

**A Test of the Private Sector's Ability to
Provide Tax Law Telephone Assistance Did
Not Produce Sufficient Information to Support
a Competitive Sourcing Decision**

July 2005

Reference Number: 2005-30-049

This report has cleared the Treasury Inspector General for Tax Administration disclosure review process and information determined to be restricted from public release has been redacted from this document.



DEPARTMENT OF THE TREASURY
WASHINGTON, D.C. 20220

INSPECTOR GENERAL
for TAX
ADMINISTRATION

July 8, 2005

MEMORANDUM FOR COMMISSIONER, WAGE AND INVESTMENT DIVISION
CHIEF, AGENCY-WIDE SHARED SERVICES

Pamela J. Gardiner

FROM: Pamela J. Gardiner
Deputy Inspector General for Audit

SUBJECT: Final Audit Report - A Test of the Private Sector's Ability to
Provide Tax Law Telephone Assistance Did Not Produce
Sufficient Information to Support a Competitive Sourcing
Decision (Audit # 200330043)

This report presents the results of our review to evaluate whether the Internal Revenue Service (IRS) designed and administered the competitive sourcing¹ Toll-Free Tax Law Services Test in a manner that would render a reliable assessment of a private vendor's ability to provide comparable service, in terms of both access and quality, to taxpayers making tax law inquiries on the IRS toll-free telephone system. Competitive sourcing is one of five initiatives in the President's Management Agenda for improving the management and performance of the Federal Government.

In summary, we concluded the private-vendor Test had design deficiencies and the results were not reliable due to the significant limitations of the data collected. Consequently, the IRS did not derive the actual expected benefit of the \$675,140 it paid the vendor that participated in the Test. The IRS also spent an undetermined amount of funds² for professional consulting services to assist it in developing and evaluating the private-vendor Test.

¹ Competitive sourcing is the process for determining whether a commercial activity will be performed by a public or private source. The goal of competitive sourcing is to achieve maximum value for tax dollars spent by allowing the private sector to propose its own methods for delivering goods and services.

² The IRS advised us on September 21, 2004, that over \$4.5 million had been expended on a multitask consulting contract that included the Toll-Free Tax Law Telephone Operation and other projects. However, the IRS did not capture the individual costs for each project.

Several factors led us to conclude that the Test had design deficiencies. First, the framework of the IRS private-vendor Test lacked conformity with the goal of competitive sourcing because the IRS tested the duplication of its Probe and Response Guide (PRG)³ method for answering tax law questions in a private-sector setting.

Second, the design of the Test created an inequitable comparison between the IRS workforce, where 77 percent of the customer service representatives (CSR) had 2 or more years' experience in providing tax law services, and the vendor's CSR staff that had little or no previous tax law knowledge. The brevity of the 60-day Test period precluded the vendor's CSRs from reaching the learning curve point at which actual experience would enhance performance.

Next, the IRS compromised the integrity of the Test by not developing an evaluation plan that predefined essential performance measurement target values with which to assess the project outcomes. Finally, the manner in which the private-vendor Test was administered increased the challenges to performing a future competition. Unlike service providers in an actual competitive marketplace, the IRS influenced its competitor (i.e., the vendor) to duplicate its PRG method for answering tax law inquiries, which negated some of the benefits of competition.

Although the IRS report⁴ on the Test included performance and telephone result metrics for the vendor, neither the report nor the IRS Competitive Sourcing Program Decision Document, issued in September 2004, stated the IRS' conclusion about the vendor's capabilities in delivering telephone tax law assistance. In discussions related to our report, IRS officials provided further clarification concerning the decision not to pursue a Business Case Analysis (BCA) by stating, "The outcome of the feasibility test (lower quality and significant taxpayer perception regarding privacy and confidentiality) coupled with the considerable adverse impact on IRS employee morale and productivity as well as the consolidation of the W&I [Wage and Investment Division] and SB/SE [Small Business/Self-Employed Division] CAS [Customer Account Services] organization, were factors which supported the decision to not pursue a BCA at this time." IRS officials also believe some benefits were realized for the funds spent. For example, the vendor responded to tax law inquiries for 43 percent of the 227,390 calls initially routed to it.⁵ The calls answered by the vendor represented an increase in calls answered beyond the volume the IRS planned to answer using its own resources.

We recommended the Deputy Commissioner for Operations Support and the Commissioner, W&I Division, rescind the September 2004 decision and eliminate use of the private-vendor Test results as justification for any decision regarding the merits of using the Office of Management and Budget Circular No. A-76 competitive sourcing

³ An IRS-developed intelligent system designed to assist customer service representatives in providing correct answers to tax law inquiries and in maintaining a consistent level of customer and procedural accuracy.

⁴ Test Study Summary Report: *IRS Competitive Sourcing Program, Delivery of Tax Law Services*, dated September 14, 2004.

