



*Individual Income Tax Return
Information Was Successfully Posted
to the Latest Release of the
Customer Account Data Engine*

August 29, 2006

Reference Number: 2006-40-134

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.

Phone Number | 202-927-7037

Email Address | Bonnie.Heald@tigta.treas.gov

Web Site | <http://www.tigta.gov>



TREASURY INSPECTOR GENERAL
FOR TAX ADMINISTRATION

DEPARTMENT OF THE TREASURY

WASHINGTON, D.C. 20220

August 29, 2006

MEMORANDUM FOR COMMISSIONER, WAGE AND INVESTMENT DIVISION

FROM: (for) Michael R. Phillips
Deputy Inspector General for Audit

SUBJECT: Final Audit Report – Individual Income Tax Return Information Was
Successfully Posted to the Latest Release of the Customer Account
Data Engine (Audit # 200640018)

This report presents the results of our review to determine whether tax return information was accurately and timely posted to taxpayers' accounts on the Internal Revenue Service (IRS) Customer Account Data Engine (CADE). We also determined whether related refunds were accurately and timely issued and whether two important changes to CADE processing since our last review – accepting address changes and two additional tax return forms – were functioning properly.

Synopsis

The CADE consists of current and planned databases and related applications that will eventually replace the IRS Master File system, which is the official IRS database that stores various types of taxpayer account information. The CADE is a critical building block in the IRS' modernization program that will enable the development of subsequent modernized systems to improve customer service and compliance. The CADE is designed to post information to taxpayers' accounts on a daily rather than weekly basis, which will facilitate faster refunds to taxpayers and allow IRS employees to provide improved service to taxpayers by having up-to-date, accurate account information available.

The CADE is being implemented in a series of releases over several years. The first CADE release, Release 1.1, began posting the simplest individual tax return information in July 2004. This first release of the CADE posted only information from an Income Tax Return for Single and Joint Filers With No Dependents (Form 1040EZ). Several CADE releases with incremental changes have been implemented since then. The current release, 1.3.2, was in place for the



*Individual Income Tax Return Information
Was Successfully Posted to the Latest Release
of the Customer Account Data Engine*

2006 Filing Season¹ and was the subject of this review. Release 1.3.2 expanded CADE functionality to post information from a U.S. Individual Income Tax Return (Forms 1040 and 1040A) in addition to Forms 1040EZ. This CADE release was also able to post address changes from the incoming returns.

From January through April 2006, the IRS successfully posted 7 million tax returns to taxpayers' CADE accounts, generating \$3.3 billion in refunds. With the exception of one problem area, tax information was accurately and timely posted to taxpayers' CADE accounts, and refunds were accurately and timely issued.

We did find that some long addresses were cut off by the CADE, resulting in an incorrect address on IRS records for those taxpayers. This caused some refunds to be returned as undeliverable by the United States Postal Service. IRS records need to accurately reflect taxpayers' addresses to ensure that refunds and notices to taxpayers are mailed to the correct addresses. We advised IRS management of this issue during our review, and they took immediate action to correct CADE programming and to identify and correct affected accounts. The IRS identified approximately 4,100 taxpayer addresses that were incorrect, and we determined that 51 of these taxpayers had refunds totaling \$27,643 returned as undeliverable.

A program error caused 4,100 taxpayers' addresses to be incorrect, resulting in 51 refunds for over \$27,000 being undeliverable.

While the IRS successfully posted taxpayer information to the CADE, the CADE is still very limited in its capabilities. Many challenges remain as the IRS expands the CADE to handle more complex tax returns and issues. Although Release 1.3.2 accepted the principal individual income tax return forms, there were still many additional restrictions. For perspective, the 7 million returns posted to the CADE from January through April 2006 represented approximately 7 percent of the total individual income tax returns processed during this period.

Recommendations

We made no recommendations in this report because IRS management took appropriate corrective action to our finding during the course of the audit.

