United States General Accounting Office

GAO

Report to the Joint Committee on Taxation, U.S. Congress

September 1995

TAX ADMINISTRATION

Sole Proprietor Identification Numbers Can Be Improved

GAO/GGD-95-160

GAO

United States
General Accounting Office
Washington, D.C. 20548

General Government Division

B-261179

September 18, 1995

The Honorable Bill Archer
Chairman
The Honorable Bob Packwood
Vice Chairman
Joint Committee on Taxation
United States Congress

This report is part of work we did at your predecessors' request on the tax compliance problems of sole proprietors.[1] Self-employed individuals are referred to as sole proprietors for tax purposes. According to Internal Revenue Service (IRS) research, the sole proprietor group is among the worst with regard to the reporting of all income.

Taxpayers are required to have identification numbers so that IRS can establish accounts for them and record transactions such as the payment of taxes. Most taxpayers are required to have only one identification number. However, individuals who are self-employed are sometimes required to have two identification numbers, a Social Security Number (SSN) for their individual income tax returns and an Employer Identification Number (EIN) for their business returns. This report discusses whether IRS (1) accurately cross references the two identification numbers that self-employed individuals report and (2) needs to take any actions to improve the accuracy of its cross-reference files.

Results in Brief

IRS uses information from different computer files to identify sole proprietors who may have tax compliance problems. To better identify such proprietors, IRS requires certain taxpayers to have a valid SSN and EIN so that it can cross-reference the taxpayers' accounts from one file to another. However, the SSN and EIN cited by a sole proprietor do not always identify that same taxpayer in IRS records.

IRS records a sole proprietor's identification numbers on three computer files. It uses the SSN to establish an account on the Individual Master File and includes the EIN in the account for cross-referencing purposes. It uses the EIN to establish an account on the Business Master File and adds the SSN as a cross-reference. It uses cross-referenced SSNs and EINs from the two master files to build the Cross-Reference Entity File, which is a file IRS

[1]We issued two other reports as a result of this request: Tax Administration: IRS Can Better Pursue Noncompliant Sole Proprietors (GAO/GGD-94-175, Aug. 2, 1994) and Tax Administration: Estimates of the Tax Gap for Service Providers (GAO/GGD-95-59, Dec. 28, 1994).

created expressly to consolidate income information on sole proprietors for use in its underreporter program.

The identification numbers IRS records for cross-referencing purposes are not always reliable. From work done at the Fresno Service Center,[2] we arrived at the following estimates:

- About 20 percent of EINs posted to tax year 1991 records on the Individual Master File from Schedule C returns filed at the Fresno Service Center were erroneous. Taxpayers had filed the returns with inappropriate EINs; often the number already belonged to another taxpayer. (Sole proprietors report nonfarm business income on Schedule C returns.)
- About 3 percent of the Business Master File records of sole proprietors who filed 1991 Schedule C returns with the Fresno Service Center contained inappropriate SSNs as cross-references.
- About 10 percent of the accounts on the Cross-Reference Entity File that related to tax year 1991 returns filed with the Fresno Service Center contained erroneous cross-referenced taxpayer identification numbers.

In addition to reducing the integrity of the files, these cross-referencing errors caused IRS to generate false underreporter leads. However, we found that IRS' computerized screening process to identify inaccurate taxpayer identification numbers on the master files and on information returns limited the number of false underreporter leads that were created by the Cross-Reference Entity File. IRS data showed that 4,700 tax year 1991 potential underreporter cases were created using the Cross-Reference Entity File for the Fresno Service Center. IRS examiners manually screened out 1,582, or 34 percent, of the cases because the income had been reported on the return or the income belonged to another taxpayer. According to Fresno Service Center officials, only 0.6 percent of the cases screened out by examiners were wrong because of erroneous cross-references on the file.

Valid identification numbers could help IRS to properly cross-reference a taxpayer's various records. IRS could take action, such as earlier computer screening of EINs, to help prevent the posting of erroneous cross-references.

Another way to reduce the number of invalid identification numbers on IRS' master files would be to eliminate the requirement that sole

[2]All IRS service centers are subject to the same returns processing procedures. IRS officials said that we could expect to find the same conditions that we found at the Fresno Service Center at the other nine service centers.

proprietors use EINS. IRS requires sole proprietors to use EINS so that it can record their business returns, such as employment tax returns, on the Business Master File, which generally uses EINS as the primary account number.

Eliminating the sole proprietor's EIN would require both IRS and the Social Security Administration to modify their computer programs to accept SSNS instead of EINS. We were not able to calculate how much the reprogramming would cost. IRS and Social Security Administration officials said that, without further study, they could not state whether such a change was operationally feasible or cost-effective.