⁵ The IRS report on the Test showed the vendor provided correct answers to only 44.6 percent of the callers it assisted.

process to determine the most efficient and cost-effective toll-free tax law telephone services provider. To ensure any future studies of competitive sourcing of the Toll-Free Tax Law Telephone Operation produce more reliable and useful results, we also recommended the Chief, Agency-Wide Shared Services (AWSS), expand the *IRS Guide to Competitive Sourcing* to require future studies to conform to a structured research design model and data collection protocol; require creation of an evaluation plan that predefines performance standards with target performance values and composite scores that constitute success; involve the IRS Research Division in the review of the design plan and data collection procedures, prior to performance of any tasks; and require executive-level approval of the research design plan prior to implementation.

To minimize the risk or appearance of unfairness in the sourcing process related to any future study of the toll-free tax law telephone services, we recommended the Commissioner, W&I Division, and the Chief, AWSS, require any future Evaluation and Source Selection team members and Source Selection Official to sign a statement certifying they have no personal impairments that inhibit their ability to make a fair and impartial decision regarding the vendor proposals. We also recommended the Commissioner, W&I Division, and Chief, AWSS, ensure any future solicitation related to the Toll-Free Tax Law Telephone Operation contains no expressed or implied requirements that private vendors use current IRS methodologies in proposing solutions to the requirements in the solicitation.

Management's Response: Management's response was due on July 5, 2005. As of July 6, 2005, management had not responded to the draft report.

Copies of this report are also being sent to the IRS managers affected by the report recommendations. Please contact me at (202) 622-6510 if you have questions or Richard J. Dagliolo, Acting Assistant Inspector General (Small Business and Corporate Programs), at (631) 654-6028.

A Test of the Private Sector’s Ability to Provide Tax Law Telephone Assistance Did Not Produce Sufficient Information to Support a Competitive Sourcing Decision

Table of Contents

Background Page 1

The Private-Vendor Test Did Not Produce Reliable Results for
Deciding Whether to Open the Toll-Free Tax Law Telephone
Operation to Public-Private Competition Page 3

Recommendations 1 and 2:Page 15

Recommendations 3 and 4:Page 16

Appendix I – Detailed Objective, Scope, and Methodology Page 17

Appendix II – Major Contributors to This Report Page 18

Appendix III – Report Distribution List Page 19

Appendix IV – Outcome Measures..... Page 20

A Test of the Private Sector's Ability to Provide Tax Law Telephone Assistance Did Not Produce Sufficient Information to Support a Competitive Sourcing Decision

Background

The Federal Activities Inventory Reform (FAIR) Act of 1998¹ requires agencies to annually submit to the Office of Management and Budget (OMB) a listing of all activities performed that are either commercial or inherently Governmental in nature. According to the FAIR Act, activities that are commercial in nature are suitable for a public-private competition to ensure the best value is delivered to the American taxpayers.

OMB Circular No. A-76, *Performance of Commercial Activities*, establishes policy and procedures for the Federal Government's reliance on the private sector for essential commercial services. Competitive sourcing is one of five objectives of the President's Management Agenda (PMA) for improving the management and performance of the Federal Government. Competitive sourcing is the process for determining whether Federal Government resources or the private sector can more efficiently and cost effectively provide services currently provided by the Federal Government.

In Fiscal Year (FY) 2002, the Internal Revenue Service (IRS) identified its Toll-Free Tax Law Telephone Operation as a possible candidate for competitive sourcing. The IRS toll-free telephone system is a primary contact method millions of taxpayers choose when seeking answers to tax law questions or trying to resolve account issues involving refunds, balance due billing activity, and changes to the amount of tax owed. During the 2004 Filing Season,² the IRS received nearly 13.6 million incoming calls for the individual income tax law services offered on its 1-800-829-1040³ toll-free telephone line.

To evaluate whether the Toll-Free Tax Law Telephone Operation was a viable candidate for competitive sourcing,

¹ Pub. L. No. 105-270, 112 Stat. 2382.

² The filing season is the period from January through mid-April when most individual income tax returns are filed. All references to the 2004 Filing Season made in this report, unless otherwise specified, are for the period January 2 to April 16, 2004.

³ The primary toll-free number called by customers with tax law questions. The IRS screens each call to this number to determine the topic of the inquiry and uses intelligent call management software to route the call to an assistor trained in that tax law topic.

A Test of the Private Sector's Ability to Provide Tax Law Telephone Assistance Did Not Produce Sufficient Information to Support a Competitive Sourcing Decision

the IRS decided to subject the task to a test of vendor capability to determine whether a vendor could deliver equal or superior quality in responding to tax law inquiries compared to the level of quality achieved by IRS resources. A second objective was to assess the public's reaction to receiving responses to tax law inquiries from a commercial vendor rather than the IRS. The IRS awarded a contract on August 28, 2003, to the vendor that offered the lowest bid of \$675,140. The selected vendor was responsible for handling approximately 10 percent of the incoming tax law calls received by the IRS during the 60-day period from February 16 through April 16, 2004.

