Self-Employed Nonfilers -- Who Are They? How Can We Identify Them?

Laura R. Rosage, Internal Revenue Service

The Taxpayer Compliance Measurement Program (TCMP) is a system created to measure the success of the IRS in its efforts to fulfill its mission. For this purpose, two nationwide surveys were conducted to measure nonfiling of individual income tax returns. The first TCMP individual survey was conducted for Tax Year 1979. The survey sample was limited to leads created by information returns received by the IRS. Thus, those individuals not subject to information return reporting or withholding were largely undetectable, and were not included in the estimate of 2.6 million nonfilers for Tax Year 1979. The most recent TCMP individual survey, from which the statistics of this paper are derived, is a random stratified sample of approximately 24,000 individual apparent nonfilers for Tax Year 1988. The sample was constructed by matching a list of all existing social security numbers (approximately 313 million), less the Death Master File (35 million), against all those who filed a tax return in 1988. The SSNs that matched were eliminated, leaving 100 million SSNs for which there were no matching returns. These were the potential nonfilers. The sample was drawn to investigate and determine if these potential nonfilers had income to require filing. The 100 million "potential" nonfilers were then distributed according to criteria indicative of the source and amount of available income information (see section on Explanation of Selected Terms):

Information Return Master File (IRMF) document present for 1988, income present: This category contains SSNs for which information documents, such as a W-2, *Wage and Tax Statement*, or Form 1099-INT, *Interest Income*, were on the IRMF, and income (wages, tips, interest, and the like) was recorded.

IRMF document present for 1988, income unknown or not present: This segment consists of SSNs on the IRMF due to the presence of information documents reporting transactions, such as real estate and stock sales, for which exact income from the transaction was unknown or not indicated.

No IRMF document present for 1988: This division encompasses all those SSNs which were not on the IRMF because no information returns were associated with them.

Individual Master File (IMF) entity, 1987 Return Transaction File (RTF) module: A social security number fell into this class if there was a name or address on the IMF for the entity, i.e., the individual, and there is a record of that person filing a return for 1987.

IMF entity, no 1987 RTF module: This category consists of SSNs where an IMF name and address existed on the IMF, but there was no record of a return having been filed for 1987.

No IMF entity: A social security number fell into this category when there was no name and address on the IMF associated with that number.

Category	IMF Entity 1987 RTF	IMF Entity No 1987 RTF	No IMF Entity
IRMF document Income present	3,272 Class 1	3,181 Class 2	4,453 Class 3
IRMF document No income	3,155 Class 4A	1,491 Class 5A	1,483 Class 5B
No IRMF document for 1988	3,992 Class 4B	1,225 Class 5C	1,052 Class 5D

These categories of information sources resulted in nine sampling classes. The stratified sample was designed to reflect the characteristics of all 100 million potential nonfilers.

In addition to the sample based on SSNs, the IRS had information on potential nonfilers where the SSN could not be processed (i.e., unmatched information returns or otherwise not included in the SSN sample). This information consisted of the following two categories:

Passport documents, domestic and foreign: This section of the sample consisted of persons with only passport documents and no identifying income information. There were 246 individuals in the sample in this category (Class 6).

Cases dropped from IRMF: This category is made up of individuals whose information returns were dropped from the IRMF for one of three reasons: the documents were filed late, the information returns were incorrect, or the SSN on the information return did not match the SSN of anyone on the IMF. There were 370 individuals in the sample in this category (Class 7).

This population is not necessarily distinct from the SSN sample. It does capture individuals who do not have SSNs, but it also includes those whose processed SSN was inaccurate, and therefore could not be matched against any other information on file. Because of this, some overlap was possible, but due to the small sample size, duplication was very limited.

Unable to Locate

Despite numerous sources and attempts, the IRS was unable to locate (UTL) or unable to contact approximately 4,700 of the 24,000 cases in the sample, which, when weighted, corresponds to 41 percent of the 100 million potential nonfilers (40,860,955). Almost 900 of the unweighted cases, 71 percent of the total weighted number of UTL potential nonfilers (28,960,000), fell into a residual category, where the Service had no more information (many times, less) than SSN, name, and age. An important point is that for the 15.6 percent of the SSNs in this category where the taxpayer was located for the study, only 3 percent of this group were actually delinquent in filing. Table 1 details the actual delinquency rates of the taxpayers located for all the sample categories.