Background

Sole proprietors are to file Form 1040, Individual Income Tax Return, and attach to it Schedule C, Profit or Loss From Business (Sole Proprietorship), or Schedule F, Profit or Loss From Farming. IRS requires most sole proprietors to write their SSN on Form 1040 and to write both their SSN and EIN on Schedules C and F. However, not all sole proprietors are required to have and use EINS. IRS requires only those sole proprietors who also file employment, pension, or excise tax returns to have and use EINS.[3] IRS received about 15 million 1040 returns from sole proprietors for tax year 1991, and about 2.2 million included an EIN on a Schedule C.

The Internal Revenue Code requires taxpayers to include taxpayer identification numbers on tax returns and gives IRS the authority to decide what identification system will be used. The Social Security Administration issues SSNS to individuals. IRS issues EINS to businesses, which must apply to IRS to obtain an EIN.

IRS has two master files containing the tax accounts of individuals and businesses. Data from Form 1040 returns, including Schedules C and F, are posted to the Individual Master File using the primary SSN as the primary account number. Data from employment and excise tax returns filed by sole proprietors are posted to the Business Master File using the EIN as the primary account number. If a sole proprietor has an account on each file, IRS can connect them by cross-referencing account numbers. The proprietor's SSN is recorded in the Business Master File account, and the EIN is recorded in the proprietor's Individual Master File account.

[3]Employment, pension, and excise tax returns are filed separately from Form 1040. Employment tax returns are Form 941, Employer's Quarterly Federal Tax Return, and Form 940, Employer's Annual Federal Unemployment Tax Return; pension returns include Form 5500, Annual Return/Report of Employee Benefit Plan; and excise taxes are generally reported on Form 720, Quarterly Federal Excise Tax Return.

IRS records sole proprietor information on a third file, the Cross-Reference Entity File (CREF).[4] IRS created CREF in response to a recommendation we made in 1987.[5] We reported then that IRS could make better use of information returns in determining whether businesses filed returns and fully reported their income. Information returns are the documents that third-party payers, such as employers, corporations, and financial institutions, send to taxpayers and IRS to report such payments as wages, dividends, interest, and other types of income.

In addition to listing the taxpayer's name, information returns identify taxpayers by either SSN or EIN. IRS created CREF so that it could consolidate information returns filed under a sole proprietor's SSN with those filed under its EIN. By computer matching these consolidated data with income tax return data, IRS can determine whether sole proprietors filed tax returns and reported all income shown on information returns.

Objectives, Scope, and Methodology

As agreed with the Joint Committee, our objectives were to determine whether IRS (1) accurately cross-references sole proprietor identification numbers on the Individual Master File, Business Master File, and CREF and (2) needs to take any actions to improve the accuracy of its cross-reference files.

To accomplish these objectives, we reviewed IRS' procedures for (1) issuing EINS, (2) processing sole proprietor tax returns, and (3) entering SSNs and EINs on the two master files. We interviewed various IRS officials who were responsible for (1) updating Individual and Business Master File records, (2) creating and updating CREF, and (3) managing and operating the information-returns-based programs for identifying taxpayers who may have failed to file returns or report all income. In addition, we reviewed samples of sole proprietor accounts to determine the accuracy of SSNs and EINS posted to them.

To help us obtain samples, IRS gave us a computer file of the 1.8 million tax year 1991 Schedule Cs filed with the Fresno Service Center in calendar year 1992, the most recent year for which complete data were available when we began our work. We picked the Fresno Service Center because it

[4]CREF includes sole proprietors who, since 1986, either filed a Schedule C or F under an EIN or simply received an EIN. It also includes sole proprietors whose Business Master File account contains a cross-reference to an SSN and does not show an active business filing requirement except for employment tax returns.

[5]See The Merits of Establishing a Business Information Returns Program (GAO/T-GGD-87-4, Mar. 17, 1987).

was a logistically and geographically convenient location in which to do our work. According to IRS officials, we could expect to find similar conditions at all IRS service centers.

We analyzed the file to identify Schedule Cs that were filed with an EIN; 243,289 Schedule Cs included a unique EIN and a primary SSN or secondary SSN that was not duplicated.[6] We matched these 243,289 EINS against EINS listed on the tax year 1991 CREF, which we had obtained from IRS. This match showed that 217,279 of these Schedule C EINS were on CREF and 26,010 were not.

From the 217,279 EINS that matched, we selected a random sample of 250 accounts to determine whether the identification numbers on CREF, the Individual Master File, and the Business Master File were accurate. For each account, we reviewed the Individual and Business Master File entity and tax year 1991 records. We judged the information on the files to be accurate when the EIN and SSN on CREF matched those on the Individual and Business Master File records.

From the 26,010 Schedule C EINS that were not on CREF, we randomly selected 250 to determine whether there were valid reasons for their absence from CREF. We made this determination by searching IRS master file and tax return records for EINS in our random sample and analyzing the records when a match occurred. We also reviewed and evaluated IRS' procedures for excluding records from CREF. Our sampling methodology, which is described in more detail in appendix I, allowed us to project the results of our samples to 217,279 of the 243,289 tax year 1991 Schedule Cs that contained EINS and were filed with the Fresno Service Center.