This audit was performed in accordance with *Government Auditing Standards (GAS)* during the period September 2003 through August 2004.⁴ The IRS management official responsible for the task we audited did not provide sufficient, competent, and relevant evidence to support the reliability of the approach the IRS used to design, administer, and evaluate the private-vendor Toll-Free Tax Law Services Test. Therefore, to comply with *GAS* in our evaluation of the design and administration of the Test, we developed evaluation criteria using applicable Federal Government regulations, policies, and procedures along with standard business and research test practices.

We applied Federal Government regulations and practices relative to competitive sourcing because the stated goal of the IRS study of the Toll-Free Tax Law Telephone Operation was to determine whether it was a viable candidate for the competitive sourcing process under the OMB Circular No. A-76 initiative. We applied standard project management practices because the composition of the private-vendor Test was consistent with the standard

⁴ To perform this audit, we interviewed IRS staff and/or analyzed documentation from the Wage and Investment Division Customer Account Services offices in Atlanta, Georgia, and Dallas, Texas; the Office of Competitive Sourcing in Washington, D.C., and Brooklyn, New York; the Office of Procurement in Oxon Hill, Maryland; and the Agency-Wide Shared Services office in Austin, Texas. We also interviewed staff and reviewed available documentation at the vendor's call center in Cumberland, Maryland.

A Test of the Private Sector's Ability to Provide Tax Law Telephone Assistance Did Not Produce Sufficient Information to Support a Competitive Sourcing Decision

The Private-Vendor Test Did Not Produce Reliable Results for Deciding Whether to Open the Toll-Free Tax Law Telephone Operation to Public-Private Competition

definition of a project⁵ under the project management discipline. We applied standard research test practices used by Federal Government agencies and private industry because they are widely recognized best practices for performing valid research tests that produce reliable and defensible results.

Detailed information on our audit objective, scope, and methodology is presented in Appendix I. Major contributors to the report are listed in Appendix II.

In September 2004, the IRS Competitive Sourcing Program Decision Document regarding the Toll-Free Tax Law Services Test announced the following: “The Deputy Commissioner Support Operations decided not to pursue a Business Case Analysis [BCA]⁶ at this time due to the re-engineering/restructuring initiative recently started to merge the SB/SE [Small Business/Self-Employed Division] CAS [Customer Account Services] organization into W&I [Wage and Investment Division].”

Although the IRS report⁷ on the Test included performance and telephone result metrics for the vendor, neither the report nor the decision document stated the IRS' conclusion about the vendor's capabilities in delivering telephone tax law assistance. The draft report of the Test results presented to the IRS Deputy Commissioner for Operations Support in August 2004 stated, “The Competitive Sourcing Office recommends that we complete the Business Case Analysis (BCA) prior to any final decision. The Business Owner, W&I, recommends conducting a subsequent test.” In a post operational analysis report submitted by the vendor that participated in the Test, it concluded that outsourcing the Toll-Free Tax Law Telephone Operation was feasible based upon the vendor's positive ratings in four of the five quality

⁵ A project is a temporary work condition involving a series of related tasks that require multiple resources to reach a predetermined goal.

⁶ A BCA is used to manage business process improvement activities through examining the benefits and risks of a proposed initiative. An accompanying experimental project to test the viability of an approach might be justified when the project involves substantial changes to service delivery and affects many stakeholders.

⁷ Test Study Summary Report: *IRS Competitive Sourcing Program, Delivery of Tax Law Services*, dated September 14, 2004.

A Test of the Private Sector's Ability to Provide Tax Law Telephone Assistance Did Not Produce Sufficient Information to Support a Competitive Sourcing Decision

factors measured by the IRS and the improved performance in the one factor that did not meet expectations.

In discussions related to our report, IRS officials provided further clarification concerning the September 2004 decision by stating, "The outcome of the feasibility test (lower quality and significant taxpayer perception regarding privacy and confidentiality) coupled with the considerable adverse impact on IRS employee morale and productivity as well as the consolidation of the W&I and SB/SE CAS organization, were factors which supported the decision to not pursue a BCA at this time."

However, based on our review, we concluded the private-vendor Test had design deficiencies and the results were not reliable due to the significant limitations of the data collected. Consequently, the IRS did not derive the actual expected benefit of the \$675,140 it paid the vendor that participated in the Test. The IRS also spent an undetermined amount of funds⁸ for professional consultant services to assist it in developing and performing the study of the Toll-Free Tax Law Telephone Operation along with the development and evaluation of the private-vendor Test. IRS officials believe some benefits were realized for the funds spent. For example, the vendor responded to tax law inquiries for 43 percent of the 227,390 calls initially routed to it.⁹ The calls answered by the vendor during the 2004 Filing Season represented an increase in calls answered beyond the volume the IRS planned to answer using its own resources.

The private-vendor Test could have been better designed to be consistent with the fundamental goal of competitive sourcing

The goal of competitive sourcing as established by Federal Government policy and the PMA is to achieve maximum value for tax dollars spent by offering the private sector an

⁸ The IRS advised us on September 21, 2004, that over \$4.5 million had been expended on a multitask consulting contract that included the Toll-Free Tax Law Telephone Operation and other projects. However, the IRS did not capture the individual costs for each project.

