

A PROFILE OF NON-FILERS

by

Jim Cilke

U.S. Department of the Treasury

OTA Paper 78

July, 1998

OTA Papers and Briefs are circulated so that the preliminary findings of tax research conducted by staff members and others associated with the Office of Tax Analysis may reach a wider audience. The views expressed are those of the author, and do not reflect Treasury policy. Additional copies of this publication may be purchased from the National Technical Information Service, 5285 Port Royal Road, Springfield, VA 22161.

Office of Tax Analysis
U.S. Treasury Department
Washington, DC 20220

I wish to thank several people for their valuable comments and suggestions; Lowell Dworin, Bob Gillette, Janet Holtzblatt, Karl Scholz, and Roy Wyscarver.

In October, 1996 the Office Of Tax Analysis (OTA) received the CPS-IRS exact match data file produced by the Bureau of the Census. To produce this file, the Bureau of the Census tried to exactly match information from 1990 Federal tax returns to every person on the March 1991 Current Population Survey (CPS).¹ In this paper, I describe how I used this data file to address one specific problem. I was interested in developing a "profile" of non-filers. Looking only at the population of CPS persons not legally required to file a tax return (e.g., with income below the filing threshold), I compared and contrasted the characteristics of people that did not file a tax return with the characteristics of those that did.

Being legally required to file a tax return is not the same as having an incentive to file. Even though not legally required, a person may have an incentive to file to obtain a refund of any taxes that were withheld from his wages. A person with children and earned income may have an additional incentive to file to claim a refundable earned income tax credit (EITC). Other agencies, such as a state government, may ask for information normally reported on a tax return. Finally, people may file a tax return for personal reasons, or for no apparent reason at all.

The first section of this paper discusses why a profile of the non-filing population is important. The next two sections briefly describe the exact match study and the steps I have taken to determine whether a CPS person was required to file. The last section contains the results of the comparison. This section also contains the results of a series of probit econometric equations. Among persons not required to file, the equations predict, for each person, the likelihood of being a non-filer.

Before proceeding, let me highlight some topics this paper does not address. First, this is not a paper about tax compliance (or non-compliance).² I only look at persons that, based on CPS information only, are apparently not required to file a Federal tax return. Second, this is not a paper about poverty or the characteristics of persons living in poverty in the United States. Although the people I examine generally have incomes below the poverty line, the focus of this paper is simply to compare and contrast the characteristics of low-income persons who file a return against those who do not. Third, I do not attempt to profile the entire non-filing U.S. population; I only examine those people within the CPS sampling framework who are not required to file a tax return. Thus, the population under investigation excludes the institutional population and all

persons under the age of 15 years. Fourth, I only present relative comparisons of subgroups in the population. Estimates of the actual number of non-filers is a topic of future research. Finally, this paper only provides descriptive statistics; I do not make any policy prescriptions.

WHY WE NEED INFORMATION ON NON-FILERS

Many agencies, including OTA, use data sampled from individual tax returns to describe certain characteristics of the population. Arguably, the best annual source of micro-level information on the sources and levels of income in the U.S. is the large sample of tax returns drawn by the Statistics of Income division (SOI) of the Internal Revenue Service.³ This data file forms the backbone of the individual tax simulation models used by the Treasury Department, as well as the Joint Committee on Taxation, the Congressional Budget Office, and other government and non-government agencies.

One obvious drawback to the SOI data is that the sample only covers the population who file an income tax return. However, many tax reform proposals affect individuals who do not currently file an income tax return; i.e., non-filers. For example, during

the health care reform debate, OTA examined ways of using the tax system to provide subsidies of health insurance costs for low-income individuals, including non-filers. More concretely, in 1993 Congress expanded the Earned Income Tax Credit so that some non-filers with low earnings and without children became eligible to receive a refundable credit if they filed a tax return.⁴ Further, OTA estimates the distribution of tax burdens from all Federal taxes including excise taxes on cigarettes and gasoline. Significant portions of these additional taxes are borne by non-filers.