¹ The period from January through mid-April when most individual income tax returns are filed.



*Individual Income Tax Return Information
Was Successfully Posted to the Latest Release
of the Customer Account Data Engine*

Response

In an email to us, IRS management concurred with the contents of an advance copy of the draft report. Because the IRS concurred and the report had no recommendations, the IRS was not required to and did not provide a formal response.

Copies of this report are also being sent to the IRS managers affected by the report results. Please contact me at (202) 622-6510 if you have questions or Michael E. McKenney, Assistant Inspector General for Audit (Wage and Investment Income Programs), at (202) 622-5916.



*Individual Income Tax Return Information
Was Successfully Posted to the Latest Release
of the Customer Account Data Engine*

Table of Contents

Background	Page 1
Results of Review	Page 3
Tax Return Information Was Posted Timely, and Tax and Refund Amounts Were Accurate	Page 3
A Program Error Caused Some Taxpayers' Addresses to Be Incorrect.....	Page 3
There Are Many Challenges Ahead As Capabilities Are Expanded to Handle More Complex Tax Returns	Page 4
Appendices	
Appendix I – Detailed Objectives, Scope, and Methodology.....	Page 6
Appendix II – Major Contributors to This Report.....	Page 8
Appendix III – Report Distribution List	Page 9
Appendix IV – Outcome Measures.....	Page 10
Appendix V – Customer Account Data Engine Deployment.....	Page 11



*Individual Income Tax Return Information
Was Successfully Posted to the Latest Release
of the Customer Account Data Engine*

Abbreviations

CADE	Customer Account Data Engine
IDRS	Integrated Data Retrieval System
IRS	Internal Revenue Service



*Individual Income Tax Return Information
Was Successfully Posted to the Latest Release
of the Customer Account Data Engine*

Background

The Customer Account Data Engine (CADE) is the foundation for managing taxpayer accounts in the Internal Revenue Services' (IRS) modernization plan. The CADE consists of current and planned databases and related applications that will eventually replace the IRS Master File system.¹ When fully operational, the CADE will house tax information for more than 200 million individual and business taxpayers.

Tax returns, both paper and electronic, are received by the IRS at various Submission Processing sites across the country. At these sites, the return information is input to the IRS' return processing computer system. The processing system validates certain taxpayer identifying information and checks the returns for mathematical errors. After the return information has been validated and errors have been corrected, the sites send the information to the IRS Computing Centers² to be posted to the taxpayers' accounts. In the past, this account information was all posted to the Master File database, the official repository for all tax accounts. It is this database that the CADE is gradually replacing.

The age and complexity of the Master File system causes delays and inaccuracies in providing service to taxpayers. Updates to taxpayers' account information on the Master File³ occur on a weekly basis, and some updates require multiple weeks to complete. Because current data are not available to IRS employees, taxpayers may be given outdated information about their accounts. In contrast, the CADE is designed to post information to taxpayers' accounts on a daily rather than weekly basis. Taxpayers should receive refunds faster, and IRS employees should be able to provide improved service to taxpayers by having up-to-date, accurate account information available.

The first phase of the CADE, for individual taxpayer accounts, is being implemented in a series of releases over several years. The simplest taxpayer accounts are being moved to the CADE first. Each successive release moves a more complex segment of taxpayer accounts. The first CADE release, Release 1.1, began posting the simplest individual tax return information from Income Tax Return for Single and Joint Filers With No Dependents (Forms 1040EZ) during July 2004. Release 1.3.2 began posting tax return information in January 2006, and was the subject of this review. Release 1.3.2 posted information from U.S. Individual Income Tax Return (Forms 1040 and 1040A), in addition to Forms 1040EZ that were included in earlier

¹ The IRS database that stores various types of taxpayer account information. This database includes individual, business, and employee plans and exempt organizations data.

² IRS Computing Centers support tax processing and information management through a data processing and telecommunications infrastructure.

³ For example, posting return information and payments.