We also reviewed the legislative history of the Code provision that authorizes the use of taxpayer identification numbers on tax returns. In addition, we discussed the need for both numbers with IRS and Social Security Administration officials.

We did our work from January 1994 to March 1995 at IRS' National Office in Washington, D.C., and its Fresno Service Center in accordance with generally accepted government auditing standards. We requested comments on a draft of this report from the Commissioner of Internal Revenue. On June 13, 1995, IRS officials provided us IRS' comments. IRS representatives at that meeting included the National Director of Submission Processing and representatives of the Assistant

[6]There were also 23,583 Schedule Cs in which more than one taxpayer used the same EIN.

Commissioners for Taxpayer Services and Collections. Their comments are summarized on pages 14-17 and incorporated in this report where appropriate.

Sole Proprietors Are the Only Taxpayers Required to Have Two Identification Numbers

Sole proprietors are the only type of business entity that has one identification number for income tax purposes and another for business tax purposes. Other business taxpayers, such as corporations, use their EINS for all tax purposes.

Sole proprietors report tax information that other individual taxpayers report, such as filing status, number of exemptions claimed, alimony paid, medical and dental expenses, interest paid on home mortgages, and taxable income. They also report tax information about their businesses, such as the accounting method used to keep the business' books, business expenses (e.g., advertising, utilities, employee wages), and cost of goods sold. They report the business information on Schedule C or F and carry to Form 1040 the net profit or loss from the sole proprietorship.

IRS records the tax information sole proprietors report on two different master files. It records the individual tax information, including Schedule C and F data, on the Individual Master File and the business tax information, such as data from employment and excise tax returns, on the Business Master File. As their names suggest, the Individual Master File holds the tax information of persons who file individual income tax returns, and the Business Master File holds the tax information submitted by sole proprietors, corporations, partnerships, and other businesses.

IRS requires sole proprietors to use two different identification numbers because each master file will accept only one type of identification number as the primary account number. The Individual Master File will accept only SSNS as the primary number for Form 1040. The Business Master File will only accept EINS as the primary account number for Form 941 employment tax returns and other business returns.[7] Therefore, in order to record tax information about them on the two files, IRS requires sole proprietors to file Form 1040 with an SSN and employment and excise tax returns with an EIN.

[7]All tax forms on the Business Master File are filed under EINs except for Form 706, Estate Tax Return, and Form 709, Gift Tax Return, which are filed under the taxpayer's SSN.

IRS Problems in Cross-Referencing Sole Proprietors' Accounts

IRS cross-references the SSN and EIN that a sole proprietor reports. However, we found that the cross-references were not always correct. Better validation of sole proprietor SSNs and EINs by IRS when processing returns could mitigate the problem but would not eliminate it.

EIN Errors on the Individual Master File

As previously mentioned, we sampled 250 of the 217,279 tax year 1991 Schedule Cs filed at the Fresno Service Center in 1992 and listed on CREF and 250 of the 26,010 returns that were not on CREF. We reviewed these cases to determine whether EINs on the Schedule Cs belonged to the taxpayers who entered them.

On the basis of our sample results, we estimate that 48,608, or 20 percent, of the 243,289 EINs posted to Individual Master File tax year 1991 records as cross-references to the Business Master File were erroneous.[8] That is, the EIN would not serve as a bridge to the sole proprietor's account on the Business Master File. Before posting, IRS did not screen EINs to detect those incorrectly reported on Schedule C.

As table 1 shows, in filing Schedule Cs, sole proprietors used EINs that either belonged to another business, did not exist on the Business Master File, or were the same as the sole proprietor's SSN.

Table 1: Estimates of Erroneous EINs on Tax Year 1991 Schedule Cs Filed at the Fresno Service Center in 1992

EIN status	Estimated number of EINs	Percent of total EINs
EIN on the Individual Master File was correct	194,681	80%
EIN belonged to another taxpayer	28,067	12%
EIN was not on the Business Master File	18,564	8%
EIN was identical to the taxpayer's SSN	1,977	1%
Total	**243,289**	**100%**

Note: Percentages do not add to 100 due to rounding.

Source: Random samples taken from tax year 1991 Form 1040 returns with Schedule Cs filed with IRS' Fresno Service Center.

[8]There were also 23,583 Schedule Cs in which more than one taxpayer used the same EIN. There were 7,423 unique EINs associated with the 23,583 schedules.

Cases involving an EIN belonging to another taxpayer included those where sole proprietors received an information return for services provided to another business and then used the EIN of that business as shown on the information return. Cases involving an EIN not on the Business Master File included those in which sole proprietors transposed digits of their EINS or otherwise entered numbers that did not reflect valid EINS.

