices to consider when determining projected amounts will provide greater assurance that these amounts will indeed reflect best practices. This is important considering the large number of activities IRS reported on during our review that were determined using projected amounts.

IRS Has Identified Key Risks and Mitigation Strategies for Its Major Investments

According to best practices, the purpose of risk management is to identify potential problems before they occur.[31] When problems are identified, risk-handling activities can be planned and invoked as needed across the life of a project in order to mitigate adverse impacts on objectives. Effective risk management involves early and aggressive risk identification through the collaboration and involvement of relevant stakeholders. Risk management includes activities related, among other things, to identifying and analyzing risks and mitigating risks.[32]

IRS has a risk management process, including risk identification, risk mitigation planning, and execution procedures. Through this process, the agency identified risks and put in place risk mitigation strategies for the seven investments we reviewed.[33] The following table lists examples and

---

[31]Carnegie Mellon Software Engineering Institute, Capability Maturity Model® Integration for Acquisition (CMMI-ACQ), Version 1.3 (November 2010).

[32]Other risk management activities are associated with preparing for risk management and executive oversight.

[33]We did not determine the effectiveness of IRS's methodology for identifying and mitigating risks.

mitigation strategies for ACA as of December 2012 and the remaining investments we reviewed as of September 2012. The project profiles in appendix III include a complete list of key risks and strategies for the seven investments.

Table 6: Examples of Investments' Key Risks and Mitigation Strategies

Investment name	Risk	Mitigation strategy
ACA	Resource constraints	• Identify current project/release resources usage. • Identify current IT delivery partner resource usage. • Identify project/release and IT delivery partner resource gaps. • Develop resource plan to address the gaps. • Execute resource plan and monitor effectiveness.
CADE 2	If the data in the CADE 2 database cannot be verified as correct and capabilities to identify data errors resolved and corrected, then incorrect taxpayer information could be introduced to downstream systems	• Error identification, error resolution, and error correction. • Incremental deployment.
e-Services	Lack of a complete disaster recovery capability	• Develop a plan for rewriting current e-Services applications in Java and moving to a new platform with full disaster recovery capability.
IRDM	Implementation of a new category code as an alternate means to identify underreporters may introduce skewed results	• Limit scope of the implementation of the new category code for FY 2012 and perform extended analysis for implementation in FY 2013. • Monitor schedules closely to ensure activities stay on track.
IRS.gov	Schedule delays due to scope or funding changes	• Officials responsible for individual task orders and overall program management executives will monitor and track the schedule and funding to ensure timely completion of all milestones and proper allocation of funds. • The Executive Governance Board will see and address all issues in a timely manner.
MeF	Volume of returns processed	• Release 8 includes significant platform configuration and application architecture changes as well as extensive performance testing to ensure that it can support filing season 2013 volumes. • As a contingency, IRS will maintain part of the 1040 legacy system for filing season 2013, comprised of 25 forms and schedules representing the highest return volume.

Investment name	Risk	Mitigation strategy
RRP	RRP technology insertion risk	• Adopting several leading-edge technologies and leading end-to-end integration, engineering, and capacity performance to allow them to work together. • Applying open-source solutions to contain cost. • Including technologies such as massive parallel processing for fast data processing, best-in-class data analytics, and a business rules engine to address dynamic fraud schemes and supporting commercial off-the-shelf Products.

Source: IRS.

By identifying risks and developing associated mitigation strategies, IRS is increasing the likelihood that its projects will be delivered on time, within budget, and with the promised functionality.

Conclusions

Between October 2011 and October 2012, IRS reported the majority of its major IT investments as being within 10 percent of cost and schedule estimates or significantly below cost. In addition, the reliability of the reported cost and schedule variances for seven major investments' was mixed and we identified opportunities for IRS for improvement. In addition, for about 25 percent of all activities IRS reported on during our review, it did not report variance information for completed activities within the 60-day time frame specified in Treasury guidance. Ensuring that updated actual information is consistently reported within the required time frame will strengthen the reliability of IRS's variances and provide information that better reflects investments' performance. Further, the guidance for determining projected cost and schedule for in-process activities lacked specificity, introducing the risk that projections may not reflect best practices. This is especially important because the majority of variances were determined using projected information. Addressing these weaknesses would help to improve the reliability of the reported variance information. To its credit, IRS has identified key investment risks and put mitigation strategies in place for all seven investments in our review. This should increase the likelihood that these investments will deliver promised functionality without significant cost and schedule variances.

Recommendations for Executive Action

To improve the reliability of reported cost and schedule variance information for the seven major investments we reviewed, we recommend that the Acting Commissioner of IRS direct the Chief Technology Officer to:

- improve the reliability of cost estimates by addressing the weaknesses we identified in this report so that each investment at least substantially meets each of the characteristics of a reliable cost estimate;

- improve the extent to which schedules are well-constructed and controlled by addressing the weaknesses we identified in this report so that each investment at least substantially meets each of these characteristics;

- ensure projects consistently follow guidelines for updating performance information 60 days after completion of an activity; and

- develop and implement guidance that specifies best practices—such as including evaluating critical path (for projected schedule), using earned value management data, evaluating the performance of completed work and comparing it to the remaining budget, assessing commitment values for material needed to complete remaining work, and estimating future conditions—to consider when determining projected cost and schedule amounts.

Agency Comments and Our Evaluation

We obtained written comments on a draft of this report from IRS's Acting Commissioner, which are reprinted in appendix IV. IRS's GAO liaison also provided e-mail comments and told us we could attribute them to the Acting Commissioner. Those e-mail comments stated that IRS agreed with three of our recommendations and partially disagreed with the fourth.

In his written comments, the Acting Commissioner stated that IRS will continue its effort to improve the management of its investment portfolio, and it appreciated our recognition of its progress in strengthening its controls and capabilities. In the e-mail comments, the Acting Commissioner agreed with our recommendations to improve the cost and schedule estimates of the investments and ensure that guidelines for updating performance information 60 days after completion of an activity be consistently followed.

The Acting Commissioner partially disagreed with our fourth recommendation addressing the use of earned value management data as a best practice to determine projected cost and schedule amounts, stating that earned value management is not part of IRS's current program management processes and the cost and burden to use earned value management outweigh the value added. While we disagree with

IRS's view of earned value management since best practices have found that the value generally outweighs the cost and burden of implementing it, we provided it as one of several examples of practices that could be used to determine projected amounts. As stated in this report, other examples include evaluating critical path (for projected schedule), evaluating the performance of completed work and comparing it to the remaining budget, assessing commitment values for material needed to complete remaining work, and estimating future conditions. Developing and implementing guidance that specifies best practices to consider when determining projected cost and schedule amounts will help improve the reliability of reported cost and schedule variance information. In addition, IRS has flexibility in determining which best practices to use to calculate projected amounts. For these reasons, we believe our recommendation is still warranted.

We are sending copies of this report to interested congressional committees and the Commissioner of IRS. This report will also be available at no charge on our website at http://www.gao.gov.

If you or your staffs have any questions on matters discussed in this report, please contact me at (202) 512-9286 or pownerd@gao.gov. Contact points for our Offices of Congressional Relations and Public Affairs may be found on the last page of this report. GAO staff who made major contributions to this report are listed in appendix V.

David A. Powner
Director
Information Technology Management Issues

Timothy M. Persons, Ph.D.
Director, Center for Science, Technology, and Engineering
Applied Research and Methods

List of Committees

The Honorable Frank Lautenberg
Chairman
The Honorable Mike Johanns
Ranking Member
Subcommittee on Financial Services and
    General Government
Committee on Appropriations
United States Senate

The Honorable Ander Crenshaw
Chairman
The Honorable José E. Serrano
Ranking Member
Subcommittee on Financial Services
  and General Government
Committee on Appropriations
House of Representatives

Appendix I: Objectives, Scope, and Methodology

Our objectives were to (1) summarize the reported cost and schedule performance for all the Internal Revenue Service's (IRS) major IT investments; (2) for selected investments—IRS Affordable Care Act (ACA), Customer Account Data Engine (CADE) 2, e-Services, Information Reporting and Document Matching (IRDM), IRS.gov, Modernized e-File (MeF), and Return Review Program (RRP)[1]—determine the reliability of reported cost and schedule variances; and (3) identify IRS's reported risks and mitigation strategies for the selected investments.