*Individual Income Tax Return Information
Was Successfully Posted to the Latest Release
of the Customer Account Data Engine*

releases, and also posted address changes from the tax returns. However, Release 1.3.2 is still very limited, posting only returns with a filing status of single, no dependents, a refund or even balance, no mathematical errors on the return, and no schedules.

To prepare for the 2006 Filing Season,⁴ the IRS moved eligible individual taxpayers' accounts from the Master File to the CADE, resulting in a total of 13.5 million accounts on the CADE as of January 2006. Based on previous tax year information, the IRS moved accounts likely to have returns filed that would meet CADE posting criteria. When a tax return that did not meet the criteria was filed for a CADE account, the account was moved back to the Master File and the return information posted there. Actions subsequent to tax return posting that could not be handled by the CADE, such as payments, address changes (from other than incoming tax returns), or amended returns, also caused accounts to be moved back to the Master File.

This review was performed at the CADE Project Office in New Carrollton, Maryland, and at the Fresno, California, Submission Processing Site during the period March through June 2006. The audit was conducted in accordance with *Government Auditing Standards*. Detailed information on our audit objectives, scope, and methodology is presented in Appendix I. Major contributors to the report are listed in Appendix II.

⁴ The period from January through mid-April when most individual income tax returns are filed.



*Individual Income Tax Return Information
Was Successfully Posted to the Latest Release
of the Customer Account Data Engine*

Results of Review

Tax Return Information Was Posted Timely, and Tax and Refund Amounts Were Accurate

From January through April 2006, the IRS successfully posted 7 million tax returns to taxpayers' CADE accounts, generating \$3.3 billion in refunds. Tax return information was timely posted to CADE accounts, and refunds were timely issued. The CADE accepted all three types of returns (Forms 1040, 1040A, and 1040EZ) for posting. The CADE also accepted returns with address changes and posted these changes to taxpayers' accounts.

***The IRS successfully posted
7 million tax returns to the
CADE, generating \$3.3 billion
in refunds.***

We reviewed a sample of 720 returns posted to the CADE during 2 weeks of the 2006 Filing Season to determine whether tax return information was accurately and timely posted and whether refunds were accurately and timely issued. This sample included each of the three types of returns that were posting to the CADE. The sample also included 118 returns with address changes. The dollar amounts of the items that we verified on the returns, such as tax and refund amounts,⁵ were posted accurately to taxpayers' accounts.

A Program Error Caused Some Taxpayers' Addresses to Be Incorrect

Some taxpayers' addresses were incorrect on IRS records because long street addresses were cut off by the CADE. IRS records have a limited length for storing a taxpayer's street address. When a street address exceeded this limit, the CADE incorrectly cut off the part of the address over the limit.

As tax return information is input to IRS computer systems during return processing, the street address is restricted to the length allowed in IRS records. Preparers of electronic returns have only this space to enter the street address. For paper returns, if the street address exceeds this space, IRS employees will shorten the address to fit the space. Employees will use abbreviations or other methods to shorten the address without affecting the accuracy of the address.

However, before posting tax returns with new addresses to taxpayers' accounts, the IRS uses a computer program with United States Postal Service data to refine the addresses. In some instances, this computer program expands the street address and causes it to exceed the IRS'

⁵ The return amounts verified were adjusted gross income, taxable income, tax, withholding, earned income tax credit, and refund.



*Individual Income Tax Return Information
Was Successfully Posted to the Latest Release
of the Customer Account Data Engine*

space limit.⁶ When street addresses exceeding this space limit were posted to taxpayers' CADE accounts, the address used for IRS records did not include the ending characters. The existing Master File, which the CADE is supposed to replace, was programmed to use the shorter return processing address in this situation; however, the CADE was not. In addition, the CADE programming had one less character space available for the street address than the Master File, which caused some addresses to be cut off even if they were not expanded by the address computer program.