IRS transcribed EINS from Schedule Cs but did not perform validity checks to determine whether EINS belonged to the sole proprietors who used them. As the returns are filed, they are received at service centers and go through various processing steps. As part of processing, data from the returns, including SSNS and EINS, are put into computer form. Once the data are in computer form, IRS makes certain computerized validity checks to ensure that the data are accurate. IRS checks SSNS against its computer file of valid SSNS to determine whether they are valid and attempts to correct those that are invalid before they are posted to the Individual Master File. However, the Fresno Service Center's computers were not programmed to check EINS on Schedule Cs against a computer file of valid EINS to identify those that were invalid.

Erroneous SSNs on the Business Master File

In 1987, IRS began entering sole proprietors' SSNS on the Business Master File as accounts were established. In 1992, IRS added sole proprietors' SSNS to accounts that were established before 1987. To add SSNS to pre-1987 accounts, IRS matched EINS from Schedules C and F to EINS on the Business Master File. When they matched, IRS added the sole proprietor's SSN to the account. More specifically, SSNS were added when the EINS matched; the Business Master File account did not contain an SSN; and the account holder was not a corporate, partnership, or tax-exempt return filer. The SSNS that were added were reported by sole proprietors on Schedules C and F for tax years 1987 through 1990.

IRS did not validate the Schedule C and F EINS before matching them with EINS on the Business Master File. As noted earlier, it is not uncommon for sole proprietors to use EINS that are assigned to other taxpayers. Therefore, some EINS that matched may have belonged to different taxpayers. In such instances, the SSN that was added to the account would not have been the account holder's SSN.

On the basis of our two samples of 250 EINS, we estimated that about 6,360, or 3 percent, of the 243,289 tax year 1991 Schedule Cs with EINS on Business Master File accounts that were filed with the Fresno Service

Center contained inappropriate SSNs.[9] The types of Business Master File accounts with inappropriate SSNs included the following:

- Corporation and partnership accounts that were inactive and for which returns were no longer required to be filed: SSNs were posted to these accounts because the unverified EIN on the sole proprietor's Schedule C matched the account EIN.
- Governmental agencies and churches: The match criteria did not include a routine to screen out these entities when a Schedule C EIN matched.[10]
- Spousal accounts of primary taxpayers: In this case, SSNs of primary taxpayers—that is, taxpayers that appeared first on the name line of the Form 1040 for married couples filing joint returns—were affixed to the Business Master File records of their spouses.

CREF Contained Erroneous EIN/SSN Combinations

CREF reflects those taxpayer identification numbers cross-referenced on the Individual and Business Master Files. Erroneous cross-references present on the two master files can appear on CREF as well. IRS' computer screening process prevented certain erroneous identification numbers from being posted to CREF. Even so, for the records we reviewed, CREF frequently contained EIN-SSN cross-references that were incorrect.

As reported in a previous section, we estimated that 48,608 of the EINs reported on 1991 Schedule Cs filed with the Fresno Service Center were in error (e.g., because the EIN the sole proprietor reported belonged to another taxpayer). Before data are posted to CREF, IRS subjects them to certain computerized tests to screen out "bad" data. IRS' screening process prevented 26,010 of the 48,608 errors from being posted to CREF. The screening process searched for EINs that belonged to businesses that were not sole proprietorships and EINs that matched a sole proprietor's SSN.[11]

However, on the basis of our sample of 250 EINs for the 217,279 CREF records, we estimated that 22,598 of the 48,608 erroneous EINs were posted

[9]The two samples consisted of 250 unique EINs for the 217,279 EINs that were on the CREF and 250 unique EINs for the 26,010 EINs that were not on CREF. At the 95-percent confidence level, the sampling error for the 2.6 percent was plus or minus 1.7 percent.

[10]The Business Master File contains an employment code that identifies governmental agencies, churches, and other nonprofit entities that are on the file because of the employment returns they must file.

[11]The Business Master File shows whether the account holder is an active filer. In the case of an active filer, it shows what type of return the account holder should be filing, such as corporation or partnership. The screening process for CREF uses this information to determine when the EIN the sole proprietor reported belongs to another business entity.

to CREF.[12] The screening process was not designed to catch the types of errors that the 22,598 EINs represented; namely:

- EINs affixed to Business Master File accounts of corporations and partnerships no longer required to file returns. In these situations, the computer could not determine that the account holder was not a sole proprietor because there were no active filing requirements.
- EINs from Schedule Cs that were not on the Business Master File.

The estimated 22,598 erroneous EINs represented about 10 percent of the 217,279 CREF accounts related to tax year 1991 Schedule Cs filed with the Fresno Service Center.