To address the first objective, we obtained from IRS a list of major IT investments and associated cost and schedule variance data for October 2011 to October 2012. The variance data provided were from the Department of Treasury's Investment Knowledge Exchange tool which the department uses for internal oversight and external reporting purposes. We interviewed relevant Treasury officials, including the Director for Capital Planning and Investment Control, about their process for collecting the data and reviewed related guidance. We identified the investments with reported variances of 10 percent or more from estimates as having significant variances. For these investments, we obtained reasons for the variances from the information reported by Treasury on the Office of Management and Budget (OMB) IT Dashboard and from IRS officials.

To address the second objective, we evaluated the reliability of cost and schedule variances for the seven selected investments by (1) analyzing the cost and schedule estimates of these investments against best practices and (2) determining IRS's process for calculating cost and schedule variances. Specifically, we determined the reliability of the cost estimates by reviewing documentation IRS submitted for its cost estimates, interviewing IRS program officials and staff from Estimation Program Office who assisted in preparing the estimates, reviewing relevant sources, and comparing the information collected to the best

---

[1]The Conference Report accompanying the Consolidated Appropriations Act, 2012 directed us to review the projects for which it asked IRS to provide a quarterly report to the Appropriations Committees. These projects are IRS.gov, RRP, Enterprise Data Access Strategy/Integrated Production Model, e-Services, MeF, and CADE 2. We did not focus on Enterprise Data Access Strategy/Integrated Production Model because IRS no longer considers it a major investment. IRS was also directed to include other IT projects that are intended to implement significant legal changes and we selected ACA for review due to the investment's criticality to IRS's mission and the expected costs for the program.

practices identified in the GAO Cost Estimating and Assessment Guide[2]
to determine whether the cost estimates were comprehensive, accurate,
well-documented, and credible. We calculated the assessment rating of
each criteria within the four characteristics by assigning each individual
assessment rating: not met = 1, minimally met = 2, partially met = 3,
substantially met = 4, and met = 5.[3] We then took the average of the
individual assessment ratings for the criteria to determine the overall
rating for each of the four characteristics. The resulting average became
the overall assessment as follows: not met = 1.0 to 1.4, minimally met =
1.5 to 2.4, partially met = 2.5 to 3.4, substantially met = 3.5 to 4.4, and
met = 4.5 to 5.0. We discussed the results of our assessment with
officials from each investment's program office and with estimators from
IRS's Estimation Program Office who assisted with developing the
estimates.

We determined whether the schedule estimates were well-constructed
and controlled by reviewing documentation IRS submitted for its schedule
estimate, conducting interviews with IRS program officials and staff from
Estimation Program Office who assisted in preparing the estimates, and
reviewing relevant sources, and comparing the information against the
criteria for each of these characteristics identified in the GAO's Schedule
Assessment Guide.[4] Similar to the assessment of the cost estimate, we
determined the overall assessment rating for each characteristic by
assigning each individual rating a number: not met = 1, minimally met = 2,
partially met =3, substantially met = 4, and met = 5. Then, we took the
average of the individual assessment ratings to determine the overall
rating for each of the characteristics. The resulting average gave us the
overall assessment as follows: not met = 1.0 to 1.4, minimally met = 1.5
to 2.4, partially met = 2.5 to 3.4, substantially met = 3.5 to 4.4, and met =
4.5 to 5.0. We discussed the results of our assessment with officials from

---

[2]GAO-09-3SP.

[3]Not met – The investment provided no evidence that satisfies any of the criterion,
minimally met – the investment provided evidence that satisfies a small portion of the
criterion, partially met – the investment provided evidence that satisfies about half of the
criterion, substantially met – the investment provided evidence that satisfies a large
portion of the criterion, and met – the investment provided complete evidence that
satisfies the entire criterion.

[4]GAO-12-120G.

each investment's program office and with estimators from IRS's
Estimation Program Office who assisted with developing the estimates.

To determine IRS's process for calculating cost and schedule variances,
we interviewed IRS project officials and staff from the Strategy and
Planning office, and Treasury officials, including the Director for the
department's Capital Planning and Investment Control office (Treasury
establishes the guidance for calculating these variances and collects the
information for its oversight purposes and for reporting to OMB and other
entities). We analyzed the four quarterly reports on the performance of IT
investments submitted by IRS to the appropriations committees between
July 2012 and January 2013 to determine whether, for the six
investments reported by IRS, the agency had documented actual cost
and schedule information for completed activities within the 60-day time
frame specified in Treasury guidance. For the in-process activities, we
asked the project management staff from each investment to describe
how they had calculated the projected cost and schedule. We also
reviewed documentation, including the *Applications Development Project
Management Office Earned Value Management Tool How to Guide and
Standard Monthly Detailed Guidance for Major IT Investments,* which IRS
cited as the key documents with guidance on how to determine projected
cost and schedule. Finally, we identified best practices for determining
projected cost and schedule information from the GAO *Cost Estimating
and Assessment Guide* and compared them against IRS's practices.

To address the third objective, we identified best practices for managing
risks.[5] For each of the selected investments, we reviewed and
summarized the risks and associated mitigation strategies reported by
IRS in risk logs and the quarterly reports submitted to the appropriations
committees. We followed up with IRS officials as necessary to clarify our
understanding of the information provided.

We conducted this performance audit from July 2012 to April 2013, in
accordance with generally accepted government auditing standards.
Those standards require that we plan and perform the audit to obtain
sufficient, appropriate evidence to provide a reasonable basis for our
findings and conclusions based on our audit objectives. We believe that

---

[5]Carnegie Mellon Software Engineering Institute, Capability Maturity Model® Integration
for Acquisition (CMMI-ACQ), Version 1.3 (November 2010).

the evidence obtained provides a reasonable basis for our findings and
conclusions based on our audit objectives.

Appendix II: Investment Variances Reported by IRS from October 2011 to October 2012

The following table shows the monthly reported cost and schedule variances for IRS's major investments from October 2011 to October 2012.

Table 7: Variances for IRS' Major IT Investments from October 2011 to October 2012

		Oct-11	Nov-11	Dec-11	Jan-12	Feb-12	Mar-12	Apr-12	May-12	Jun-12	Jul-12	Aug-12	Sep-12	Oct-12
Account Management Services	Cost	0.00%	2.04%	2.97%	4.64%	4.64%	4.64%	4.64%	7.24%	7.24%	5.43%	5.43%	11.62%	11.62%
	Schedule	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Affordable Care Act (ACA) Administration	Cost			29.31%			19.47%			11.19%			-48.36%	
	Schedule			N/A			N/A			N/A			N/A	
Current CADE	Cost	0.00%	0.00%	0.00%	0.00%	-37.48%	-1.41%	N/A	N/A	N/A	N/A	N/A	N/A	N/A
	Schedule	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Customer Account Data Engine 2 (CADE 2)	Cost	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	-0.32%	-0.32%	-0.32%	-0.32%	-0.32%	-0.32%
	Schedule	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Electronic Fraud Detection System	Cost	-2.15%	-2.15%	-2.36%	-2.36%	-2.11%	-1.69%	-1.69%	-0.95%	-0.95%	-2.55%	-2.55%	-2.55%	-2.55%
	Schedule	-6.02%	-6.02%	-6.14%	-7.74%	-6.14%	-6.14%	-6.14%	-7.00%	-7.00%	-7.00%	-2.83%	-2.83%	-2.83%
e-Services	Cost	0.00%	-0.40%	-0.40%	-0.28%	-2.14%	-2.14%	-2.14%	12.66%	12.66%	21.72%	21.98%	17.31%	17.31%
	Schedule	0.00%	-9.24%	-9.24%	-5.96%	-1.64%	-1.64%	-1.64%	7.45%	7.45%	-3.61%	-3.61%	-1.91%	-1.91%
Foreign Account Tax Compliance Act	Cost	N/A	N/A	N/A	N/A	N/A	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	Schedule	N/A	N/A	N/A	N/A	N/A	0.00%	0.00%	0.00%	0.00%	-17.53%	0.00%	0.00%	0.00%
Implement Return Review Program (RRP)	Cost	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	Schedule	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Individual Master File (IMF)	Cost	0.00%	0.00%	0.00%	0.00%	0.00%	14.06%	14.06%	14.06%	14.06%	14.06%	14.06%	14.06%	14.06%
	Schedule	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Information Reporting and Document Matching (IRDM)	Cost	0.00%	0.00%	0.00%	0.00%	0.02%	0.02%	0.02%	0.02%	0.02%	0.88%	1.62%	1.62%	1.62%
	Schedule	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Integrated Customer Communication Environment	Cost	0.00%	0.00%	26.16%	26.16%	26.16%	33.47%	-0.84%	-0.84%	-7.94%	-7.94%	1.75%	6.47%	17.50%
	Schedule	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	-66.35%	-24.26%	0.00%	0.00%	0.00%