Because addresses for some accounts handled by the CADE were cut off, IRS records showed an incorrect address for some taxpayers. IRS records need to accurately reflect taxpayers' addresses to ensure that refunds and notices to taxpayers are mailed to the correct addresses.

Management Action: We advised the IRS of this condition during our review, and management took immediate action to correct the programming and affected accounts. IRS management informed us that CADE programming was corrected to prevent this address problem in the future. The IRS identified approximately 4,100 taxpayer accounts with incorrect addresses caused by this condition. We determined that 51 of these taxpayer accounts had refunds totaling \$27,643 that were returned to the IRS as undeliverable by the United States Postal Service.

There Are Many Challenges Ahead As Capabilities Are Expanded to Handle More Complex Tax Returns

The IRS is making incremental progress in adding capabilities to the CADE. The IRS plans to continually increase these capabilities by increasing the number of accounts added to and retained by the CADE and by expanding the types of returns and issues the CADE can handle. This will allow the IRS to provide better service to more taxpayers. The IRS has expanded the scope of the current CADE release to include two additional individual income tax return forms (Form 1040 and Form 1040A) and to accept address changes on tax returns. This expansion increased the number of accounts that were eligible to be moved to the CADE.

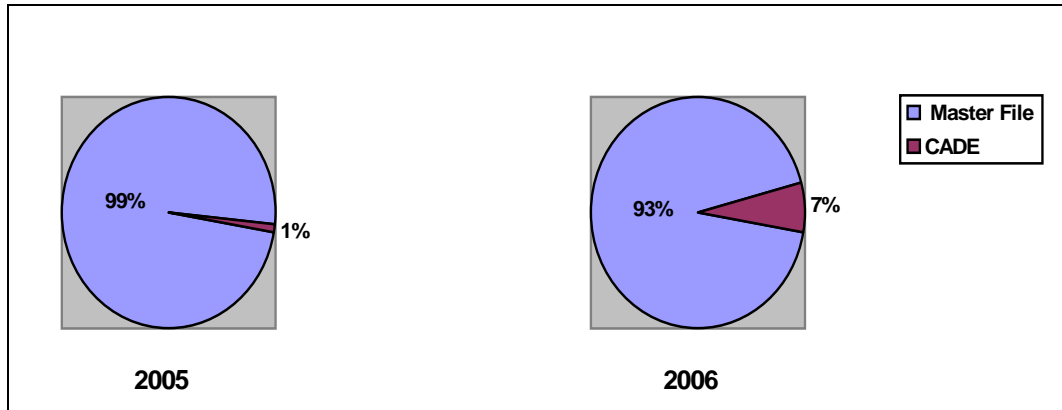
The IRS substantially increased the number of tax returns posted to the CADE with Release 1.3.2. From January through April 2005, there were 1.3 million returns posted to the CADE; this increased to 7 million in 2006. However, for perspective, a total of 107 million returns were posted during this same period in 2006. While there has been a marked increase in returns posted to the CADE, the vast majority of tax returns still are posted to the Master File (see Figure 1).

⁶ For example, the abbreviation "lndg" was expanded to "landing."



*Individual Income Tax Return Information
Was Successfully Posted to the Latest Release
of the Customer Account Data Engine*

**Figure 1: Master File and CADE Return Posting Comparison
January Through April**



Source: IRS statistics for total individual tax returns.

There are still many challenges ahead as the IRS expands the CADE capabilities to handle more complex tax returns. The current capabilities are still very limited. For example, Release 1.3.2 posted information only for tax returns with a filing status of single, no dependents, a refund or even balance, no schedules, and no mathematical errors identified during processing. In addition, actions subsequent to tax return posting, such as payments, address changes, or amended returns, caused accounts to be removed from the CADE and returned to the Master File. Of the 13.5 million taxpayer accounts on the CADE in January 2006, approximately 5.2 million (38.5 percent) were moved back to the Master File through April 2006, either because the tax returns could not be posted to the CADE, or there was subsequent activity on the account that the CADE could not process.