⁹ The IRS report on the results of the Test showed the vendor provided correct answers to only 44.6 percent of the callers it assisted.

A Test of the Private Sector’s Ability to Provide Tax Law Telephone Assistance Did Not Produce Sufficient Information to Support a Competitive Sourcing Decision

opportunity to propose its own methods for delivering goods and services defined by Federal Government requirements. The framework of the IRS project did not conform to this concept because the IRS tested the duplication of its Probe and Response Guide (PRG)¹⁰ method for answering tax law inquiries in a private-sector setting. The IRS report on the Test states, “The mandated use of the Probe and Response Guide had an impact on the accuracy rate. The vendor did not have the opportunity to develop their own probe and response system.”

Prior to initiating the Test, the IRS surveyed eight private-sector vendors offering a wide range of customer service support to their clients. The IRS report on the survey stated, “An assessment was performed on the feasibility of duplicating the tax law and Area Distribution Center¹¹ telephone operations in a vendor setting. The industry leaders were surveyed for interest and capabilities in providing tax law services for the IRS.” Although seven vendors responded to the survey expressing interest in the opportunity to provide telephone tax law services for the IRS, none had any experience in offering this specialized service.

The project management discipline contains several business tools that can be employed to assist management in planning, implementing, and evaluating a project. A feasibility study¹² and BCA are structured analysis methodologies within this discipline. The project plan is a fundamental element of the project management process. It establishes the project boundaries, support structures, relationships between issues, and detailed steps to accomplish the intended goals. The application of standard project management techniques in the planning process

¹⁰ An IRS-developed intelligent system designed to assist customer service representatives in providing correct answers to tax law inquiries and in maintaining a consistent level of customer and procedural accuracy.

¹¹ An Area Distribution Center is a multimedia storage facility where internal and external requests for published products are processed.

¹² A feasibility study is a structured methodology for documenting the operational, economical, and technical analysis of a business problem. It includes a determination of whether the problem can be solved effectively.

A Test of the Private Sector's Ability to Provide Tax Law Telephone Assistance Did Not Produce Sufficient Information to Support a Competitive Sourcing Decision

could have assured the development of a reliable, systematic, goal-oriented test approach that connected the objective and scope of the Toll-Free Tax Law Services Test to the overarching objectives of the W&I Division strategic plan and the competitive sourcing initiative.

Although the IRS defined the objective of the Test and estimated its Toll-Free Tax Law Telephone Operation included 1,100 Full-Time Equivalents (FTE),¹³ it did not formalize plans for implementing the Test. In discussions about the Test, the project manager confirmed the IRS did not formally document the project management framework and other key elements specific to its implementation. Consequently, the project objective was never directly linked to the goal of the competitive sourcing initiative.

To properly align the private-vendor Test with the goals of competition¹⁴ and competitive sourcing,¹⁵ the scope of work needed to include the assembling of sufficient information for an unbiased assessment of a vendor's alternative methods for delivering toll-free tax law telephone services. The planned project outcomes should have included the demonstration of whether service delivery could be improved by the competitive sourcing process. By not using accepted project management tools, the project was put on a course that would not likely provide relevant evidence for decision-making purposes.

The design of the private-vendor Test created an inequitable comparison between the IRS and a novice in providing tax law services

The IRS contract with the selected vendor was awarded on August 28, 2003, with a planned Test start date of

¹³ An FTE is a measure of labor hours in which 1 FTE is equal to 8 hours multiplied by the number of compensable days in a particular fiscal year. For FY 2004, 1 FTE was equal to 2,096 labor hours.

¹⁴ The term competition as defined in this report is, "the effort of two or more parties acting independently to secure the business of a third party by offering the most favorable terms."

¹⁵ Competitive sourcing is holding public-private competitions that compare the performance of a Federal Government organization to that of a private sector or other non-Federal organization. Federal Acquisition Council's *Manager's Guide to Competitive Sourcing* (dated October 2, 2003).

A Test of the Private Sector's Ability to Provide Tax Law Telephone Assistance Did Not Produce Sufficient Information to Support a Competitive Sourcing Decision

February 16, 2004. The Statement of Work (SOW)¹⁶ required the vendor to provide sufficient staff trained in 11 tax law topic areas ranging from basic procedural and filing requirement issues to very complex individual return issues such as installment sales. Five of the 11 tax law topics form the fundamental building blocks for the more complex topics. The 65 customer service representatives (CSR) the vendor assigned to this project did not initially possess the body of knowledge needed to fulfill the terms of the contract. In addition to hiring and training the required CSR staff, the vendor had to establish and train its quality assurance team, purchase the necessary telecommunications and computer equipment, establish connectivity to the IRS, and meet the requirements for providing telephone system call data in a format comparable to the IRS' telephone system data.

The IRS' August 2003 risk assessment stated, "The timeframe for vendor ramp up is considerably aggressive for the complex scope of the test study. Vendors have not previously handled tax law subject matter." Expertise in tax law was the most critical comparability factor for weighing the success or failure of the CSRs that responded to tax law inquiries from the public. The IRS allowed the vendor only 5½ months to prepare for and implement the Test.