		Oct-11	Nov-11	Dec-11	Jan-12	Feb-12	Mar-12	Apr-12	May-12	Jun-12	Jul-12	Aug-12	Sep-12	Oct-12
Integrated Data Retrieval System	Cost	0.00%	0.00%	0.00%	0.54%	1.03%	1.46%	1.70%	2.20%	2.56%	2.70%	3.04%	3.11%	3.04%
	Schedule	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Integrated Financial System/CORE Financial System	Cost	0.00%	0.00%	0.00%	-0.02%	-0.02%	0.87%	0.87%	0.87%	0.87%	-2.26%	-2.26%	-3.42%	-4.47%
	Schedule	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Integrated Submission and Remittance Processing System	Cost	0.00%	0.00%	0.00%	0.00%	0.00%	1.11%	1.11%	7.41%	7.41%	7.41%	0.00%	0.00%	0.00%
	Schedule	0.00%	0.00%	0.00%	0.00%	-5.48%	-10.97%	-5.48%	-5.48%	-5.48%	-5.48%	-8.24%	-8.24%	-8.24%
IRS End User Systems and Services	Cost	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	-0.04%	0.00%
	Schedule	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	-4.51%	-63.89%	-63.89%	0.00%
IRS Main Frames and Servers Services and Support	Cost	N/A	N/A	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	Schedule	N/A	N/A	-112.05%	-246.59%	-246.59%	0.00%	-2.44%	-2.44%	-2.44%	-2.44%	-1.22%	-0.49%	-0.49%
IRS Telecommunications Systems and Support	Cost	0.00%	0.00%	0.00%	0.00%	6.76%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	Schedule	0.00%	-9.84%	-9.84%	-9.84%	-2.40%	-2.40%	-2.40%	-2.40%	-2.40%	-2.40%	-2.40%	-2.40%	-2.40%
IRS.gov	Cost	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	-1.52%	-1.52%	-1.52%	-1.52%	-1.52%	-1.52%	-1.66%
	Schedule	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	-0.07%	-0.07%	-0.07%	-0.07%	-0.07%	-0.06%	-0.06%
Modernized e-File (MeF)	Cost	0.00%	-9.35%	-9.35%	-9.35%	-10.39%	-18.77%	-18.77%	-17.95%	-20.20%	-20.20%	-20.20%	-20.93%	-20.93%
	Schedule	0.00%	0.00%	0.00%	0.00%	0.00%	-7.29%	-7.29%	-0.84%	-0.84%	-0.84%	-0.84%	-0.84%	-0.84%
Service Center Recognition/Image Processing System	Cost	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	Schedule	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%

Source: GAO analysis of IRS data.

Note: Positive variance indicates that cost or schedule is less than originally planned.

Note: Negative variance indicates that cost or schedule is greater than originally planned.

Note: Positive variances still warrant attention because they may be an indication of incomplete/deferred work.

Appendix III: Investment Profiles, Detailed Cost Assessments, Detailed Schedule Assessments, and Key Risks and Mitigation Strategies

This appendix contains the profiles, detailed cost and schedule assessments, and list of risks and mitigation strategies for the seven key investments we reviewed.

Patient Protection and Affordable Care Act (ACA)

ACA — Investment Details

Program start date:	June 2010
Full operational capability:	2017
Costs through FY 2017:	$767 million
Current life-cycle phase:	Mixed (development/ operations and maintenance)

Source: IRS data.

In March 2010, Congress enacted the Patient Protection and Affordable Care Act which impacted the work performed by IRS and resulted in the need for new systems. IRS is deploying the ACA investment in a series of releases comprised of multiple projects. Specifically, it has the following releases and projects planned:

- **Release 1.0:** Non-exchange provisions.

- **Release 2.0:** Branded prescription drugs.

- **Release 3.0:**

  - Income and family size verification: intended to calculate household income and family size based on most recent filed tax return for each household member.

  - Information sharing and reporting: intended to provide ACA data-sharing capabilities via a secure bidirectional data channel between the Department of Health and Human Services and IRS and provide reporting capabilities.

  - Individual coverage compliance: provides an online estimator for individuals to estimate their penalties.

  - Customer service.

  - Employer coverage compliance: an online estimator for employers to estimate their assessable payment.

  - Coverage data repository: a retrieval and storage system for ACA-related data.

  - Premium tax credit: intended to be a computation engine hosted at IRS as a web service that the Department of Health and Human Services will consume.

- **Release 4.0:** Coverage data repository, information sharing and reporting, and customer service.

- **Release 4.1:** ACA information returns and insurance provider fee.

  Release 5.0: ACA verification service, information returns, coverage data repository, and information sharing and reporting.

- **Release 6.0:** Coverage data repository.

- **Release 6.1:** ACA compliance validation, releases deployed and in production.

Table 8: Assessment of the ACA Program's Cost Estimate

Characteristic	Assessment	Key examples of rationale for assessment
Comprehensive	Partially met	ACA's estimate was partially comprehensive because it documented ground rules and assumptions, but not the entire life cycle. Further, details of the work necessary did not capture all costs.
Well-documented	Partially met	The estimate was partially well-documented because methodologies behind the calculations were described; however, several sources of data were unclear or relied on IRS employee input without data to back up expert claims. Historical data were also not normalized to ensure consistency of cost data.
Accurate	Partially met	ACA's estimate was partially accurate because it was unbiased and based on most likely and historical costs, but it had not been recently updated (it was last updated in October 2010). In addition, analyses of variance between actual and projected costs were not documented.
Credible	Minimally met	The estimate was minimally credible because it only included some information for conducting a high-level risk and sensitivity analyses, which would help identify variables most likely to affect the estimate.

Source: GAO analysis of ACA's October 2010 cost estimate.

Table 9: Assessment of the ACA Program's Schedule Estimate

Characteristic	Assessment	Key examples of rationale for assessment
Well-constructed	Partially met	ACA's schedule was partially well-constructed because there were few missing dependencies and few activities with dangling logic. In addition, the schedule contained no start-to-finish logic and only one Summary activity with logic. However, there were several activities with missing or incorrect logic. In addition, there were convergence issues with the schedule. For example, 25 activities within the master schedule had more than 10 predecessors, including 1 activity with 65. Activities with a large number of predecessors can represent an unrealistic plan because they imply the need to accomplish a large amount of work on time before an activity or event can occur as planned. We also found a significant number of lags in the schedule, as well as a number of unjustified soft constraints. In addition, activities that appear as critical that may not actually be impacting the start date of the key Go Live milestone. Finally, the schedule appeared to contain an unreasonable amount of total float—the amount of time by which a predecessor activity can slip before the delay affects the program's estimated finish date. For example, according to the schedule 15 percent of all remaining activities had over 1,000 days of total float, with some activities implying an ability to slip more than 7 working years before impacting the project finish date. Incorrect float estimates may result in an invalid critical path and an inaccurate assessment of project completion dates.

Characteristic	Assessment	Key examples of rationale for assessment
Controlled	Substantially met	The schedule substantially met the controlled characteristic because the program management office regularly updated, reviewed, and assessed the health of the schedule, and tracked trends of key metrics. However, there were a few instances of date anomalies, such as inconsistent status dates in subprojects, start and finish dates in the past with no actual dates, and actual start dates in the future. Program officials stated that a baseline schedule is the basis for measuring performance, and we found baseline dates were being monitored by the program. However, we found only a portion of activities had corresponding baseline dates. In addition, we found date variances ranging from -305 days (almost 14 months ahead of schedule) to 203 days (or 9 months behind schedule). In addition, no schedule baseline documentation was provided, which provides a single document that defines, among other things, the organization and logic of the integrated master schedule.

Source: GAO analysis of ACA's May 2012 integrated master schedule.