Here then is the fundamental modeling issue: OTA uses data from tax returns for many policy analyses. However, for some analyses, tax data does not cover an important part of the population. To address this problem, OTA augments the tax data with microdata information on non-filers. The solution, adopted by OTA is to create non-filers from low-income persons on the CPS.

In OTA's current tax model based on 1989 tax data, non-filers were drawn almost entirely from CPS persons apparently not required to file a tax return. Non-filers were largely generated as the residual group after persons on the CPS were linked to returns on an SOI file using a constrained statistical match.⁵

OTA extracted from the CPS, a weighted number of "constructed tax units" that exactly equaled the tax filing population. Every CPS person not in this group became a non-filer in the tax model.

Other researchers are struggling with the same dilemma.⁶ With the availability of the CPS-IRS exact match study, OTA and other researchers may be able, in future tax models, to more-closely define the non-filing population. This paper assists this process in two ways. First, the paper provides descriptive statistics comparing low income filers to low income non-filers on the CPS. Second, the paper presents a set of probit econometric equations that may be used to impute non-filers from a CPS file.

SUMMARY OF THE CPS-IRS EXACT MATCH STUDY

The Current Population Survey (CPS), conducted by the Bureau of the Census, obtains economic, employment, and demographic information from a sample of the non-institutional population of the United States. The universe for the survey is the United States civilian non-institutional population living in housing units plus members of the Armed Forces living in civilian housing on or off a military base. The Bureau of the Census used a

probability sample to select approximately 60,000 households for the March 1991 survey. Detailed information is collected for the approximately 121,000 persons in the survey aged 15 and over.

The Bureau of the Census attempted to exactly match all persons aged 15 and over on the March 1991 CPS to tax returns on the IRS's Individual Returns Transactions File (IRTF) filed for tax year 1990. A record linkage between these two data files involved two major tasks. First, SSNs reported on the CPS file needed to be validated. Second, tax return data needed to be found and extracted for each valid SSN.

The Census Bureau worked in cooperation with the Social Security Administration (SSA) to validate reported SSNs. Each SSN as well as other person-specific information was compared to administrative data held by the SSA. Of the 121,000 persons on the CPS approximately 8,000 CPS respondents refused (or were unable) to provide an SSN. The SSNs for another 8,000 respondents did not match records from the Social Security Administration and were declared invalid. In the end, approximately 106,000 CPS records had validated SSNs. Only these CPS persons were included in the matching process.

The next step was to find and extract the tax return information for those persons with a valid SSN. The IRS provided the Census Bureau with an extract data file from the IRTF containing all 1990 tax returns filed as of the end of August 1991.⁷ A taxpayer claiming an automatic extension had until August 15, 1991 to file his 1990 tax return. Approximately 86,000 CPS persons were matched to tax returns.⁸

The Bureau of the Census then created a data file based on the results of the exact match.⁹ The file contains 23 variables for each of the 121,000 persons on the CPS file over the age of 15. One variable indicates whether or not a successful match was obtained. This variable also indicates whether a match was possible (i.e., whether the CPS person reported a valid SSN). Two variables are used to exactly link each record on the match file to the corresponding record on the full 1991 March CPS. The remaining 20 variables each contain tax information. The Census Bureau "masked" all of the tax information to remove the possibility of uniquely identifying any person on the file.¹⁰

For purposes of this exercise, the only variables used from the Exact Match file were the two file linking variables and the matching indicator variable. Further, I assumed that persons who did not report a valid SSN are randomly distributed among filers

and non-filers. I also assumed that the number of people mis-identified as a filer or as a non-filer is very small.

FORMING RETURNS AND DETERMINING FILING REQUIREMENTS

The next step is to identify, using CPS information only, those persons who were not required to file a Federal tax return. All of the analyses presented in this paper are based on this particular population. In general, a filing requirement can be determined with three pieces of information; filing status (single, joint, or head-of-house), dependency status (dependent or non-dependent), and adjusted gross income (AGI). The Office of Tax Analysis has well-established procedures for determining these three pieces of information for each person on the CPS.¹¹ These procedures attempt to follow the IRS rules as closely as possible.