For the 2007 Filing Season, the IRS plans to significantly expand the scope of the CADE to include tax returns with most filing statuses, limited dependent exemptions, and various schedules for Forms 1040 and 1040A.⁷ Based on this expanded scope, the IRS again expects a considerable increase in the number of tax returns that will post to the CADE during the 2007 Filing Season. However, because recent changes to tax laws and other issues could impact the expansion of the scope of CADE, no firm volume projections were available at the time of our review.

⁷ See Appendix V.



*Individual Income Tax Return Information
Was Successfully Posted to the Latest Release
of the Customer Account Data Engine*

Appendix I

Detailed Objectives, Scope, and Methodology

The overall objective of this review was to determine whether tax return information was accurately and timely posted to taxpayers' accounts on the CADE. We also determined whether related refunds were accurately and timely issued and whether two important changes to CADE processing since our last review – accepting address changes and two additional tax return forms – were functioning properly. To accomplish our objectives, we:

- I. Determined whether tax return information was accurately and timely posted to the CADE.
 - A. Obtained computer extracts of all returns processed by the Austin and Fresno Submission Processing Sites and posted to the CADE during the weeks ending March 3 and March 17, 2006.
 - B. Selected random samples of paper and electronically filed tax returns from the returns extracted in Step I.A. We selected 90 returns of each type, for each week, from each Submission Processing site for a total of 720 returns – 360 paper returns and 360 electronically filed returns. This sampling method was sufficient for our test purposes, and we did not plan to project the results to the population. For paper tax returns, we obtained the actual tax returns; for electronically filed tax returns, we obtained a data extract of the return information.
 - C. Researched the returns in Step I.B. on the IDRS¹ and compared the IRS account information with the information on the tax returns to determine whether the returns were accurately posted. The return information verified was: name, address, adjusted gross income, taxable income, tax, withholding, earned income tax credit, and refund.
 - D. Researched the returns in Step I.B. on the IDRS to determine the number of days from the date the tax return information was sent to the Computing Centers² to the date it was posted to the CADE. We compared this to the expected time periods to determine if the tax returns were posted timely.

¹ The IRS computer system capable of retrieving or updating stored information. It works in conjunction with a taxpayer's account records.

² IRS Computing Centers support tax processing and information management through a data processing and telecommunications infrastructure.



*Individual Income Tax Return Information
Was Successfully Posted to the Latest Release
of the Customer Account Data Engine*

- E. Determined whether the data extracted in Steps I.A. and I.B. were valid. We used IRS production reports to verify that the population of returns from our computer extracts matched the number of returns processed. We performed a computer analysis of the data to determine if the data were valid. We researched selected extract records on the IDRS to verify that the extract data matched the data on the IDRS.
- II. Determined whether refunds were timely issued for CADE accounts.
- A. Researched the returns in Step I.B. on the IDRS to determine the dates they were posted. From this, we determined the expected refund dates based on standard CADE time frames for both direct deposit and paper check refunds.
 - B. Obtained an extract from the CADE of actual refund issuance dates provided to the IRS by the Department of the Treasury Financial Management Service, which issues all tax refunds. Compared the expected refund date determined in Step II.A. to the actual refund date from the Financial Management Service.
- III. Determined whether the additional functionality added to CADE Release 1.3.2 was operating as intended.
- A. Determined whether the CADE was accepting address changes on returns. During the evaluation of the accuracy of the sample returns in Step I.C., we identified accounts with address changes on the return. The accuracy of the address posting was evaluated along with the accuracy of other return items as part of Step I.C.
 - B. After identifying a problem with how the CADE handles long street addresses, we obtained data files from the IRS of the 1,934 accounts that had incorrect addresses and had been moved from the CADE to the Master File. Accounts with an undelivered refund would have been moved back to the Master File. We researched all these accounts to determine the number and dollar amount of undelivered refunds.
 - C. Determined whether the CADE was accepting all the intended individual income tax return form types. We evaluated the sample returns selected in Step I.B. to determine whether the three types of tax returns – Forms 1040, 1040A, and 1040EZ – were included.
- IV. Obtained information on the expected additions to CADE functionality for upcoming releases and on the volumes of tax returns processed, posted to the CADE, and returned to the Master File from the CADE.