Table 10: List of Key Risks and Mitigation Activities for ACA as of December 2012

Risk	Mitigation
ACA transaction portal availability	• Manage implementation plan to meet various milestones. • Develop contingency plan.
Incident management	• Form an incident management sprint to collaboratively develop an incident management strategy and plan.
Resource constraints	• Identify current project/release resources usage. • Identify current IT delivery partner resource usage. • Identify project/release and IT delivery partner resource gaps. • Develop resource plan to address the gaps. • Execute resource plan and monitor effectiveness.
Environment management	• Conduct environment management workshops to review ACA environments, document environment configurations, and align IRS stakeholders on environment nomenclature.
Security controls and audit requirements	• Provide detailed requirements to contractor for validation of the security requirements. • Cyber security will work with the enterprise architecture office and the projects to ensure they have clear understanding of the requirements. • A cyber security contractor will provide the 'build to' details. • The enterprise architecture office will finish up the design efforts.
Consistency checks for periodic data	• Work with the business to solidify the list of consistency checks and manage the implementation of the checks.
Legacy impacts to release 5	• Work with the business on forms design discussions to understand future requirements and impact on legacy systems. • Develop a schedule for requirements gathering and architecture/roadmap development for release 5. • Gather requirements from the business and build out the solution architecture and program roadmap/guidebook for future releases. • Allocate requirements and architecture to projects. • Further refine the level of detail on the legacy system impact for 2013 and beyond, and determine complexity of system impacts and lead time needed. • Develop implementation timeline. • Communicate requirements, architecture, and legacy impacts to all affected stakeholders.

Source: IRS.

Customer Account Data Engine 2 (CADE 2)

CADE2 — Investment Details

Program start date:	January 2010
Full operational capability:	2018
Costs for FY11 to FY16:	$1,147.98 billion
Current life-cycle phase:	Mixed (development/ operations and maintenance)

Source: IRS data.

CADE 2 began in 2010 as a new strategy for accelerating completion of a modernized database and converting to a single processing system sooner than CADE (which was intended to provide a modernized system of taxpayer accounts, with the ultimate goal of eventually replacing IMF) would allow. CADE 2 is expected to deliver its functionality incrementally through three phases known as transition states. Transition state 1, for which a key segment was implemented in January 2012, includes (1) daily batch processing of individual taxpayer returns provided by modifying the IMF to run on a daily, rather than weekly, basis; and (2) a comprehensive database for housing all individual taxpayer accounts and loaded with data from CADE and IMF to provide timelier updates of taxpayer information for use by IRS employees for compliance and customer service. Transition state 2, which IRS has begun planning, is to include (1) target technology developed and deployed (single processing system; IMF retired); (2) high-priority downstream service and compliance applications modified to take advantage of the new database; and (3) fixes to some key financial material weaknesses.[1]

According to IRS, the database component of transition state 1 has been in production as of May 2012, and in August 2012, daily updates to the database began. In addition, IRS has scheduled downstream feeds from the database to relevant systems. Before putting the CADE 2 database into full production, IRS is completing the remaining database implementation activities, including synchronization with the latest taxpayer account data.

Table 11: Assessment of the CADE 2 Program's Cost Estimate

Characteristic	Assessment	Key examples of rationale for assessment
Comprehensive	Substantially met	CADE 2's cost estimate was substantially comprehensive because it included all anticipated project costs, including both government and contractor costs, and recurring operations and maintenance costs over the life cycle of those systems implemented as part of transition 1. The estimate also provided detailed costs and total cost of ownership for each of six different major elements of the work breakdown structure, and that costs were developed according to it. However, we were unable to match the work breakdown structure elements with elements in the updated milestone 3 schedule, which increases the potential for costs to be overlooked.

[1]CADE 2 will not completely resolve IRS's existing financial management deficiencies related to unpaid tax assessments. This is because CADE 2 is designed only to replace IRS's IMF and not the Business Master File and resolving the financial management deficiencies would require addressing issues related to both master files.

Characteristic	Assessment	Key examples of rationale for assessment
Well-documented	Substantially met	CADE 2's cost estimate was substantially well-documented because it contained detailed information on the source data used to develop the estimates for each work breakdown structure element. It also included primary sources, such as delivery partner agreements and existing contracts for labor rates, as well as actual cost data expended to date. Where analogous data were used, it was normalized through the application of factors to tailor them to the specific project being estimated. Detailed descriptions of the knowledge base settings used with cost models were provided. However, there was no discussion of the reliability of the data or how actual costs were normalized to estimate future costs. Since the bulk of the estimation relied upon a proprietary model, only those individuals or organizations possessing the license will be able to access the model.
Accurate	Partially met	CADE 2's cost estimate was partially accurate because it was evaluated to determine where it fell against a range of possible costs, and most costs were estimated between the 55 percent to 65 percent confidence level unless there was reason to believe that more conservative estimates were appropriate because of the complexities associated with the tasks involved. However, there was no overall confidence level associated with the total cost estimate. In addition, there was no discussion in the documentation about the reliability of the data and there was no earned value management system that is generating actual costs for CADE 2.
Credible	Substantially met	The cost estimate for CADE 2 was substantially credible because a sensitivity analysis was included, but it was not performed on the infrastructure portion due to the nature of how architecture is constructed. Further, key cost drivers were identified and their parameters examined. In addition, simulation models were used to develop distributions of total possible costs under a range of assumptions and estimates were compared to the present cost update as were the estimates from the Delivery Partners. Where differences existed between the estimates, they were explained. According to IRS, it has an independent cost estimate because an independent organization developed it; however, a second estimate is to be developed independently so that it may be used to confirm the results developed in the IRS estimate. Although the group that developed the estimate is independent of the acquiring organization, there is still only one estimate that has been developed.

Source: GAO analysis of CADE 2's February 2011 transition state 1 milestone 3 cost estimate.

Table 12: Assessment of the CADE 2 Program's Schedule Estimate

Characteristic	Assessment	Key examples of rationale for assessment
Well-constructed	Partially met	CADE 2's schedule was partially well-constructed because it contained no activities with missing logic, a minimal number of date constraints, no start-to-finish logic, and no logic on Summary activities. However, the schedule contained a relatively high number of activities with dangling logic. That is, while these activities had predecessors and successors, either the predecessors were not affecting the start date of the activity, or the activity's finish date had no impact on the start date of successor tasks. In addition, while each activity had a predecessor and successor, the schedule's ability to calculate reasonable values of total float and, therefore, a valid critical path, was severely hampered by the high number of predecessors—158—tied into the exit milestone. As a consequence of the unreasonable amount of predecessor paths merging into the exit milestone, as well as the high number of activities with dangling logic and long-duration critical tasks, both the critical path and the driving path were unnecessarily complex, and resulted in an unrealistic plan. Total float values in the schedule appeared reasonable for the most part, but of the 93 remaining predecessor activities to the exit milestone, over half of them could slip between 29 to 94 percent of the remaining 35 days of the project.

Characteristic	Assessment	Key examples of rationale for assessment
Controlled	Partially met	CADE 2's schedule was partially controlled because there were no instances of date anomalies, such as actual start or finish dates in the future; a relatively small number of out-of-sequence activities; and many baseline dates stored in the schedule. However, nearly 8 percent of activities in the schedule were missing associated baseline dates. While officials stated trend analysis is performed on a weekly basis and includes a critical and near-critical path analysis, a schedule change log, and a 4-week look-ahead assessment, we had no documentation to support this assertion. In addition, we found some baseline date inconsistencies. Specifically, some activities had actual start dates that differed from their baseline start dates, even though they occurred prior to the date the baseline dates were set in the file.

Source: GAO analysis of CADE 2's August 2012 database implementation integrated master schedule.

Table 13: List of Key Risks and Mitigation Activities for CADE 2 as of September 2012

Risk	Mitigation
Incorrect taxpayer information could be introduced to downstream systems	Error identification, error resolution, error correction, and incremental deployment.

Source: IRS.

GAO-13-401 IRS Major IT Investments

e-Services

e-Services — Investment Details

Program start date:	August 2011
Full operational capability:	Time frame not available
Costs for FY11 to FY16:	$82.53 million
Current life-cycle phase:	Mixed (development/ operations and maintenance)

Source: IRS data.

e-Services is a suite of web-based products that are intended to allow tax professionals and payers to conduct business with IRS electronically. These services are only available to approved IRS business partners and not available to the general public. The program is available via the Internet 24 hours a day, 7 days a week, and it contains products such as registration, an e-file application, and taxpayer identification number matching (a pre-filing service which allows authorized payers to match up to 25 payee taxpayer identification numbers and name combinations against IRS records prior to submitting an information return). In addition, tax professionals who are active participants in the IRS e-file program and e-file five or more accepted individual or business returns in a season are eligible to use the following incentive products: disclosure authorization, electronic account resolution, and transcript delivery system.