In addition, a number of SSNs classified as UTL corresponded to persons with multiple SSNs or deceased individuals who were not listed on the Death Master File. As a result, many of the nonfiler leads designated UTL were generated by SSNs that were duplicative or should have been inactive because the person was deceased or because no such person existed. If all taxpayers designated UTL were located and the leads generated by erroneous SSNs eliminated, the percent of potential nonfilers who were actually delinquent in filing would likely decrease.

Background

Because self-employed have sources of income not subject to information reporting (and tax withholding), there is a concern that the self-employed who fail to file required income tax returns may escape detection. (Information reporting is the process in which employers and other financial entities report wages and other monies paid out to taxpayers on Form W-2, *Wage and Tax Statement*, Form 1099-MISC, *Miscellaneous Income*, and the like. These reports are sent to the payee for his or her records and tax return preparation and to the IRS for tax return verification.) For this paper, an individual taxpayer was considered self-employed if he or she had an Employer's Identification Number (EIN), or reported self-employment tax (on a delinquent return), or was identified as self-employed by an IRS revenue officer [1].

Table 1

Sample	UTL	Located	Actually Delinquent (of those located)
Class 1	16.1%	83.9%	27.8%
Class 2	19.0%	81.0%	24.5%
Class 3	13.8%	86.2%	4.6%
Class 4A	19.1%	80.9%	6.3%
Class 4B	20.7%	79.3%	5.0%
Class 5A	15.4%	84.6%	4.5%
Class 5B	10.6%	89.4%	1.3%
Class 5C	33.2%	66.8%	3.8%
Class 5D	84.4%	15.6%	3.1%
Class 6	22.6%	77.4%	7.4%
Class 7	24.9%	75.1%	3.7%

Disposition

Of the estimated 100 million individuals who were potential nonfilers for 1988, approximately 5 million were delinquent and required to file an income tax return. (This estimated number of nonfilers includes only those individual taxpayers who failed to file a required individual income tax return for Tax Year 1988 only [2].) More than 3 million (3,054,474) of these nonfilers had self-employment income. Since leads were generated by the presence of an SSN, the elderly, children, deceased, and unemployed were also included in the sample. Based on the sample used for this study, only 56 percent (1,697,081) of the self-employed potential nonfilers were delinquent and were required to file an income tax return. Another 10 percent of those classified as self-employed sampled were not required to file a return. In addition, there were 29 percent who had previously filed a return, 4 percent could not be found and were classified as UTL, and for 1 percent there was insufficient information for classification. For some there was an indication of self-employment before investigation (Table 1). The indicator is based in the presence of a Schedule C, *Profit or Loss from Business*, or Schedule F, *Profit or Loss from Farming*, attached to Form 1040 in previous years' records on the IRS Individual Master File. This shows that only 13 percent of the self-employed potential nonfilers who were required to file a return had filed Schedule C or Schedule F with their tax return in earlier years. Eighty-seven percent were not classified as self-employed until after the investigation (Table 2).

Table 2

Disposition	Identified as self-employed before investigation	Identified as self employed after investigation	Total
Total	291,818	2,762,656	3,054,474
Required to file	220,635	1,476,446	1,697,081
Previously filed	48,358	841,510	889,868
Not required to file	15,526	293,220	308,746
Unable to locate	6,521	117,963	124,484
Insufficient information	778	33,517	34,295

Of the estimated 1,697,081 persons who were required to file, 88 percent (1,485,544) of their returns were secured, 10 percent (176,470) were sent for further IRS examination, criminal investigation, or substitutes for returns were created; and only 2 percent (35,067) were not secured.

■ Classification of Nonfiler Cases

In normal operations, the IRS investigates individuals who have not filed a return based on the potential tax yield of the required returns, using information reported on such other returns as Form W-2, *Wage and Tax Statement* and Form 1099-INT, *Interest Income*. These cases are classified as "in the system."