Erroneous Cross-References Can Lead to False Underreporter Leads

IRS collects tax information by taxpayer identification number and decides whether to contact a taxpayer for potentially underreporting income on the basis of the information collected. When taxpayer identification numbers are erroneously cross-referenced, the information generated for a taxpayer may be inaccurate and thus provide IRS examiners with false leads to pursue. IRS examiners have been instructed to screen the information manually to detect inaccuracies so that taxpayers will not be inappropriately contacted.

The problem associated with erroneous cross-referencing is highlighted by the underreporter leads that CREF generates. For tax year 1991, IRS received about 5.9 million information returns with EINs for about 1.1 million sole proprietors. Using CREF data, IRS developed a nationwide list of about 28,400 sole proprietors who may have underreported their income for tax year 1991. About 4,700 of these cases were handled by the Fresno Service Center.

These potential underreporter cases were created because the sole proprietor's tax return appeared not to have reported all of the income shown in information returns. (IRS computers match the tax return and information return data to produce the potential underreporter list.) However, Fresno Service Center examiners manually screening these 4,700 cases found that for 1,582, or 34 percent, either the income had been reported on the return or the information returns were wrong.[13] For

[12]The associated statistical confidence interval at the 95-percent confidence level for the 22,598 EINs ranges from 13,845 to 31,349.

[13]For the remaining 3,118 cases, the taxpayers were sent notices inquiring about the potential unreported income.

example, the sole proprietor may have reported interest income on the state income tax refund line of the Form 1040, or the name on the information return may not have been that of the sole proprietor.

According to Fresno Service Center officials, 10, or 0.6 percent, of the 1,582 cases screened out before taxpayer contact were caused by erroneous cross-referencing of taxpayer identification numbers. IRS had no data on how many potential underreporting cases were generated nationwide through erroneous cross-referencing of taxpayer identification numbers. IRS has no information on whether any underreporting case generated by erroneous cross-referencing slipped through the manual screening process, resulting in taxpayer contact.

Earlier Screening for Erroneous Identification Numbers Is Possible

As shown by our tests, the Individual Master File, the Business Master File, and CREF all contain inappropriate cross-references of SSNs and EINs. IRS is faced with the problem of locating and removing erroneous cross-references already on the files and preventing cross-referencing errors in the future. As discussed below, some aspects of the problem may be more readily correctable than others.

Having accurate taxpayer records is extremely important as IRS moves forward with its multibillion-dollar Tax Systems Modernization (TSM) efforts. IRS envisions that under TSM it will have an integrated case processing system in place by the year 2001, and that this system will allow all compliance issues for a taxpayer to be identified and investigated at the same time. However, IRS needs valid cross-referenced taxpayer identification numbers to properly integrate its various taxpayer records. Without accurate cross-referencing, IRS could create false noncompliance leads and have a less effective integrated case processing system than envisioned.

There are several validity checks that IRS could run to help prevent certain types of erroneous EINs from being added to sole proprietor records on the Individual Master File and CREF. One check would involve programming service center computers to identify instances where the EIN listed on Schedule C or F is the same as the sole proprietor's SSN. When the two are identical, the EIN would not be added to the Individual Master File.

The first two digits of an EIN represent an IRS district. The second validity check would involve programming service center computers to identify EINs that are invalid because the first two digits are not an IRS district code.

Our analysis of the tax year 1991 CREF showed that almost 1 percent of the 4.3 million EINs on CREF had invalid district office codes.

The third validity check would involve validating EINs in a way similar to that used by IRS to validate SSNs. IRS could computer match EINs against its file of valid EINs to identify those that are invalid. And, where appropriate, procedures could be developed to notify taxpayers that the EIN they used was invalid or wrong and to obtain the correct EIN from them.

We discussed the technical possibility and usefulness of creating these computerized screens with officials of the Fresno Service Center. They believed that there were no technical barriers to creating the screens, although there would be some cost involved in creating and testing them. They believed as well that the screens would be useful for creating accurate files. The officials did not know what the cost of creating the screens would be nor whether the screens would save examiners time in reviewing potential underreporter cases.

The erroneous cross-referenced SSNs that were incorrectly added to Business Master File records of governmental agencies and other non-income tax paying entities in 1992 are readily identifiable by computer. A computer program could be written to remove the erroneous cross-referenced SSNs from these records. However, other types of errors may not be readily identifiable by computer because they contain SSNs that were attached to Business Master File records of inactive businesses, such as defunct corporations. Therefore, IRS may have to effect these removals on a case-by-case basis, such as when underreporter examiners detect cases with erroneous cross-reference data. IRS does not currently have procedures that instruct examiners to initiate action to correct erroneous cross-references on the master files or CREF.

Two Identification Numbers May Be Unnecessary

IRS' difficulties with regard to cross-referencing a sole proprietor's two identification numbers could be eliminated if sole proprietors used a single identification number for all tax information. In addition to aiding IRS, the use of a single identification number would lessen the compliance burden that sole proprietors shoulder, which would be in keeping with IRS' long-range business plan and Congress' intent when it gave IRS the authority to require taxpayer identification numbers.