Table 14: Assessment of the e-Services Program's Cost Estimate

Characteristic	Assessment	Key examples of rationale for assessment
Comprehensive	Partially met	e-Services's cost estimate was partially comprehensive because it reviewed projects within the program, but did not include costs. While it was possible to link costs of specific projects to the level of effort, not all projects were documented. The projects included in the work breakdown structure generally did not match those covered by the estimate. Where projects matched, the estimated level of effort was different.
Well-documented	Minimally met	e-Services's cost estimate was minimally well-documented because it provided project impact assessments and a cost worksheet to support the cost estimate, but the cost worksheet lacked detailed documentation that described how costs were derived. Within the project impact assessments and, estimates of level-of-effort to complete tasks did not reference the factors used. The project impact assessments also did not discuss how the data were normalized. Where changes in projects have occurred, updates to the project impact assessments and show a change in the estimated hours of work required for completion, but without an explanation of what drives these changes.
Accurate	Partially met	e-Services's cost estimate was partially accurate. While roll-ups of the estimate appeared to be correct, it was not possible to determine whether the cost estimate was overly conservative or optimistic because no uncertainty analysis was performed. Further, adjustment for inflation was not documented and the estimate was missing information to verify whether or not it contained mistakes. In addition, while the estimate had been updated to reflect significant changes, it did not explain how changes drive the estimate. The estimate documentation asserted that cost information from previous or comparable programs was used, but there was no documentation to verify this claim.
Credible	Not applicable	

Source: GAO analysis of e-Services's October 2011 cost estimate.

Table 15: Assessment of the e-Services Program's Schedule Estimate

Characteristic	Assessment	Key examples of rationale for assessment
Well-constructed	Partially met	The e-Services schedule was partially well-constructed because there were a relatively low number of Summary activities with logic links, no activities with start-to-finish links, and no potential issues with convergence (that is, a high number of predecessor activities). However, approximately one out of five remaining activities was missing logic. Without the correct linkages, activities that slip early in the schedule do not transmit delays to activities that should depend on them. There were also a relatively high number of date constraints: there were 58 remaining activities (31 percent) that had start-no-earlier- than constraints, which prevent activities from starting earlier than a defined date, yet allow them to slip if predecessors are delayed. The critical path was a straightforward waterfall sequence of activities, yet it began with a start-no-earlier-than constraint and was thus not continuous throughout the schedule. As a consequence, there was no near-term work directly impacting the overall duration of the project. Finally, there were large values of total float in the schedule, which was most likely not indicative of the true flexibility within the schedule. Because the schedule was missing logic, float estimates were not accurate and may result in an invalid critical path and an inaccurate assessment of project completion dates.
Controlled	Minimally met	e-Services's schedule was minimally controlled. The schedule did not have a valid status date because it was over 13 months in the past. Unless a valid status date is provided, the schedule cannot be used to reliably convey past and remaining effort. Using the date embedded in the file name as the current status date, there were no out-of-sequence activities. However, we found 18 activities with start dates in the past, but no actual start date; 21 activities with finish dates in the past, but no actual finish date; 4 activities with actual finish dates in the future; and 1 activity with an actual start date in the future. If the schedule has not been updated, then it is impossible to tell what activities have been completed, are in progress, are late, and are planned to start on time. While there were two sets of baseline dates in the schedule file, it was not clear how valid these data were. For example, there was an inconsistent number of baseline start dates, and for those baseline start dates that were stored in the schedule, there were start variances that varied greatly. For example, some were as low as -322 working days and some as high as 312 working days. Without a formally established baseline schedule to measure performance against, management cannot identify or mitigate the effect of unfavorable performance.

Source: GAO analysis of e-Services's October 2012 integrated master schedule.

Table 16: List of Key Risks and Mitigation Activities for e-Services as of September 2012

Risk	Mitigation
Lack of a complete disaster recovery capability	Develop a plan for re-writing current e-Services applications in Java and moving to a new platform with full disaster recovery capability.

Source: IRS.

Information Reporting and Document Matching (IRDM)

IRDM — Investment Details

Program start date:	August 2010
Full operational capability:	Time frame not available
Costs for FY11 to FY16:	$110.76 million
Current life-cycle phase:	Mixed (development/ operations and maintenance)

Source: IRS data.

IRDM is aimed at helping close the tax gap—the difference between what taxpayers should have paid and actually did. It is intended to improve voluntary compliance and accurate reporting of income by establishing a new business information matching program initially focused on merchant card payments, securities basis reporting, and withholding on government payments. IRDM is intended to support a methodology embedded in a new process that pre-sorts, matches, identifies, manages, and reports on returns that are likely sources of gap-reducing taxes that are missed by the current system. To close the tax gap and achieve the performance goals, IRS requires operational resources and systems to be put in place to implement the changes across the enterprise that are intended to expand and improve its automated matching of data on information returns to the data submitted on tax returns filed. IRS is implementing IRDM primarily through IRS's Information Technology and Business Operating Division. Information Technology intends to build IT systems to deliver functionality through four software projects.

Table 17: Assessment of the IRDM Program's Cost Estimate

Characteristic	Assessment	Key examples of rationale for assessment
Comprehensive	Substantially met	IRDM's cost estimate was substantially comprehensive because it contained nonrecurring and recurring costs from 2012 through 2016, as well as costs for both groups responsible for the investment. However, only 2 years of operations and maintenance costs were shown because the Estimation Program Office's practice is to show 5 years of costs at a time. Further, the estimate reflected the current schedule at a high level and the technical baseline was captured in a document dated July 16, 2012. The document contained sufficient detail to assist the cost estimators and was developed by a cross-functional group of experts including engineers. In addition, there was a work breakdown structure that contained each of the four IRDM projects and each was product-oriented in that they broke down the end products and outlined major work. They also contained more than three levels of indenture and the sum of the children equaled the parent, as shown in the investment summary report spreadsheet. However, the earned value management data were not broken down at the same level as the work breakdown structure because, according to IRS officials, IRS's financial system data are used to populate the earned value management data. As a result, IRS tracks projects using internal order codes that correspond only to projects. Therefore, IRDM's earned value management data were only available at a high level. Finally, ground rules and assumptions were contained within the basis of estimate document. For example, there were global assumptions, software application development assumptions, project infrastructure assumptions, business operating division assumptions, deployment assumptions, and program management assumptions.
Well-documented	Substantially met	The estimate was substantially well-documented. While the documentation provided a thorough description of the calculations used, only some of the source data were provided. The technical baseline the cost estimate was based on was consistent with the technical information provided. A briefing to management was conducted but did not include all the information necessary to fully explain how the estimate was derived.

Characteristic	Assessment	Key examples of rationale for assessment
Accurate	Substantially met	The estimate was substantially accurate because IRS reported that historical information was used to inform the cost estimate. For example, actual costs and data from release 1 were used to cross check the costs of release 2, while the estimate was created using a commercial software package where historical costs are embedded as part of the underlying computations and we were therefore unable to verify what specific historical data were used to support the estimate. IRS stated the cost estimate was unbiased and based on most likely costs; however, the documentation stated that the estimate was at an 80 percent confidence level, which is outside the best practice 55 percent to 65 percent range. Further there were no calculations mistakes seen; however, there was no visibility into the model calculations. Finally, inflation was used properly.
Credible	Partially met	The estimate was partially credible because IRDM did not have an independent estimate completed. However, IRDM performed cross checks on the major cost elements. For the risk analysis, IRS stated that the cost estimate was at the 80 percent confidence level by definition of being an early (i.e., Vision and Strategy) estimate; however, a simulation analysis was not conducted so there was no quantitative analysis to back up that claim. Finally, documentation showing the output of a sensitivity analysis was provided, but no evidence of the analysis itself was provided.

Source: GAO analysis of IRDM's July 2012 release 2 cost estimate.

Table 18: Assessment of the IRDM Program's Schedule Estimate

Characteristic	Assessment	Key examples of rationale for assessment
Well-constructed	Substantially met	IRDM's schedule was substantially well-constructed because it was only missing a few logic links and there were relatively few instances of hard date constraints, lags, and leads. However, we found that there were three activities with a large number of predecessors (between 11 and 51 predecessors) which, as noted previously, may affect the ability of the activities to finish on time. While the program had a continuous critical path that followed a typical waterfall sequence, there were relatively few remaining activities that actually appeared critical. Of these activities marked critical by the scheduling software, two were recurring support type activities that should not be driving the finish date. The critical path also had one constraint and six activities with lags and leads. Further, total float was not reasonable. For example, the schedule had total float as high as 289 days (meaning an activity could slip as much as a year without affecting the final finish date). High total float values are usually caused by missing logic, a condition that we found in the schedule. When total float is not reliable, management cannot trust the forecasted dates and cannot use the critical path to proactively manage the program. However, until the issues regarding missing logic and unreasonable total float are entirely resolved, we cannot confirm whether the critical path is valid.