Of the approximately 3,359¹⁷ IRS CSRs to which the vendor's performance was compared, about 2,600 (77 percent) had 2 or more years' experience in tax law subject matter while about 350 CSRs were newly hired for the 2004 Filing Season. Within the IRS, new CSRs are initially trained in only 2 of the 11 tax law topics included in the scope of the Test and are allowed time to gain proficiency in those topics before advancing to other topics in their second year. The brevity of the 60-day Test period precluded the vendor's CSRs from reaching the learning

¹⁶ The SOW is a description of the Federal Government's technical requirements and the work to be performed by the contractor. It specifies the terms of the contract and is the standard for measuring contractor performance.

¹⁷ The IRS advised that in calculating the number of CSRs with 2 or more years' experience, the CSRs that transferred to other positions and returned to the CAS were included in more than 1 entry on duty date group and combined in the total.

A Test of the Private Sector's Ability to Provide Tax Law Telephone Assistance Did Not Produce Sufficient Information to Support a Competitive Sourcing Decision

curve point at which the training and actual experience would enhance performance. Therefore, it was unrealistic to expect the vendor's group of CSRs newly trained in 11 tax law topics to effectively perform at a level that would approach parity with that of the IRS' experienced CSRs. These events created unfair Test conditions with a high probability for unfavorable results.

The IRS could have accomplished a reliable comparison of the private sector's capabilities and competence in providing tax law assistance to the American public by using an experimental test design commonly used by Federal Government and private-sector organizations that perform research studies. In an experimental test design, relevant characteristics of the control group¹⁸ and experimental group¹⁹ populations should be similar before a variable condition is applied to the experimental group. The comparability factor is critical because the performance of the control group usually serves as a baseline for measuring the effect of variable conditions on the experimental group.

If an experimental test design had been used, the vendor that participated in the IRS Test would have represented the experimental group. In this instance, the only viable control group the IRS could have formed for making a fair comparison to the vendor's group would have been from the CSRs newly hired in October 2003. The scope of the tax law topics included in the vendor's responsibilities would have been limited to the topics handled by the IRS control group. The variable conditions applied to the experimental group would have consisted of the vendor's own methods and procedures for delivering service to the public.

The IRS could have taken reasonable actions to prevent the formation of Test parameters that produced skewed results if it had properly applied the knowledge gained from its market research and obtained essential staffing information from the vendor after the contract was awarded. However, in its written response to our request for information about

¹⁸ The control group is the group in an experimental study that does not receive the experimental condition.

¹⁹ The experimental group is the group in an experimental study that receives the experimental condition.

A Test of the Private Sector’s Ability to Provide Tax Law Telephone Assistance Did Not Produce Sufficient Information to Support a Competitive Sourcing Decision

the tax law skill level of the vendor staff, the IRS stated, “It is not a requirement of the contract for the vendor to give us information about the people they hired.”

The Test reliability was compromised because the IRS emphasized use of its methodology for accurately responding to tax law inquiries

The SOW for the private-vendor Test established the following requirement: “The service provider Quality Control Plan shall include a method for monitoring, measuring, and maintaining: Customer Accuracy, Regulatory Accuracy, Procedural Accuracy.” The SOW also stipulated the IRS would perform a quality review of the services provided by the vendor using the same methods it employs to review the quality of its own Toll-Free Tax Law Telephone Operation. In measuring its own Operation for procedural accuracy, the IRS rates its CSRs on their use of the PRG. Thus, the SOW implied the vendor’s adherence to the PRG would be an integral factor in the rating of its performance in accurately responding to tax law inquiries.

The IRS visited the vendor in January 2004 and stated the following in its evaluation of the vendor’s readiness to implement the test: “It will be critical to the service provider’s success to ensure employees receive a maximum amount of practice in using the on-line guide before taking live calls.” The IRS emphasized vendor use of the PRG even though some of its CSRs seemed to encounter problems using it. During the 2004 Filing Season, the customer accuracy rates achieved by the IRS declined from 81 percent to 79 percent when compared to the 2003 Filing Season. On April 2, 2004, the IRS’ assessment regarding its latest revisions to the PRG stated, “Our revisions to the Probe and Response Guide in FY 2003 have not enhanced the usability of this tool as we anticipated. We have begun the process of redesigning this critical tool to make it easier for our CSRs to complete all the needed probes and to provide correct answers.”

The IRS report on the Test showed the vendor achieved a customer accuracy rate²⁰ of only 44.6 percent during the

²⁰ The customer accuracy rate was a measure of whether the taxpayer who called received a correct and complete response.

A Test of the Private Sector's Ability to Provide Tax Law Telephone Assistance Did Not Produce Sufficient Information to Support a Competitive Sourcing Decision

60-day Test period. Conversely, the IRS report showed the vendor's performance ranged from 89.9 percent to 99.8 percent in other quality measures (i.e., regulatory accuracy, procedural accuracy, timeliness, and professionalism).²¹ The vendor's performance in these measures was comparable to that of the IRS and was above the IRS' performance goals set in two areas.