The Social Security Administration uses EINs to track employers' contributions to the Social Security trust funds. Therefore, any change in

EIN structure or use affects the Social Security Administration and its records. Social Security Administration officials with whom we talked saw no major technical obstacles for their agency if IRS eliminated the EIN requirement for sole proprietors. However, the officials said that, without further study, they could not determine whether it would be cost-effective or operationally feasible to use SSNs in place of EINs on agency computer files.

Having sole proprietors use only their SSNs for all tax purposes would be in keeping with the legislative history behind the taxpayer identification system and with the concept of the sole proprietor. Section 6109 of the Internal Revenue Code requires that taxpayer identification numbers be included on tax information and gives IRS the authority to decide what identification system will be used. Section 6109 further provides that, except as otherwise specified by regulation, the Social Security account number shall be used as the identifying number for individuals. The legislative history of section 6109 includes a request to IRS to work out an identification system that would involve the least possible burden to taxpayers.

Sole proprietors differ from other business entities in that the business does not exist separately from the owner. The sole proprietor accepts the risks of business, including tax obligations, to the extent of all of his or her assets, whether used in the business or personally. For that reason, it appears unnecessary to treat sole proprietors as two separate entities, requiring two identification numbers, for tax purposes. IRS officials said that, although sole proprietors are different from other business entities, eliminating their EIN requirement would require computer programming and possible reconfiguration changes to be made to the Business Master File. The officials did not know how much these changes would cost or whether the changes would be operationally feasible.

Conclusions

Under certain conditions, sole proprietors are required to file tax returns using two different identification numbers, a filing requirement not extended to other business taxpayers. However, sole proprietors frequently report the wrong identification number, especially in the case of the EINs that must be obtained from IRS.

For its part, IRS has not screened out all erroneous identification numbers, which means that the numbers posted to sole proprietors' records as cross-references may identify someone other than the intended taxpayer.

No data are available to discern the total effects of such mispostings; however, several false underreporter cases conducted by the Fresno Service Center were created because of erroneous cross-references. More screening is also needed if IRS is to properly integrate a taxpayer's various records under TSM.

The problems of cross-referencing identification numbers could be avoided entirely if sole proprietors could use their SSNs for all tax information. However, IRS would have to evaluate the feasibility of eliminating the EIN requirement for sole proprietors before such a change could be made.

Recommendations

We recommend that the Commissioner of Internal Revenue

- establish returns-processing and compliance-screening procedures to help remove erroneous cross-referenced taxpayer identification numbers from sole proprietors' tax records, and
- evaluate the feasibility of eliminating the requirement that sole proprietors use EINs for filing business returns.

Agency Comments and Our Evaluation

We requested comments on a draft of this report from the Commissioner of Internal Revenue. Responsible IRS officials, including the National Director of Submission Processing and representatives of the Assistant Commissioners for Collections and Taxpayer Services provided IRS' comments in a meeting on June 13, 1995. IRS' Chief Compliance Officer also provided us with comments on our draft report on July 5, 1995. These officials generally agreed that data on CREF should be perfected and said that IRS will begin evaluating how to do this. They also said that a single taxpayer identification number would facilitate reporting compliance by sole proprietors. However, they said that, due to major obstacles to implementing such a change, IRS does not agree with our recommendation to evaluate the feasibility of having sole proprietors use their SSNs instead of EINs on business returns.[14] A more detailed discussion of IRS' comments follows.

[14]In our draft report, we recommended that IRS undertake this evaluation as part of an initiative in its Business Master Plan that dealt with changing the EIN system. IRS officials said that this initiative was no longer valid, and thus we deleted references to the initiative from this report.

Cleaning Up Taxpayer Records

The officials said that while IRS' goal is to ensure accurate taxpayer account information, the cost of implementing the recommendation to clean up taxpayer records must be weighed against other competing resource demands. They said IRS must first determine the scope and impact of the problem before any decision can be made about correcting the records. IRS plans to begin evaluating the extent of the problem and to decide by May 1996 on how to address the errors in CREF.

We recognize that IRS has limited resources and needs to use those resources judiciously. At the same time, we also recognize that accurate taxpayer identification numbers are vital for an effective and efficient tax administration system. Therefore, although we are encouraged that IRS plans to clean up its taxpayer records, we do not believe that IRS has to wait until May 1996 before starting to do so. For example, IRS could begin now to institute some procedures to clean up the records. As one approach, it could have examiners who work underreporter cases correct erroneous cross-reference data when they are detected.