Characteristic	Assessment	Key examples of rationale for assessment
Controlled	Partially met	The schedule partially met best practices for being controlled. For instance, though program officials said the schedule had been recently updated there was no status date recorded in the schedule. Unless a status date is provided, the schedule cannot be used to reliably convey past and remaining effort. Program officials provided evidence that showed they held weekly schedule status meetings and that they were actively managing the program using information from the schedule. Our analysis found some date anomalies, such as activities that were scheduled to start or complete in the past that had not been rescheduled to resume after the status date. When we altered the schedule to bring forward uncompleted work in the past, the finish date for the schedule slipped almost a month. While it was clear that the program had set several baselines from which to measure progress against, there was no evidence that the program had a schedule baseline document which would describe the schedule's network logic, the program's approach to managing resources, how to use the schedule file, what acronyms and custom fields were used, and how risks were assessed and accounted for in the schedule.

Source: GAO analysis of IRDM's May 2012 case inventory selection and analytics integrated master schedule.

Table 19: List of Key Risks and Mitigation Activities for IRDM as of September 2012

Risk	Mitigation
Implementation of a new category code as an alternate means to identify underreporters may introduce skewed results	Limit scope of the implementation of the code for fiscal year 2012 and perform extended analysis for implementation for fiscal year 2013. Monitor schedules closely to ensure activities stay on track.
Selection of new category code alternate case types may result in working identical cases using traditional Automated Underreporter processing	Implement a means to supply functionality that deconflicts the ability to select identical underreporter cases by providing a manual workaround. The workaround will provide a manual list of underreporter cases worked by an alternate system to tax analysts to prevent working identical cases.

Source: IRS.

IRS.gov

IRS.gov — Investment Details

Program start date:	July 2009
Full operational capability:	2020
Total life-cycle costs:	$1.128 billion
Current life-cycle phase:	Mixed (development/ operations and maintenance)

Source: IRS data.

IRS.gov is a new web environment intended to replace the current web environment. It consists of public user portal—IRS.gov, a registered user portal, and an employee user portal. The key goals of the program include simplifying and transforming the user web experience, consolidating and advancing IRS web technology to industry standards, implementing a high-performing contract structure and terms, and marketing competitive costs throughout the program's life cycle.

Additionally, this effort supports the framework and functionality of IRS portals services used by the American taxpayer, IRS employees, and registered users. The overall objectives of this program are to:

- provide a high-performing partnership between IRS and its contractor;

- provide industry-leading web practices and innovations;

- provide a compelling program performance management framework;

- create a new IRS web environment that is the trusted taxation website;

- create a simple and manageable IRS web environment;

- provide a single point, end-to-end operational accountability and visibility;

- provide a cost effective and affordable program cost structure; and

- transition successfully from the old programs to the new program.

Table 20: Assessment of the IRS.gov Program's Cost Estimate

Characteristic	Assessment	Key examples of rationale for assessment
Comprehensive	Substantially met	IRS.gov's cost estimate was substantially comprehensive. The cost estimate provided by the program office included annual roll-ups and functional roll-ups. It also included a detailed cost work breakdown structure. The solicitation issued by the program office included a technical baseline with deliverables, exit criteria, and performance requirements to assess the technical reasonableness of the program. However, some cost elements in the work breakdown structure applied to more than one portal, so it was not clear whether or not some cost elements were either double-counted or omitted. The work breakdown structure relied on historical data to estimate the cost associated with most of the cost elements.

Characteristic	Assessment	Key examples of rationale for assessment
Well-documented	Partially met	IRS.gov's cost estimate was partially well-documented because it lacked detailed documentation that describes how costs were derived. While it cited historical data and referred to the previous version of the portal, it did not explain what, if anything may have changed between versions of the portal that may influence costs. IRS.gov provided evidence that the estimate was reviewed and approved by management.
Accurate	Partially met	The cost estimate for IRS.gov was partially accurate because it was not possible to determine whether the cost estimate was overly conservative or optimistic because a risk confidence level was not calculated. Adjustment for inflation in the portal budget forecast should be explained. The estimate has not been updated to reflect any significant changes (if any occurred).
Credible	Minimally met	The cost estimate did not discuss any limitations of the analysis and did not carry out any sensitivity analysis. Further, cost drivers were not identified. Although IRS reports having an independent cost estimate, it was a high-level roll-up and did not include details to show how it was calculated.

Source: GAO analysis of IRS.gov's cost estimate.

Table 21: Assessment of the IRS.gov Program's Schedule Estimate

Characteristic	Assessment	Key examples of rationale for assessment
Well-constructed	Minimally met	The schedule for IRS.gov was minimally well-constructed because it was missing some (14 percent) predecessor and successor logic, had too many constraints (40 percent), contained a few lags and leads (11 percent), and had several activities (13 percent) with too many predecessors, all of which call into question the ability of the activities to finish on time. Because of these issues, the critical path was not valid. For instance, a third of the 12 activities marked critical were level of effort activities. In addition, there were constraints, lags, and leads on the critical path as well as an activity missing a predecessor. Moreover, when we traced the driving path for the "IRS.gov project complete milestone"—that is, the sequence of activities that determine its planned start date—we found only five activities that were actually affecting the milestone end date. Of these five activities, two had start-no-earlier-than constraints. Finally, the entire duration of the driving path was only 7 working days in duration and spanned the period from October 2012 to November 2012 and did not include any effort prior to October 25, 2012. Until the schedule can produce a true critical path, the program office will not be able to provide reliable time line estimates or identify when problems or changes may occur and their effect on downstream work. In addition, we found that total float was not reasonable. In particular, 61 percent of the remaining activities had total float greater than 44 days with one activity reflecting total float of 226 days, which means it could slip more than 10 months without affecting the project's finish date. High total float values are usually caused by missing logic or incorrect logic that results in high convergence of predecessors. Both of these issues were present in this schedule. When total float is not reliable, management cannot trust the forecasted dates and cannot use the critical path to proactively manage the program.

Characteristic	Assessment	Key examples of rationale for assessment
Controlled	Partially met	IRS.gov's schedule was partially controlled because it had been recently updated and a baseline had been set from which to measure progress against. In addition, program officials held monthly schedule status meetings and provided evidence that they were actively managing the program using information from the schedule. The schedule also did not have any instances of actual start or finish dates recorded beyond the status date of July 13, 2012, which aligns with best practices. There was also a schedule guide that stated baseline start and finish dates should be included in the schedule and defined several custom flag fields for tracking work products, cross task order dependencies, and exit criteria. Moreover, IRS.gov provided evidence of a change control process as well as documentation of a change control log to show that it was using it to track changes to the schedule. However, there were some date anomalies, such as a few activities that were scheduled to start more than a year ago that had not been rescheduled to resume after the status date and some activities were performed out of sequence causing breaks in the original logic. There was also no schedule narrative that accompanied each status update and addressed the status of key activities, a description of the critical path, changes in network logic, or a schedule baseline document that existed outside the schedule that described the reasons for lags and constraints, a description of critical risks and any schedule contingency, or how the critical path was derived. Finally, because the schedule was not well constructed, this calls into question the validity of the critical path and any variances being measured against the baseline. Until the schedule has been updated to address the issues discussed in this report, the schedule baseline cannot be considered reliable and any variances resulting from comparing actual status to the baseline are questionable.

Source: GAO analysis of IRS.gov's July 2012 integrated enterprise portal integrated master schedule.

Table 22: List of Key Risks and Mitigation Activities for IRS.gov as of September 2012

Risk	Mitigation
Schedule delays due to scope or funding changes or delays	Officials responsible for individual task orders and overall program management executives will monitor and track the schedule and funding to ensure timely completion of all milestones and proper allocation of funds. In addition, the Executive Governance Board will see and address all issues in a timely manner.
Unexpected legislative mandates will negatively affect program	Program officials will monitor legislative mandates for impact to program and assess impact and determine approach, and escalate to executive level as needed based upon the type of mandate.
Program scope changes due to changes to operational standards	Program officials will monitor environment for impact and determine approach. They will escalate changes to the Executive Governance Board to resolve issues about Request for Proposal requirements and new standard requirements.

Source: IRS.

Modernized e-File (MeF)

MeF — Investment Details

Program start date:	November 2005
Full operational capability:	Time frame not available
Costs for FY11 to FY16:	$292.14 million
Current life-cycle phase:	Mixed (development/ operations and maintenance)

Source: IRS data.