Overall, the private-vendor Test did not produce new knowledge on the merits of competitively sourcing the Toll-Free Tax Law Telephone Operation. In addition, the IRS missed the opportunity to benchmark its quality control methods against that of the private sector. Under genuine competitive test conditions, the vendor would have developed its own method for maintaining consistency in providing accurate responses to tax law inquiries and for measuring performance in accordance with agency requirements. A test structured in this manner would have had the added benefit of providing valuable information about a private vendor's capacity to systematically furnish complete and accurate answers to tax law inquiries.

The integrity of the Test was compromised because the IRS did not properly predefine essential performance measurement target values with which to evaluate the private-vendor Test

We requested the IRS provide its plan for evaluating the results of the private-vendor Test and the adjoining study. The IRS never provided a formal evaluation plan for our assessment. Instead, its written reply to one request stated, "We'll be reviewing customer accuracy, performance metrics for calls answered, and customer perception. There is no pre-determined 'passing grade' but rather our decision will be made based on the performance of the service provider."

Sound project management practices require that the initial framework of a project and any associated test(s) include

²¹ Regulatory accuracy measures those attributes that are required to be covered per regulations; procedural accuracy measures internal IRS requirements; timeliness measures the ability to handle the call timely; and professionalism measures the courtesy and professionalism of the CSR.

A Test of the Private Sector's Ability to Provide Tax Law Telephone Assistance Did Not Produce Sufficient Information to Support a Competitive Sourcing Decision

verifiable quantitative and qualitative performance evaluation criteria. The performance evaluation criteria should define the performance measurement target values and any appropriate composite scores that constitute achievement of a desired goal. This provides the basis for an objective assessment of the outcomes of the project activities and for an effective evaluation of potential alternatives.

The IRS design of the private-vendor Test did not include binding contract requirements with the vendor for achievement of any specific performance metric target values. In its written statement concerning the reason for not adhering to performance-based contracting methods,²² the IRS stated, “. . . so that at the conclusion of the test, we would be comparing like performance.” We believe this position did not negate the need for performance standards but, rather, required the development of measurable criteria with predefined target values and appropriate threshold composite scores to ensure an equitable comparison.

The IRS has developed and classified the following corporate performance indicators as critical to determining its success in providing high-quality service to each toll-free tax law caller:

- **CSR Level of Service (LOS)** – measures the relative success rate of taxpayers seeking assistance from a live assistor.
- **Customer Accuracy** – measures the percentage of taxpayers who received the correct answers to their tax law inquiries.
- **Assistor Calls Answered** – measures the volume of calls that received assistance from a CSR.

²² In procuring services, the Federal Acquisition Regulation, 48 C.F.R. pt 1-53 (2002), requires the use of performance-based contracts with measurable performance standards.

A Test of the Private Sector's Ability to Provide Tax Law Telephone Assistance Did Not Produce Sufficient Information to Support a Competitive Sourcing Decision

- **Timeliness** – measures responsiveness to callers in using elements of time controlled by the CSR (i.e., talk time, hold time, and wrap time).²³
- **Professionalism** – measures the effectiveness of the CSR communication skills.
- **Customer Satisfaction and Dissatisfaction** – survey results indicate the caller's overall satisfaction with the service provided.

Management officials at the IRS advised us they had not planned to compute the CSR LOS due to the difference in the way incoming calls were handled by the vendor. Since the CSR LOS is considered one of the most critical quantitative performance measures for determining the IRS' success, it was critical for the IRS to develop a practical method to assess the outcome of this and other vital performance information.

The vendor's sampling techniques and data compatibility issues negated the reliability of the information collected during the Test

Although the IRS computed qualitative indicators (e.g., customer accuracy, professionalism) for the tax law services provided by the vendor, the calls selected by the vendor's system for the IRS' review did not represent a statistically valid sample. Application of a standard statistical sampling methodology was required to ensure all calls had an equal opportunity for inclusion in the sample and to assure the sample results were objective.

However, the vendor could not provide reasonable assurance of the validity of the sampling method because the call-recording system used a sampling procedure that was proprietary to the manufacturer. The sampling procedure as described by the vendor indicated a valid sample of calls had not been recorded from the universe of calls received during the first 43 days of the 60-day Test period. Although a different sampling procedure was

²³ Talk time is the time a CSR spends with a caller during a transaction, hold time is the time a caller spends on hold after talking to a CSR, and wrap time is the time a CSR spends performing call-related work after the call is over.

A Test of the Private Sector’s Ability to Provide Tax Law Telephone Assistance Did Not Produce Sufficient Information to Support a Competitive Sourcing Decision

employed during the last 17 days of the Test period, there was no reasonable assurance that this method yielded a statistically valid sample.

The IRS report on the Test states, “Per the Statistics of Income Division, sample reliability remained an issue throughout the test period for a variety of reasons which impacted on the actual quality as determined during the test.” Therefore, the IRS’ quality review results cannot be considered a reliable indicator of the vendor’s performance in providing accurate answers to tax law inquiries.