IRS officials stated that our suggestions to establish EIN validity checks that would identify erroneous EINs may not work. They said that, prior to 1985, taxpayers could request a specific EIN from IRS and that many taxpayers chose to use their SSNs as their EINs. They also said that, prior to 1985, if taxpayers filed business returns under their SSNs—and if there were no EINs with the same sequence of digits, and the taxpayers did not already have EINs—then IRS would assign the SSNs as EINs and process the business returns.[15] Therefore, in the view of these officials, if IRS were to adopt our validity-check suggestions, it would potentially remove valid EINs from its records.

We believe that these problems could be overcome by validating EINs in a way similar to that used to validate SSNs. IRS does this by computer matching reported SSNs against a computer file of valid SSNs. It seems that IRS could computer match EINs against its file of valid EINs to identify those that are invalid.

Allowing Sole Proprietors to Use SSNs Instead of EINs

The officials said that relieving sole proprietors of the requirement of using EINs, and permitting them to use their SSNs for all tax reporting purposes, might reduce reporting errors to some extent. However, IRS

[15]IRS officials told us that since 1985, when new computer systems were placed in service centers, each service center has been given blocks of EINs to issue to taxpayers. The officials also said that taxpayers can no longer chose their own EIN and that IRS will not process business returns in cases where taxpayers have used their SSNs instead of EINs.

officials said that IRS will not pursue such a change due to major implementation obstacles, including the necessity of

- extensively reprogramming IRS, Social Security Administration, and private sector record systems;
- massively reconfiguring the Business and Information Returns Master Files;[16]
- allocating significant IRS resources to educate taxpayers in the new requirement;
- imposing the added burden on the majority of sole proprietors who now report correctly of changing their reporting responsibilities; and finally,
- requiring sole proprietors to disclose their SSNs on Forms W-2, which raises privacy concerns.

Our draft report had already noted that computer programs and computer files would have to be changed. We did not know how extensive the changes would have to be, which is why we are recommending that IRS make an evaluation rather than recommending the actual changes. Also, we agree with IRS that taxpayers would need to be educated in the new requirement and to change their reporting responsibilities. We assumed that these factors and their costs would be part of IRS' evaluation. With regard to having sole proprietors disclose their SSNs on Forms W-2, sole proprietors who do not have EINS currently submit information returns under their SSNs.

IRS noted another obstacle concerning the use of SSNS and, in doing so, provided us with information that we were not aware of when we did our work. The officials stated that there are about 23 million pairs of EINs and SSNs on the Individual and Business Master Files that share the same sequence of digits. They said that, while some of these pairs may be due to error, others may be valid.

In our view, the problem of having EINs and SSNs with the same sequence of digits could be overcome without causing major complications. IRS currently receives information returns in which the EIN of one taxpayer has the same sequence of digits as the SSN of another taxpayer. In these situations, IRS has computer checks that allow it to associate these information returns with the proper taxpayer. Similar computer checks could be developed for business returns. The feasibility and costs of this

[16]The Information Return Master File is where INS captures the data from information returns (such as W-2s) filed by third parties. The privacy concerns associated with having all sole proprietors use their SSN rather than an EIN could also be included in this effort.

solution to the problem of having matching EINs and SSNs could be part of IRS' evaluation.

Planned IRS Study Could Address Multiple Concerns

Although IRS is opposed to our recommendation, we still believe that eliminating the EIN requirement for sole proprietors is worthy of further evaluation before a decision is made on its feasibility and cost-effectiveness, especially since in the past IRS allowed sole proprietors to use their SSNs as EINs. A similar policy for those cases where IRS has not assigned an EIN with the same digits as the sole proprietor's SSN should not involve major Business Master File reprogramming and reconfiguration. As previously discussed, IRS is proposing to do a study on the extent of the problems with CREF and ways to address them. This study could also include an evaluation of the feasibility of sole proprietors using their SSN rather than an EIN. Thus, we encourage IRS, as it proceeds with its modernization and simplification efforts, to include in its study a reconsideration of its position not to undertake an evaluation of the sole proprietor EIN requirement.

As agreed with Committee staff, we will send copies of this report to the Ranking Minority Members of the House Committee on Ways and Means and the Senate Committee on Finance, the Secretary of the Treasury, the Commissioner of Internal Revenue, and other interested parties. We will also make copies available to others upon request. Major contributors to the report are listed in appendix II. If you have any questions, please contact me at (202) 512-8633.

Lynda D Willis

Lynda D. Willis
Associate Director, Tax Policy
and Administration Issues

Contents

Abbreviations

CREF	Cross Reference Entity File
EIN	Employer Identification Number
IRS	Internal Revenue Service
SSN	Social Security Number
TSM	Tax Systems Modernization

Sampling and Data Analysis Methodology

This appendix describes our sampling approach for selecting sole proprietors with EINS to achieve a 95-percent confidence level for our estimates. Statistical sampling enables us to make estimates and draw conclusions about the universe on the basis of information in a sample of that universe. Our samples were of sole proprietors who reported EINS on tax year 1991 returns filed with the Fresno Service Center.