Determining Filing Status

Married persons (except separated) are assigned a joint filing status. A person deemed to have a head-of-house filing status had the following characteristics. First, the person was unmarried or married but separated. Second, the person was the "reference" person in the household. In general, the housing

unit is owned or rented in this person's name. Third, the household had at least one of the following: a) had an unmarried child, grandchild, or foster child, b) a married child or grandchild that could be claimed as a dependent, or c) a non-child relative that could be claimed as a dependent. I assumed all other unmarried persons had a single filing status.¹²

Determining Dependency Status

A person could be a dependent or a non-dependent. A person was a dependent if he passed a "relationship test," a "support test," and for certain people, a "gross income test." In general, a person met the relationship test if he was a relative of the primary person or spouse in the family. The support test was met if the person's total income was less than the average total income among all family members. The gross income test was met if the person's AGI was less than the value of a personal exemption (\$2050 in 1990.) The gross income test did not apply to children under 19 years of age. Nor did the test apply to children or foster children who were going to school and were at least 19 years old.¹³

Determining Adjusted Gross Income

An estimate of individual's adjusted gross income (AGI) is simply the sum of the following income sources reported on the

CPS. For joint returns, the amounts for each spouse were combined.

- o Wages,
- o Net non-farm business income,
- o Net farm income,
- o Unemployment compensation,
- o Pensions,
- o Interest,
- o Dividends,
- o Net rental income,
- o Alimony,
- o Survivors benefits except from workers compensation,
- o Disability income except from workers compensation and from military retirement benefits,
- o Certain educational assistance,
- o Up to one half of Social Security and Railroad Retirement benefits, according to the taxation of benefits rules.

Determining a filing requirement

As a general rule a person (or joint couple) is required to file a tax return if the sum of his (their) AGI exceeds a specified threshold amount. The threshold amount is equal to the sum of the return's standard deduction plus personal exemptions for self and spouse.¹⁴ In 1990, the threshold amount for non-dependent single returns was \$5,300 plus an additional \$800 if the person was aged 65 or older. The threshold for dependent single returns was the greater of \$500 or earned income up to the non-dependent single's threshold. For joint returns the threshold was \$9,550 plus an additional \$650 for each spouse aged 65 or older. For head-of-house returns, the threshold was \$6,800 plus \$800 if aged 65 or older.

In addition to this threshold test, any person with net self-employment earnings greater than \$400 is required to file. For purposes of this study, I further assumed any person with negative total income, which could occur with business losses, was required to file.

Several important differences between tax return information and CPS data are worth noting. These differences make it impossible to exactly determine whether a low-income CPS person was required to file a tax return. First, demographic information on the CPS reflects each person's status in March of 1991 and may be legitimately different from what was reported on a 1990 tax return.

Second, the CPS does not collect information on certain types of income included in AGI. In particular, CPS does not collect realized capital gains. Further, the definitions of individual income items on the CPS may differ somewhat from the definitions used by the IRS. For example, wages on the CPS may include tax-deferred contributions to 401(k) or other retirement plans. Interest income on the CPS may include tax-exempt interest. However, such differences are likely to be negligible among low-income persons.

Third, the CPS does not collect certain pieces of information that would indicate a person is required to file a tax return, regardless of the level of his AGI. For example, the CPS does not capture whether a person had a penalty tax from an early retirement plan distribution. In addition, the tax rules for determining filing status or dependency status are much more complicated than modeled here.¹⁵

Fourth, the Bureau of the Census recognizes that income nonresponse is an important problem. So, the CPS imputes sources and amounts of income to "non-respondents." Such imputations are unlikely to exactly hit the true amount of income for any given person.