MeF is the primary system interface for all electronic tax filings. When MeF receives an electronic tax return, the system validates the return and creates an acknowledgement. If the electronic return passes validation, MeF generates an accepted acknowledgment. If it fails validation, the system generates a reject acknowledgement and sends information concerning the cause(s) back to the transmitter. Returns that successfully pass validation are accepted and forwarded to the IRS systems used for processing tax returns. MeF stores all tax return data in a modern format in a modernized tax return database, allowing authorized viewers to see tax returns securely online using the IRS intranet. The most current release, MeF Release 7, included the rollout of over 129 remaining 1040-family schedules and forms, including forms 1040A and 1040EZ, which expanded the reach of MeF to the entirety of the e-File population, or approximately 98.3 million filers. Release 8 will implement added performance tuning, stabilization and monitoring. Release 9 is to add the employment and unemployment tax family of forms and the U.S. Income Tax Return for Estates and Trusts to the MeF environment along with an RRP interface.

Table 23: Assessment of the MeF Program's Cost Estimate

Characteristic	Assessment	Key examples of rationale for assessment
Comprehensive	Minimally met	MeF's cost estimate was only minimally comprehensive because it did not include both government and contractor costs for the life cycle of Release 8. It also did not include documentation that completely defined the program, reflected the current schedule, and was technically reasonable. While work breakdown structure documents, technical baselines, briefing to management, and contract documentation were provided, these documents did not comprise the entire life cycle of Release 8. In addition, the work breakdown structure was incomplete, inconsistent, and untraceable to other documentation and therefore could not be used to track the program costs and schedule by defined deliverables.
Well-documented	Not met	The cost estimate was not well-documented because the program did not provide data used to derive the estimate, a data dictionary, or a correct independent government cost estimate. Further, some documentation, such as work breakdown structure documents, were consistently incomplete, inconsistent, and at such a level as to preclude their use as decision-making documentation. While the technical baseline was provided and corresponded to the task order descriptions in the contract award, it was provided at such a high level of detail it is difficult to see how a reliable cost estimate could be derived from its information.
Accurate	Minimally met	The estimate for MeF was minimally accurate because neither a risk analysis nor an uncertainty analysis was included and therefore we could not determine if the estimate was unbiased. When documentation was provided, such as the cost work breakdown structure documents, there were inconsistencies, mistakes, and no data to support high-level analyses. Furthermore, there was conflicting evidence as to whether the cost work breakdown structure documents were updated to reflect changes in assumptions or actuals as the program developed. While MeF asserted that it had access to and used historical data for development of cost estimate, no documentation was provided to validate this.

Characteristic	Assessment	Key examples of rationale for assessment
Credible	Minimally met	The estimate for MeF was minimally credible because it did not identify the limitations of the analysis due to uncertainty or bias surrounding the data and assumptions. In addition, no risk or uncertainty analysis was provided, suggesting the program office did not factor uncertainty into their cost estimate or three-point estimates of effort hours. Furthermore the three-point estimates for the cost elements did not provide data to substantiate their values of best case, most likely case, and worst case effort hours estimate. MeF also did not provide sufficient documentation of an independent cost estimate and how it was reconciled with its cost estimate.

Source: GAO analysis of MeF's October 2011 release 8 cost estimate.

Table 24: Assessment of the MeF Program's Schedule Estimate

Characteristic	Assessment	Key examples of rationale for assessment
Well-constructed	Minimally met	The MeF schedule was minimally well constructed because 48 percent of the remaining activities were missing predecessor and successor logic. In addition, 44 percent of the remaining activities had dangling logic; that is, while these activities had predecessors and successors, either the predecessors were not affecting the start date of the activity, or the activity's finish date had no impact on the start date of successor tasks. When the schedule logic is incomplete, the ability of the schedule to dynamically respond to changes is compromised and the schedule loses its effectiveness as a management tool. In addition, there were a significant number of constraints, lags, and leads present in the schedule, which also affected the schedule network logic and the schedule's ability to forecast valid dates. Our analysis also determined that the critical path was not valid. We found 29 percent of the activities marked critical were level of effort activities that cannot be on the critical path because they do not represent discrete effort. There were also two critical activities with constraints and 36 percent of the critical activities contained lags or leads. Until missing and incomplete logic is resolved and constraints, lags, and leads are either removed or justified, the critical path will remain invalid. Further, total float was not reasonable. For example, the schedule had total float ranging from -556 days to 157 days (meaning that some activities were behind schedule more than 2 years while others could slip up to 7 months without affecting the final finish date). Abnormal total float values are usually caused by missing logic, a condition present in the schedule. When total float values are not reliable, management cannot trust the forecasted dates and cannot use the critical path to proactively manage the program.

Characteristic	Assessment	Key examples of rationale for assessment
Controlled	Partially met	We found that the schedule was partially controlled. For example, the schedule had a baseline that had been recently set from which to measure progress against. However, we found several activities with planned start and finish dates in the past that had not been updated or brought forward to occur after the status date. There were also several activities with actual start and finish dates more than a month after the status date. Furthermore, 14 activities had out-of-sequence logic, meaning they had started before their predecessors had finished causing the network logic to break. In accordance with best practices, the MeF program management plan stated that the program should hold weekly schedule status meetings. However, we found no evidence that these meetings were taking place. We also did not find evidence of a schedule narrative that should accompany each status update to address the status of key activities, the critical path, and changes in network logic. Though the program had set baselines from which to measure progress against, there was no corresponding baseline schedule document which defined custom fields in the schedule, justifications for constraints, lags, and long activity durations, a description of critical risks and any schedule contingency, or how the critical path was derived. Finally, because we found that the MeF schedule was not well constructed, this causes us to question the validity of the critical path and any variances being measured against the baseline. Until the schedule has been updated to address the issues discussed in this report, the schedule baseline cannot be considered reliable and any variances resulting from comparing actual status to the baseline are questionable.

Source: GAO analysis of MeF's November 2012 release 8 integrated master schedule.

Table 25: List of Key Risks and Mitigation Activities for MeF as of September 2012

Risk		Mitigation
Interface files	•	Release 8 will include interface-specific activities to ensure optimal performance in Filing Season 2013.
	•	Release 8 will include several automated monitoring techniques aimed at providing alert notification and reporting, as well as specific activities to ensure optimal performance in Filing Season 2013.
Volume of returns processed	•	Release 8 will include significant platform configuration and application architecture changes and extensive performance testing to ensure that it can support Filing Season 2013 volumes.
	•	IRS will maintain part of the 1040 legacy system for Filing Season 2013, comprised of 25 forms and schedules representing the highest return volume.

Source: IRS.

Return Review Program (RRP)

RRP — Investment Details

Program start date:	April 2010
Full operational capability:	Time frame not available
Costs for FY11 to FY16:	$87.31 million
Current life-cycle phase:	Mixed (development/ operations and maintenance)

Source: IRS data.

RRP is an integrated and unified system/application and is intended to be deployed in 2014. It is to provide the following functionality:

- enhance productivity of the scheme development center, investigative analysts, and aids;

- enable more effective routing of returns;

- detect noncompliant and fraudulent returns;

- ensure timely issuance of refunds and credits;

- prevent issuance of refunds and credits not legally due to filers;

- provide the platform to implement the pre-refund program new business model (functional capabilities and business rules);

- grow and change to meet changing noncompliant and fraud detection needs;

- streamline business processes used by the IRS criminal investigative analyst, aides, and civil employees; and

- replace Electronic Fraud Detection system with modernized technology to provide the foundation for the pre-refund new business model and achieve some high-priority pre-refund capabilities.

The new system is comprised of three major activities:

- **Detection.** The system will incorporate several existing models as well as new models to enhance detection of probable noncompliance. Using algorithms and business rule sets, the system will detect questionable information on each return as the return is processed. The system will also detect returns with potential fraud characteristics, thereby allowing criminal investigators to link and analyze groups of returns to identify schemes for potential criminal prosecution.

- **Resolution.** The system will accommodate existing treatment streams and new treatment streams. Returns will be routed systemically to the best treatment stream, opened into that treatment stream's inventory and, if applicable, the system will send an initial contact letter to the taxpayer.

- **Prevention.** The system will automatically integrate the results of each return's resolution into the detection models. The results can be used to help target education and outreach efforts to taxpayers and preparers on how to avoid unintentional noncompliance. The system will also allow analysis and identification of fraud and noncompliance not identified by the predictive detection models.