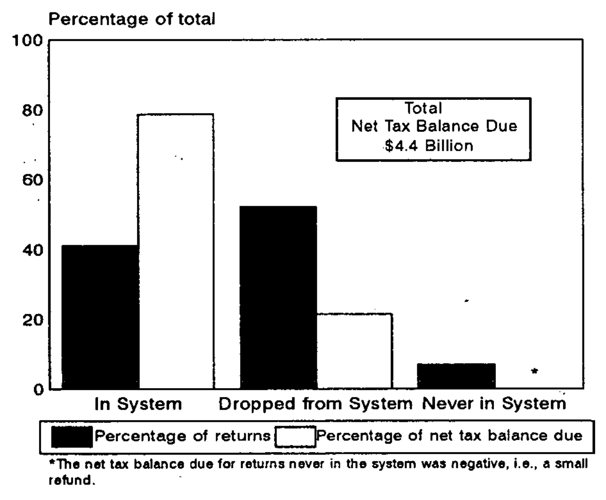
In some cases, information returns show that taxpayers have income or transactions, but have not filed a return. Because it is not feasible to investigate all these potential nonfilers, criteria have been developed to exclude from follow-up those where available information indicates a potentially lower tax yield. Those cases eliminated by the screening process are classified as "dropped from the system."

Finally, there are potential nonfilers for whom IRS has no information returns and no record of the individual on the Individual Master File. These persons are classified as "never in the system." Although the IRS does receive and investigate informants' allegations, and has special projects to detect and secure delinquent returns, returns secured in this manner are not "in the system" because the IRS was not aware of the delinquent taxpayer prior to the allegation.

For the purposes of an accurate profile of the nonfiler population, an appropriate number of cases classified as "in the system," "dropped from the system," and "never in the system" were included in the sample. The results show that 5 percent were "in the system," 49 percent were "dropped from the system," and the remaining 46 percent were "never in the system."

The criteria used to classify nonfiler cases into the three groups proved valid (Figure A). The system is designed to yield the most taxes, rather than the most returns. More than half of the self-employed nonfilers accounted for only one-fifth of the net tax due. In normal IRS operations, the low tax yield expected from these cases would cause them to be dropped from follow-up so that resources could be focused on higher tax-yielding

Figure A.--Ability of Present System to Identify Self-Employed Nonfilers



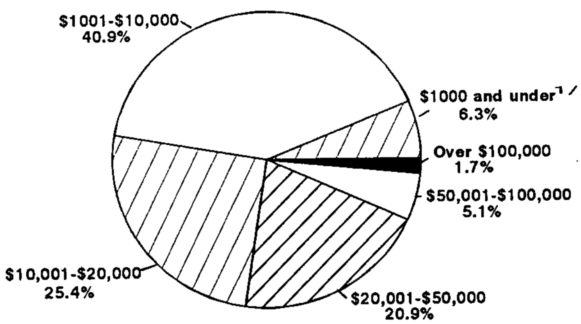
cases. More than 78 percent of the self-employed nonfiler tax balance due for 1988 was owed by those who were "in the system" (41 percent) and would be identified in normal operations. The seven percent who were "never in the system" accounted for only a small amount of tax due (actually as a group, they were due a small refund). Despite the small amount of taxes owed by these taxpayers, future efforts will need to be developed that will identify them and integrate them into the Federal tax system.

■ Characteristics of Self-Employed Nonfilers

Income

Although the average adjusted gross income (AGI) of the self-employed nonfilers was almost $19,000, AGI ranged from negative $867,000 to positive $5,200,000 (Figure B). However, the largest percentage of self-employed (41 percent) had income ranging from $1,001 to $10,000. More than 6 percent of the self-employed nonfilers had income of $1,000 or less. For slightly more than 25 percent, the range was $10,001 to $20,000, and almost 21 percent had income between $20,001 and $50,000. Less than 7 percent had income above $50,000.

Figure B.--Size of Adjusted Gross Income Self-Employed Nonfilers



1/ Includes negative AGI.

Place of Residence

Self-employed nonfilers were divided into four geographic segments based on their place of residence: urban, suburban, rural, and unknown. Results show that over four in ten reside in non-metropolitan cities or towns with populations of 50,000 or less. Roughly three in ten reside in cities with populations over 50,000, and almost one in four live in the suburbs of cities with populations over 50,000. Slightly over four percent could not be classified because there was no address known for them.

Type of Return

Of the estimated 1,697,081 self-employed nonfilers, returns were obtained through IRS enforcement contact for an estimated 1,485,544, or 88 percent. These returns were classified as "balance due" if the taxpayers owed taxes before any remittances with the returns were credited. Returns were classified as "refund" returns if prepayments (i.e., taxes withheld or estimated tax payments) and earned income credits exceeded the tax liabilities. Returns for which there was no tax liabilities or prepayments, or where prepayments exactly matched tax liabilities, were classified as "even."