As noted earlier, even though they are not legally required, many people still file a tax return. A person with income below the filing requirement may file to obtain a refund of any taxes that were withheld from his or her wages. A person with children and earned income may have an incentive to file a return to claim a refundable earned income tax credit (EITC).¹⁶ The CPS does not indicate whether a person made estimated tax payments or had tax withheld from his or her wages. But, as I shall show, the presence of earned income substantially increases the likelihood a person will be a filer.

RESULTS

To reiterate, the population examined here are people in the CPS who: a) reported a valid SSN (or, in some cases, was the spouse of someone with a valid SSN), and b) were apparently not required to file a tax return (using CPS information only).

The population was further divided into nine unique groups:

- 1) Unmarried dependents,
- 2) Unmarried, without children, and under 62 years of age,
- 3) Unmarried, without children, and aged 62 and over,
- 4) Married, without children, and age under 62 years,
- 5) Married, without children, and aged 62 and over,
- 6) Married, with children, and age under 62 years,
- 7) Married, with children, and age 62 and over.
- 8) Unmarried, with children, and under 62 years of age,
- 9) Unmarried, with children, and aged 62 and over,

The characteristics examined are:

- a) Gender,
- b) Primary activity (in the previous week of the survey),
- c) Highest education grade completed,
- d) Head of a household status,
- e) Race class (Black, Asian+Indian+Other, Hispanic, and White),
- f) Type of living quarters (home or apartment vs. any other living quarters including mobile homes),
- g) Public or assisted housing status,
- h) Presence of means tested transfer income (general assistance including AFDC, SSI, and food stamps).
- i) Presence of non-taxable, non-means tested transfer income (veteran's benefits, workers compensation, and other non-taxable transfers except Social Security).
- j) Presence of Social Security income
- k) Presence of earned income (wages, farm, and non-farm business income),
- l) Presence of taxable unearned income (dividends, interest,

- net rent, and alimony),
- m) Presence of taxable transfers income (pensions and annuities, taxable survivors income, taxable disability income, and unemployment compensation)
- n) Presence of Adjusted Gross Income.

For purposes of this study, I assumed that incomes of married persons are shared. So, for example, a person without earned income would be classified as having earnings if his spouse had earnings. For all persons, certain household characteristics are assumed to be shared. For example, if the household receives public housing assistance, then each person in the household receives public housing assistance.

Table 1 shows three columns of data. The first column shows the percent distribution, across the nine groups, of all persons on the CPS who apparently are not required to file a tax return. The second column presents the percent distribution of actual non-filers as determined by the exact match data file. The third column shows the percent of each group who are non-filers. That is, the third column shows, for each group, the ratio of actual non-filers to the total population not required to file. For example, among unmarried dependents not required to file, 64.8% do not file a tax return.

TABLE 1

PERCENT DISTRIBUTION OF PEOPLE NOT REQUIRED TO FILE A TAX RETURN
 BY TYPE OF PERSON
 (all values in percentages)

	Persons Not Required to File		
	all people	actual non- filers	ratio of actual non-filers to total
	-----	-----	-----
Unmarried Dependents	31.1	36.4	64.8
Unmarried, No Children, Age < 62	13.2	10.4	43.7
Unmarried, No Children, Age >=62	16.0	19.5	67.3
Married, No Children, Age < 62	4.1	2.6	35.6
Married, No Children, Age >=62	16.7	12.9	43.0
Married, With Children, Age < 62	6.2	4.0	36.1
Married, With Children, Age >=62	1.3	1.2	51.3
Unmarried, With Children, Age < 62	8.8	9.9	62.2
Unmarried, With Children, Age >=62	2.5	3.1	69.7
All Persons	100.0	100.0	55.5

Table 2 compares, for several assorted demographic and economic characteristics, actual non-filers with the total population not required to file. Each row in Table 2 classifies a person by a particular demographic or economic characteristic. Table 2 shows two columns of data for each of the nine groups under investigation. The first column shows the percent of actual non-filers with a particular attribute; the second column shows the same statistic as it applies to all persons not required to file. By comparing these two columns, one can see how actual non-filers compare to persons not required to file a return.