Table 26: Assessment of the RRP Program's Cost Estimate

Characteristic	Assessment	Key examples of rationale for assessment
Comprehensive	Substantially met	RRP's cost estimate was substantially comprehensive because it documented all relevant costs associated with the program; however, weaknesses were noted. Specifically, while cost estimate documents were developed that anticipated life-cycle costs at the point in time that the estimate was completed based on information available at that time, these were not technical baselines and the cost estimate did not contain all the typical elements expected in a technical baseline. Additionally, while the estimate documented ground rules and assumptions, it did not capture all of the general ground rules and assumptions in accordance with best practices. Specifically, government-furnished equipment and ground rules and assumptions associated with delivery partners were excluded from both documents. Finally, the estimate did not provide a product-oriented work breakdown structure that was traceable to the statement of work.
Well-documented	Partially met	The estimate for RRP was partially well-documented. While it contained procedures that consisted of the major steps associated with the methodology used to develop the estimate, it lacked details including information about (1) the work breakdown structure, (2) how the specific cost elements were defined, (3) calculations used, (4) links to the input data, and (5) validation of the estimate. Further, an independent estimate was not developed and the estimate did not capture the source data used or the reliability of the data. Finally, information was not provided showing management's acceptance of the cost estimate including recommendations for changes, feedback, and the level of contingency reserves decided upon to reach a desired level of confidence.
Accurate	Substantially met	RRP's cost estimate was substantially accurate. For example, the basis of estimate states that the probability of the effort is at the 50 percent confidence level and IRS provided evidence that the nonrecurring cost estimate represented a 52 percent confidence level. However, the data backing up the 52 percent confidence level were not transparent and there was no level of confidence provided for the recurring cost estimate which comprises almost half of the life cycle cost estimate. As a result, we cannot be sure that the total cost estimate represents the most likely costs. We checked the data in the ISR Extract that contains the detailed cost breakdown for the program and found them to be reliable as there were only a few small rounding errors. Additionally, IRS said that since January 2012, inflation-adjusted costs are contained in the 5-year Investment Summary Report in the Estimation Program Office's estimates. We verified this claim by reviewing the Investment Summary Report Extract which showed cost information for 5 years that was indeed adjusted for inflation. The RRP program has updated the cost estimate three times. The original cost estimate was developed in 2008 and then updated again in 2010 and in March 2012. According to officials, these updates contained the most current cost information known at the time. Furthermore, the estimate for transition state 1 MS3 contained information comparing the project initiation and estimate to actual cost data which showed slight variances between Project Initiation Estimated full-time employees and actual full-time employees. However, variance data from October 2011 to July 2012 showed no reported cost or schedule variance which was inconsistent with the estimate documentation. Finally, while IRS had evidence showing that it relied on historical data, there was no documentation associated with lessons learned for noted differences.

Characteristic	Assessment	Key examples of rationale for assessment
Credible	Minimally met	RRP's cost estimate was minimally credible because it was developed by an independent office. However, the program did not provide any evidence that it developed its own estimate. In addition, while the program did not conduct a sensitivity analysis as part of the basis of estimate, it explained that IRS tends to use standard confidence levels for various pieces of the work (e.g., software uses an 80 percent confidence level). Finally, while the basis of estimate provided minimal information on how data were cross-checked, we were unable to verify what was done to determine if similar results were produced using various methods and what those results were.

Source: GAO analysis of RRP's March 2012 transition state 1 milestone 3 cost estimate.

Table 27: Assessment of the RRP Program's Schedule Estimate

Characteristic	Assessment	Key examples of rationale for assessment
Well-constructed	Minimally met	RRP's schedule was only minimally well constructed because it was missing predecessor and successor logic, had too many constraints, contained lags and leads, and had several activities with too many predecessors, all of which call into question the ability of the activities to finish on time. Because of these issues, the critical path was not valid. For instance, the schedule had hundreds of activities marked as critical due to the overabundance of constraints in the schedule. Specifically, almost 40 percent of the transition state 1 activities and 80 percent of the transition state 2 activities had constraints that were preventing activities from starting or finishing earlier, thereby causing them to be critical. We traced the driving paths and found far fewer activities were actually affecting the finish date. As a result, monitoring the critical path was not a useful exercise because the activities on it were not the most important ones impacting the end date. In addition, we found that total float was not reasonable. For example, the transition state 1 schedule had total float ranging from -88 days (meaning 4 months behind schedule) to 723 days (meaning an activity could slip almost 3 years without affecting the final finish date). On the other hand, the transition state 2 schedule had only positive float ranging from 0 days (meaning each day of slippage results in the same amount of delay on the end date) to a maximum of 194 days (meaning an activity could slip almost 9 months). High total float values are usually caused by missing logic, a condition that we found in the schedule. When total float is not reliable, management cannot trust the forecasted dates and cannot use the critical path to proactively manage the program.

Characteristic	Assessment	Key examples of rationale for assessment
Controlled	Partially met	RRP's schedule was partially controlled because it had been recently updated and had set baselines from which to measure progress against. In addition, program officials held weekly schedule status meetings and provided evidence that they were actively managing the program using information from the schedule. However, there were some date anomalies, such as activities that were scheduled to start or be completed in the past that had not been rescheduled to resume after the status date. When we fixed the schedules to bring forward uncompleted work in the past, the finish date for the transition state 1 schedule slipped from December 30, 2013, to February 14, 2014. However, the finish date did not change for the transition state 2 schedule due the many constraints placed on remaining activities. In addition, there were a few activities that were marked as complete several months into the future, which is not reasonable. For example, there were several activities that had actual start and finish dates recorded in May and June 2013 which is well past the status date of October 26, 2012. There were also some activities that were performed out of sequence causing breaks in the original logic. Program officials stated that the scheduler responsible for maintaining the schedules had suddenly left and the new scheduler agreed that there were many problems with the schedule that needed to be addressed. The new scheduler is working to remediate the issues we found and is committed to improving the schedule in the near future. While these efforts are positive, because the schedule was not well constructed, this calls into questions the validity of the critical path and any variances being measured against the baseline. Until the schedule has been updated to address the issues discussed in this report, the schedule baseline cannot be considered reliable and any variances resulting from comparing actual status to the baseline are questionable.

Source: GAO analysis of RRP's October 2012 transition states 1 and 2 integrated master schedule.

Table 28: List of Key Risks and Mitigation Activities for RRP as of September 2012

Risk	Mitigation
Technology insertion risk	• Adopting several leading-edge technologies and leading end-to-end integration, engineering, and capacity performance to allow them to work together. • Applying open-source solutions to contain cost. • Including technologies such as massive parallel processing for fast data processing, best-in-class data analytics, and a business rules engine to address dynamic fraud schemes and supporting commercial off-the-shelf Products .

Source: IRS.

Appendix IV: Comments from the Department of Treasury

DEPARTMENT OF THE TREASURY
INTERNAL REVENUE SERVICE
WASHINGTON, D.C. 20224

COMMISSIONER

April 4, 2013

Mr. James R. White
Director, Tax Issues, Strategic Issues Team
U.S. Government Accountability Office
441 G Street, NW
Washington, DC 20548

Dear Mr. White:

Thank you for the opportunity to comment on the draft report titled, *Information Technology: Consistently Applying Best Practices Could Help IRS Improve the Reliability of Reported Cost and Schedule Information*, (GAO-13-401). We recognize the importance of reliable cost and schedule estimating and reporting. We will continue our effort to improve management of our investment portfolio and value your support and guidance.

We also appreciate that GAO recognized our progress in strengthening our controls and capabilities resulting in the removal of the Business Systems Modernization program from the GAO high-risk list.

We will provide the detailed corrective action plan addressing the recommendations with our response to the final report. If you have any questions, please contact me or a member of your staff may contact Terence V. Milholland, Chief Technology Officer, at (202) 622-6800.

Sincerely,

Steven T. Miller
Acting Commissioner

Appendix V: GAO Contact and Staff Acknowledgments

GAO Contact

David A. Powner, (202) 512-9286 or pownerd@gao.gov

Staff Acknowledgments

In addition to the contact name above, individuals making contributions to this report included Sabine Paul (Assistant Director), Remmie Arnold, Tisha Derricotte, Jennifer Echard, Emile Ettedgui, Rebecca Eyler, Mary Fike, Lisa Hardman, Sairah Ijaz, Sandra Kerr, Jason Lee, Paul Middleton, Karen Richey, and Jeremy Sebest.