Prior to the start of the Test, the IRS did not confirm whether the vendor could actually produce call detail data²⁴ in a format that met its specifications. An IRS assessment of the vendor’s performance stated the vendor did not provide accurate data in the format needed to produce performance data comparable to its own corporate performance metrics. We could not determine the cause of the call detail data incompatibilities due to conflicting information received from the IRS and the vendor.

Guidelines for planning and performing valid research studies and tests are needed

The Office of Competitive Sourcing has drafted the *IRS Guide to Competitive Sourcing* for accomplishing required tasks within the IRS Competitive Sourcing Program. In August 2004, we reported, “Better resource planning is needed before starting the competitive sourcing process.”²⁵ The report explained that the IRS needs to consider how projects should be prioritized based on resources and the likelihood of savings and improved service before nominating activities for a BCA. The report further stated, “. . . the guide does not include any procedures about nominating activities for a Business Case Analysis. Without procedures for nominating activities, the IRS may commit resources that are not available and may not be nominating the best candidates for the program.”

²⁴ Call detail data is the record of call event information that the telephone system stores for each call transaction.

²⁵ *The Office of Competitive Sourcing Needs to Improve the Planning, Scheduling, and Costing of Projects* (Reference Number 2004-10-146, dated August 2004).

A Test of the Private Sector's Ability to Provide Tax Law Telephone Assistance Did Not Produce Sufficient Information to Support a Competitive Sourcing Decision

Although the January 2004 draft of the *IRS Guide to Competitive Sourcing* describes the four phases of the IRS' competitive sourcing process, the *Guide* does not define specific methods to be employed to produce statistically valid and reliable research data. Research and data collection protocols are necessary to assure consistency and objectivity in the data collection and evaluation processes. Without these research protocols, sufficient and competent evidence is not available for informed decision making.

Challenges to performing a future competition are increased by the manner in which the private-vendor Test was administered

Unlike service providers in an actual competitive marketplace, the IRS influenced its competitor (i.e., the vendor) to duplicate the IRS' PRG method for answering tax law inquiries, which negated some of the benefits of competition. In addition, the IRS did not properly predefine its plan for evaluating the outcomes of the private-vendor Test. The IRS needs to structure any future public-private competition in a manner that promotes fairness and assures objectivity in the evaluation process to protect the integrity of the sourcing process.

The Federal Acquisition Regulation prescribes guidelines for the purchase of products and services and promotes the exercise of common sense, good judgment, and sound discretion to prevent unfairness in the acquisition process. The Commercial Activities Panel (CAP)²⁶ has adopted and recommended a set of principles to guide sourcing decisions in the Federal Government. In an April 2002 report,²⁷ the CAP stated that the use of a clear, transparent, and consistently applied sourcing process is the key to ensuring integrity and creating trust in the process. The CAP concluded accountability provides assurance that the sourcing process is efficient and effective. Accountability was characterized as requiring defined objectives, processes,

²⁶ A group of subject matter experts convened by the Comptroller General of the United States to improve the current sourcing framework and processes so they reflect a balance among taxpayer interests, Federal Government needs, employee rights, and contractor concerns.

²⁷ CAP, *Improving the Sourcing Decisions of the Government*, Final Report dated April 2002.

A Test of the Private Sector's Ability to Provide Tax Law Telephone Assistance Did Not Produce Sufficient Information to Support a Competitive Sourcing Decision

and controls with methods to track success or deviation from objectives. The CAP also described fairness as ensuring competing parties receive comparable treatment, which includes access to relevant information.

At a minimum, any future solicitation related to the Toll-Free Tax Law Telephone Operation should provide other interested vendors access to the same information and materials given to the vendor that participated in the Test. The SOW should be performance based and include relevant qualitative and quantitative performance standards with target performance values and appropriate composite scores that indicate success. Also, the IRS should exercise necessary precautions to prevent personnel with any bias regarding the outcome of the private-vendor Test from participating in the evaluation and source selection decisions on any future competition related to the Operation.

Recommendations

1. The Deputy Commissioner for Operations Support and the Commissioner, W&I Division, should rescind the September 2004 decision and eliminate use of the private-vendor Test results as justification for any decision regarding the merits of using the OMB Circular No. A-76 competitive sourcing process to determine the most efficient and cost-effective toll-free tax law telephone services provider.

Management's response: Management's response was due on July 5, 2005. As of July 6, 2005, management had not responded to the draft report.

2. The Chief, Agency-Wide Shared Services (AWSS), should expand the *IRS Guide to Competitive Sourcing* to define minimum standards for performing reliable and valid research studies and tests to prevent inefficient use of Federal Government funds in the performance of future studies/tests. These standards should require use of a structured research design model and data collection protocol that are most appropriate for meeting the objectives of the project and require creation of an evaluation plan that predefines performance standards with target performance values and composite scores

A Test of the Private Sector's Ability to Provide Tax Law Telephone Assistance Did Not Produce Sufficient Information to Support a Competitive Sourcing Decision

that constitute success. The research standards should also require involvement by the IRS Research Division in the review of the research design plan and data collection procedures, prior to the performance of any tasks, and executive-level approval of the research design plan prior to implementation.