Sample Selection and Scope

We obtained from IRS a computer file of the 1.8 million tax year 1991 Schedule Cs filed with the Fresno Service Center in calendar year 1992. Our analysis of the file identified 266,872 Schedule Cs that were filed with an EIN. Further analysis showed that there were 23,583 Schedule Cs in which more than one taxpayer used the same EIN. There were 7,423 unique EINS associated with the 23,583 Schedule Cs. We matched the remaining 243,289 EINS against EINS on the tax year 1991 CREF, which we had obtained from IRS and which contained about 4.3 million records (i.e., EIN/SSN combinations).

This match showed that, of the 243,289 Schedule Cs with EINS, 217,279 of these EINS were on CREF and 26,010 were not. From the 217,279 EINS that matched, we selected a random sample of 250 accounts to determine whether the information on CREF, the Individual Master File, and the Business Master File was accurate. For each account, we reviewed the Individual and Business Master File entity record and tax return data. We determined that the information on the files was accurate when the EIN and SSN on CREF were the same as those on the Individual and Business Master File entity records.

Table I.1 shows the point estimates and the statistical confidence intervals at the 95-percent level of our case analysis by types of erroneous EINS listed on the Schedule Cs found on CREF.

Table I.1: Point Estimates and Confidence Intervals by the Types of Erroneous EINs Found on CREF

EIN status	Percent	Point estimate	Lower limit	Upper limit
No erroneous EINs	90%	194,681	185,930	203,434
EIN was not on the Business Master File	8%	18,252	12,675	25,266
EIN belonged to another taxpayer	2%	4,346	1,545	9,384
Total	**100%**	**217,279**	**Not applicable**	**Not applicable**

Source: Tax year 1991 Schedule Cs filed with the Fresno Service Center.

From the 26,010 Schedule C EINs not on CREF, we randomly selected 250 to determine whether there were valid reasons for their omission from CREF. We made this determination by searching IRS master file and tax return records for EINs in our random sample and analyzing the records when a match occurred. Table I.2 shows the point estimates and the statistical confidence intervals at the 95-percent level of our analysis of cases not on CREF.

Table I.2: Point Estimates and the Statistical Confidence Intervals at the 95-Percent Level for the Reasons Why EINs Were Not on CREF

Reason	Percent	Point estimate	Lower limit	Upper limit
EIN belonged to another taxpayer	91%	23,721	22,749	24,693
EIN was identical to the SSN	8%	1,977	1,337	2,794
EIN was not on the Business Master File	1%	312	70	853
Total	**100%**	**26,010**	**Not applicable**	**Not applicable**

Source: Tax year 1991 Schedule Cs filed with the Fresno Service Center.

Table I.3 shows the overall results of our analysis of EINs on tax year 1991 Schedule Cs filed at the Fresno Service Center.

Table I.3: Statistical Results of the Estimated Numbers of Erroneous EINs

EIN status	Percent	Point estimate	Lower limit	Upper limit
EIN correct	80%	194,681	185,930	203,434
EIN belonged to another taxpayer	11%	28,067	24,183	31,951
EIN not on Business Master File	8%	18,564	11,074	26,054
EIN was taxpayer's SSN	1%	1,977	1,337	2,794
Total	**100%**	**243,289**	**Not applicable**	**Not applicable**

Source: Tax year 1991 Schedule Cs filed with the Fresno Service Center.

Major Contributors to This Report

General Government Division, Washington, D.C.

Ralph Block, Assistant Director, Tax Policy and Administration Issues
Nancy Peters, Assignment Manager

San Francisco Field Office

Arthur L.Davis, Evaluator-in-Charge
Sam Scrutchins, Technical Advisor
Hans Bredfeldt, Statistician

Ordering Information

The first copy of each GAO report and testimony is free. Additional copies are $2 each. Orders should be sent to the following address, accompanied by a check or money order made out to the Superintendent of Documents, when necessary. Orders for 100 or more copies to be mailed to a single address are discounted 25 percent.

Orders by mail:

U.S. General Accounting Office
P.O. Box 6015
Gaithersburg, MD 20884-6015

or visit:

Room 1100
700 4th St. NW (corner of 4th and G Sts. NW)
U.S. General Accounting Office
Washington, DC

Orders may also be placed by calling (202) 512-6000
or by using fax number (301) 258-4066, or TDD (301) 413-0006.

Each day, GAO issues a list of newly available reports and testimony. To receive facsimile copies of the daily list or any list from the past 30 days, please call (301) 258-4097 using a touchtone phone. A recorded menu will provide information on how to obtain these lists.

For information on how to access GAO reports on the INTERNET, send an e-mail message with "info" in the body to:

info@www.gao.gov

United States
General Accounting Office
Washington, D.C. 20548-0001

Bulk Rate
Postage & Fees Paid
GAO
Permit No. G100

Official Business
Penalty for Private Use $300

Address Correction Requested