Almost three out of every four returns obtained from the self-employed nonfilers showed a balance due. The statistics show that only one in five was a refund return and one in 12 was "even." If all self-employed nonfilers were identified, the distribution of balance due and refund returns would result in an estimated total net balance due of nearly $4.4 billion. The majority (78 percent) of this amount would be owed by those self-employed who are "in the system" and would be subject to enforcement activities under normal procedures.

For this study, balance due and refund amounts were further distributed as follows: $1 to $100, $101 to $500, $501 to $1,000, $1,001 to $5,000, $5,001 to $10,000, $10,001 to $20,000, $20,001 to $50,000, $50,001 to $100,000, and over $100,000. Of those required to file with balance due, the largest segment (37 percent or 450,000) was in the $1,001 to $5,000 range. Almost 28 percent (336,000) fell between $101 and $500, and nearly 15 percent (178,000) fell in the $501 to $1,000 range.

Although the number of self-employed nonfilers who would have gotten refunds was less than one-third of those with a balance due, the size distributions were somewhat similar. The largest (29 percent or 98,000) was also in the $1,001 to $5,000 class. However, almost 26 percent (87,000) were refund returns between $1 and $100 and 25 percent (85,000) were refunds in the $101 to $500 class.

An analysis of balance due and refund returns with respect to place of residence showed that over 44 percent of the balance due returns were attributable to self-employed nonfilers living in non-metropolitan areas. Thirty-three percent of the balance due returns were attributable to taxpayers residing in cities, leaving only 22 percent attributable to those residing in suburbs. The distribution is similar for refund returns. The majority of refund returns, 40 percent, were for self-employed living in non-metropolitan areas. Slightly more than 36 percent of the refund returns would have been filed by taxpayers in cities and 24 percent by persons residing in the suburbs.

Age

An estimated 10 percent of the self-employed nonfilers were age 25 and under, and only 9.0 percent were 26 to 30. The majority of self-employed nonfilers were in their prime earning years: 12.4 percent were age 31 to 35, while 9.3 percent were 36 to 40, and 15.2 percent were age 41 to 50. While nearly seven percent were age 51 to 60, only 6.5 percent were over 60 years of age. For 31 percent of the sample, age was not determinable, because the date of birth of the taxpayer was unavailable.

Reason for Not Filing

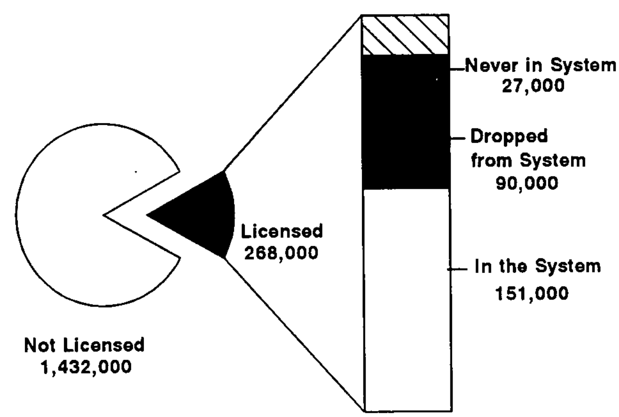
The investigating IRS revenue officers were asked to categorize the reason for which they believed the taxpayer did not file a return. A distribution of the reasons shows procrastination was the primary factor of noncompliance for more than one in four (27 percent) self-employed nonfilers. Slightly more than 12 percent did not file because of inability to pay the tax and 11 percent, because they were unaware of the filing requirement. Nearly one in 10 claimed to be missing vital information or documents, such as a Form W-2 or a Form 1099-INT. Some less frequent reasons given were misinterpretation of the tax laws, not wanting to bother to file for a small refund, forgetting to mail the return, assuming a relative or spouse had filed the return, taxpayer was deceased, and nonfiler was a tax protestor.

Occupation

Several different construction trades, together, comprise the largest (15.2 percent) occupation group for self-employed nonfilers. The specific groups within this category least likely to comply were carpenters (5.0 percent); painters, paperhangers, and plasterers (3.3 percent); and "other construction trades" (4.6 percent). The "miscellaneous occupations" category was the next largest, with 13.8 percent of the total. Within this category, occupations of those least likely to file included housewives (5.0 percent), children and students (4.0 percent), and retired and disabled persons (4.0 percent), all of whom were self-employed in some capacity. In addition to these major categories, the single occupation of self-employed individuals least likely to file was insurance, real estate, and securities sales. These persons, alone, accounted for 5.5 percent of the total number of required-to-file self-employed nonfilers.