To minimize the risk or appearance of unfairness in the competitive sourcing process related to any future study of the toll-free tax law telephone services, the Commissioner, W&I Division, and the Chief, AWSS, should:

3. Require any future Evaluation and Source Selection team members and Source Selection Official to sign a statement certifying they have no personal impairments that inhibit their ability to make a fair and impartial decision regarding the vendor proposals.
4. Ensure any future solicitation related to the Toll-Free Tax Law Telephone Operation contains no expressed or implied requirements that private vendors use current IRS methodologies in proposing solutions to the requirements in the solicitation.

A Test of the Private Sector's Ability to Provide Tax Law Telephone Assistance Did Not Produce Sufficient Information to Support a Competitive Sourcing Decision

Appendix I

Detailed Objective, Scope, and Methodology

The objective of this review was to determine whether the Internal Revenue Service (IRS) designed and administered the competitive sourcing Toll-Free Tax Law Services Test in a manner that would render a reliable assessment of whether a private vendor can provide comparable service, in terms of both access and quality, to taxpayers making tax law inquiries on the IRS toll-free telephone system. To accomplish this objective, we:

- I. Evaluated whether the framework of the private-vendor Test was aligned with the fundamental goal of the competitive sourcing initiative.
 - A. Reviewed the project activities as defined in the Statement of Work (SOW)¹ and other project documentation for consistency with Federal Government policy, procedures, and goals set for competitive sourcing in Office of Management and Budget Circular No. A-76, *Performance of Commercial Activities*; the President's Management Agenda; the Federal Acquisition Regulation;² and the Federal Acquisition Council's *Manager's Guide to Competitive Sourcing*.
- II. Assessed whether the IRS applied standard project management, business case, and/or research study methodologies in designing the private-vendor Test.
- III. Evaluated whether the private-vendor Test was implemented and performed in a manner that provided reasonable assurance the Test results would be bias free.
 - A. Assessed whether the IRS made an equitable comparison between its workforce and the vendor's staff in terms of tax law skills and experience.
 - B. Assessed whether performance measures and target values were properly established and whether valid performance data were obtained to objectively judge the success or failure of the Test.
 - C. Discussed implementation and performance issues with officials at the IRS and the participating vendor.

¹ The SOW is a description of the Federal Government's technical requirements and the work to be performed by the contractor. It specifies the terms of the contract and is the standard for measuring contractor performance.

² 48 C.F.R. pt 1-53 (2002).

A Test of the Private Sector's Ability to Provide Tax Law Telephone Assistance Did Not Produce Sufficient Information to Support a Competitive Sourcing Decision

Appendix II

Major Contributors to This Report

Richard J. Dagliolo, Acting Assistant Inspector General (Small Business and Corporate Programs)

Philip Shropshire, Director

William E. Stewart, Audit Manager

Una K. Smith, Senior Auditor

Ali Vaezazizi, Auditor

A Test of the Private Sector's Ability to Provide Tax Law Telephone Assistance Did Not Produce Sufficient Information to Support a Competitive Sourcing Decision

Appendix III

Report Distribution List

Commissioner C
Office of the Commissioner – Attn: Chief of Staff C
Deputy Commissioner for Operations Support OS
Deputy Commissioner for Services and Enforcement SE
Deputy Commissioner, Wage and Investment Division SE:W
Director, Competitive Sourcing OS:A:C
Director, Customer Account Services, Wage and Investment Division SE:W:CAS
Director, Procurement OS:A:P
Deputy Director, Procurement OS:A:P
Director, Accounts Management, Wage and Investment Division SE:W:CAS:AM
Director, Joint Operations Center, Wage and Investment Division SE:W:CAS:JOC
Chief Counsel CC
National Taxpayer Advocate TA
Director, Office of Legislative Affairs CL:LA
Director, Office of Program Evaluation and Risk Analysis RAS:O
Office of Management Controls OS:CFO:AR:M
Audit Liaisons:
 Deputy Commissioner for Operations Support OS
 Deputy Commissioner for Services and Enforcement SE
 Commissioner, Wage and Investment Division SE:W
 Chief, Agency-Wide Shared Services OS:A
 Director, Competitive Sourcing OS:A:C

A Test of the Private Sector’s Ability to Provide Tax Law Telephone Assistance Did Not Produce Sufficient Information to Support a Competitive Sourcing Decision

Appendix IV

Outcome Measures

This appendix presents detailed information on the measurable impact that our recommended corrective actions will have on tax administration. This benefit will be incorporated into our Semiannual Report to the Congress.

Type and Value of Outcome Measure:

- Inefficient Use of Resources – Actual; \$675,140 (see page 3).

Methodology Used to Measure the Reported Benefit:

This is the actual value of the contract awarded to test the viability of a commercial vendor to provide toll-free tax law telephone services. The Toll-Free Tax Law Services Test results produced no valid information that would support a decision for or against opening the Internal Revenue Service Toll-Free Tax Law Telephone Operation to competition.