Many of the results of this study are not surprising. Some highlights of important results seen in Tables 1 and 2 are as follows (in bullet form).

Results relating to population subgroups.

- S** The largest group of non-filers, comprising 36.4% of the non-filing population, are single dependents. Not surprisingly, most people in this class are going to school. Recall that the stringent gross income test does not apply persons age of 19 and over who going to school.
- S** The second and third largest classes of non-filers are unmarried and married persons without children and aged 62 and over. These groups, respectfully accounting for 19.5% and 12.9% of the population of non-filers, can largely be characterized as retired or semi-retired persons living primarily on Social Security income. Interestingly, almost 79% of the unmarried non-filers are female.

- S The next largest class of non-filers, accounting for 10.4% of the non-filing population, are unmarried persons without children and under 62 years of age. This is probably the most difficult group to characterize. The group is fairly evenly split between male and female. About 60% of this group have no or little taxable income and about 32% have some earned income. About 53% are the primary person in the household. Most have at least a high-school education.
- S Unmarried persons with children comprise 13.0% of the non-filing population. An overwhelming majority of these persons are female; 94% of persons under 62 years of age and 86% of persons aged 62 and over are female.
- S As seen in column 3 of Table 1, married people who are not required to file are more likely to file a return than unmarried people. Similarly, non-dependents under the age of 62 are more likely to file a return than non-dependents aged 62 and over.

Results relating to sources of income included in AGI

- S The presence of earned income increases the likelihood of filing a return. Table 2 shows that 33.5% of all persons in the examined population have some earnings. However, among non-filing persons, only 17.3% have earnings. The presence of earnings is often associated with wage withholding. Interestingly, this approximate 2-to-1 ratio applies to people with children as well as people without children.
- S Although not as dramatic, the presence of unearned income included in AGI also increases the likelihood of filing a return. Here, 32% of all examined persons have some taxable unearned income, while 26% of non-filing returns have unearned income.
- S Following a similar pattern, the presence of taxable transfer income (pensions and annuities + unemployment + certain types of disability and survivors income) included in AGI slightly increases the likelihood of filing a return. Here, 14.5% of all examined persons have some taxable transfer, while just over 10% of non-filers have a taxable transfer.

Results relating to sources of government transfer income.

- S People with means-tested government transfer income (food stamps, SSI, or AFDC) are less likely to file a return. Here, nearly 25% of persons not required to file have some means-tested transfer income. However, 32% of non-filers have some means-tested transfer. Among unmarried persons with children and under 62 years old, the corresponding percentages are 73% and 83%. The corresponding percentages for unmarried persons without children and under 62 years of age are 30% and 50%.
- S Interestingly, the pattern is reversed for persons with selected non-taxable non-means tested transfers (veteran's benefits, workers compensation, and other benefits.) Non-filers are slightly less likely to receive these types of benefits.

Results relating to gender, education, and race

- S An overwhelming majority of unmarried persons with children who are not required to file are female. Further, the percent of filers who are female is roughly the same as the percent of non-filers who are female.
- S Interestingly, among married persons without children and under 62 years, 62.5% are female. Correspondingly, among married without children and over 62, 47.5% are female. The obvious explanation is that many females are married to an older, retired male. (Obviously, the number of married male and female non-filers must be equal.)
- S Persons with an education level at or below the 10th grade are less likely to file. Persons with some college education are likely to be filers.
- S Non-filers who are household heads occur at approximately the same percentage as all persons below the filing threshold who are household heads.
- S Minorities are somewhat less likely to file than whites.