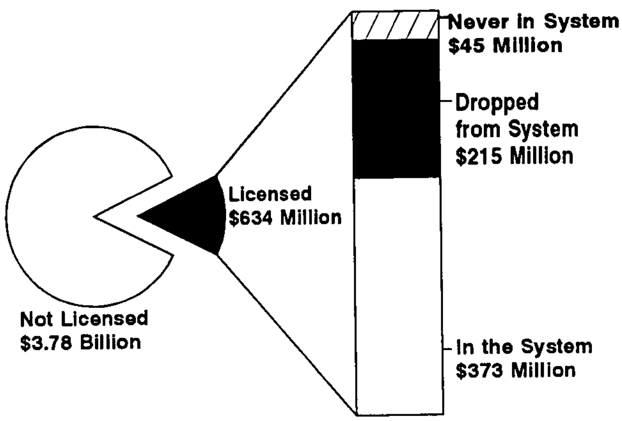
Many states or local jurisdictions require a license before a person can practice certain trades. Using listings of the licensed occupations and professions from the State of Minnesota and from Seminole and Brevard Counties in Florida (known for having strict licensing laws), 30 occupations were determined to have some licensing requirements. Based on the Minnesota and Florida experience, it is calculated that of the total estimated 1,697,000 self-employed nonfilers, only 268,000 (15.8 percent) had any sort of licensing requirement (Figure C).

Figure C.--Number of Licensed Self-Employed Nonfilers



The group of licensed taxpayers (268,000) was only responsible for $634 million, or 14.5 percent, of the $4.4 billion net tax due by required-to-file self-employed nonfilers (see Figure D). Further, 56 percent were "in the system," 34 percent were "dropped from the system," and only 10 percent were "never in the system." Those in the system with licensing requirements were responsible for $373 million of the tax due, more than half the total tax due (59 percent) by licensed self-employed persons who were required to file, and almost 9 percent of the total $4.4 billion net tax balance due. Those "dropped from the system" had only 5 percent ($215 million) of the total tax balance outstanding, and those "never in the system" were responsible for only 1 percent ($45 million).

Figure D.--Net Balance Due Licensed Self-Employed Nonfilers



Thus, current IRS procedures would ordinarily collect an estimated 59 percent of the tax due from these licensed taxpayers. If additional procedures were instituted based on licensing, another $261 million (or 6 percent) could be collected.

■ Explanation of Selected Terms

Individual Master File (IMF).--This data file, housed at the IRS National Computing Center in Martinsburg, West Virginia, contains tax accounts for individual taxpayers. Identifying information, such as name, address, and social security number, as well as tax information, such as adjusted gross income or balance due, is recorded there. The file is updated periodically with changes to taxpayer accounts.

Information Returns Master File (IRMF).--This data file contains copies of Forms W-2 and Forms 1099-INT, and any other information returns issued to taxpayers for a given year. The IRS uses the file to verify the amount of income claimed on an individual's income tax return. This file contains over one billion information returns annually.

Returns Transaction File (RTF).--This file contains taxpayer return information as transcribed directly from the income tax return. RTF files are created for each tax year. Unlike the IMF, they are not updated.

Substitute for Return (SFR).--When a taxpayer does not file a return, the IRS issues a series of reminder and delinquency notices. If the taxpayer continues to be delinquent, one option the Service can choose is to file a "Substitute for Return." Using the standard deduction and income information available from Form W-2, *Wage and Tax Statement* and other information returns, the Service computes the tax liability and penalties of the taxpayer. A bill for the outstanding amount is then issued. Because only the standard deduction is included in the calculations, for instance, the tax amount may be higher than if the taxpayer had filed directly.

■ Notes and References

[1] The cases selected for this study were thoroughly investigated by Revenue Officers during 1990 and 1991. Contrary to normal operations, cases where the estimated potential tax yield was low, or a refund, were worked as completely as those where the potential tax due was very high. At the conclusion of the investigation, Revenue Officers completed detailed questionnaires, which provide the basis of the statistics for this paper.

[2] Graeber, Michael J., Nichols, Bonnie L., and Sparrow, D. Arthur, "Characteristics of Delinquent Returns," *The IRS Research Bulletin*, Internal Revenue Service, Department of the Treasury, 1992. ■