*Individual Income Tax Return Information
Was Successfully Posted to the Latest Release
of the Customer Account Data Engine*

Appendix II

Major Contributors to This Report

Michael E. McKenney, Assistant Inspector General for Audit (Wage and Investment Income Programs)

Scott A. Macfarlane, Director

Richard J. Calderon, Audit Manager

Steven Stephens, Lead Auditor

Carola Gaylord, Senior Auditor

Bonnie Shanks, Auditor

Arlene Feskanich, Information Technology Specialist



*Individual Income Tax Return Information
Was Successfully Posted to the Latest Release
of the Customer Account Data Engine*

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
Chief Information Officer OS:CIO
Deputy Chief Information Officer OS:CIO
Deputy Commissioner, Wage and Investment Division SE:W
Director, Customer Account Services, Wage and Investment Division SE:W:CAS
Director, Submission Processing, Wage and Investment Division SE:W:CAS:SP
Director, Strategy and Finance, Wage and Investment Division SE:W:S
Chief, Performance Improvement, Wage and Investment Division SE:W:S:PI
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 Internal Control OS:CFO:CPIC:IC
Audit Liaisons:
 Senior Operations Advisor, Wage and Investment Division SE:W:S
 Chief, GAO/TIGTA/Legislative Implementation Branch SE:S:CLD:PSP:GTL



*Individual Income Tax Return Information
Was Successfully Posted to the Latest Release
of the Customer Account Data Engine*

Appendix IV

Outcome Measures

This appendix presents detailed information on the measurable impact that our findings will have on tax administration. These benefits will be incorporated into our Semiannual Report to Congress.

Type and Value of Outcome Measure:

- Reliability of Information – Actual; approximately 4,100 taxpayer accounts had an incorrect address caused by the CADE (see page 3).

Methodology Used to Measure the Reported Benefit:

IRS management identified the accounts with an incorrect address and provided this information to us. As part of the corrective action taken during our review, the contractors for the CADE performed computer analyses of CADE data. To identify the addresses that had been incorrectly shortened, the contractors used the individual address elements to identify taxpayer addresses with a length in excess of that allowed in the combined address field.

Type and Value of Outcome Measure:

- Taxpayer Rights and Entitlements – Actual; 51 undelivered refunds totaling \$27,643 for the accounts with incorrect addresses caused by the CADE (see page 3).

Methodology Used to Measure the Reported Benefit:

The IRS provided us with the identity of approximately 1,900 accounts with an incorrect address and a potential undelivered refund. If an account had an undelivered refund, it would be moved from the CADE back to the Master File. Therefore, the accounts with incorrect addresses that were moved back to the Master File potentially had an undelivered refund. We researched these accounts on IRS records and determined the number and amount of undelivered refunds.



*Individual Income Tax Return Information
Was Successfully Posted to the Latest Release
of the Customer Account Data Engine*

Appendix V

*Customer Account Data Engine
Deployment*

Deployment	Filing Status	Return Types Added	Expected Volumes
January 2005	Single No Dependents	Forms 1040EZ Refund or Even Balance	1.3 Million (actual)
January 2006	Single No Dependents	Forms 1040 & 1040A With No Schedules Address Changes	7 Million (actual)
January 2007 (projected)	Single Married Filing Jointly Married Filing Separately Head of Household Limited Dependents	Forms 1040 and 1040A With Various Schedules	To Be Determined

Source: Condensed from the IRS CADE Master Plan, dated October 18, 2005, and updated through discussions with the IRS.