TABLE 02 (continued)
 PERCENT OF ACTUAL NON-FILERS AND PERSONS NOT REQUIRED TO FILE:
 BY ASSORTED DEMOGRAPHIC AND ECONOMIC CHARACTERISTICS
 BY TYPE OF PERSON

	DEPENDENTS ONLY		UNMARRIED NO KIDS, AGE < 62		UNMARRIED NO KIDS, AGE >=62		MARRIED NO KIDS, AGE < 62		MARRIED NO KIDS, AGE >=62	
	NON- FILERS (NO IRTF MATCH)	ALL NOT REQUIRED TO FILE	NON- FILERS (NO IRTF MATCH)	ALL NOT REQUIRED TO FILE	NON- FILERS (NO IRTF MATCH)	ALL NOT REQUIRED TO FILE	NON- FILERS (NO IRTF MATCH)	ALL NOT REQUIRED TO FILE	NON- FILERS (NO IRTF MATCH)	ALL NOT REQUIRED TO FILE
	LIVING QUARTERS									
HOUSE OR APARTMENT	94.4	95.1	88.3	91.6	91.4	92.5	85.3	86.9	89.6	91.6
OTHER	5.6	4.9	11.7	8.4	8.6	7.5	14.7	13.1	10.4	8.4
PUBLIC HOUSING ASSISTANCE										
NO ASSISTANCE	94.7	95.6	83.9	90.1	78.6	83.8	90.4	94.5	94.3	96.8
WITH ASSISTANCE	5.3	4.4	16.1	9.9	21.4	16.2	9.6	5.5	5.7	3.2
MEANS TESTED TRANSFERS										
NO TRANSFERS	80.0	83.9	50.0	69.5	75.0	81.9	59.3	76.9	86.5	92.8
WITH TRANSFERS	20.0	16.1	50.0	30.5	25.0	18.1	40.7	23.1	13.5	7.2
NON-MEANS TESTED TRANSFERS										
NO TRANSFERS	98.5	98.3	86.4	88.5	92.1	92.8	78.0	81.0	88.6	89.2
WITH TRANSFERS	1.5	1.7	13.6	11.5	7.9	7.2	22.0	19.0	11.4	10.8
SOCIAL SECURITY BENEFITS										
NO SOCIAL SECURITY	90.4	92.4	73.4	84.2	6.0	5.0	44.7	54.9	3.7	3.1
WITH SOCIAL SECURITY	9.6	7.6	26.6	15.8	94.0	95.0	55.3	45.1	96.3	96.9
COMPONENTS OF AGI										
NO EARNED INCOME	78.4	63.3	68.5	38.3	95.9	93.0	74.1	51.1	93.3	85.3
WITH EARNED INCOME	21.6	36.7	31.5	61.7	4.1	7.0	25.9	48.9	6.7	14.7
NO UNEARNED INCOME	87.4	89.2	82.0	72.9	54.2	45.5	77.6	59.3	45.1	34.2
WITH UNEARNED INCOME	12.6	10.8	18.0	27.1	45.8	54.5	22.4	40.7	54.9	65.8
NO TRANSFERS IN AGI	99.9	99.9	96.0	94.9	82.6	77.6	87.7	77.6	64.4	56.9
WITH TRANSFERS IN AGI	0.1	0.1	4.0	5.1	17.4	22.4	12.3	22.4	35.6	43.1
NO AGI	68.7	55.4	54.2	28.0	44.9	34.9	52.2	23.5	31.4	18.8
WITH AGI	31.3	44.6	45.8	72.0	55.1	65.1	47.8	76.5	68.6	81.2
ADDENDUM AGI < \$100	77.3	61.9	59.7	30.9	53.7	42.0	56.6	27.9	36.2	22.3

TABLE 2 (continued)
 PERCENT OF ACTUAL NON-FILERS AND PERSONS NOT REQUIRED TO FILE:
 BY ASSORTED DEMOGRAPHIC AND ECONOMIC CHARACTERISTICS
 BY TYPE OF PERSON

	MARRIED W KIDS, AGE < 62		MARRIED W KIDS, AGE >=62		UNMARRIED W KIDS, AGE < 62		UNMARRIED W KIDS, AGE >= 62		ALL RETURNS	
	NON- FILERS (NO IRTF MATCH)	ALL NOT REQUIRED TO FILE	NON- FILERS (NO IRTF MATCH)	ALL NOT REQUIRED TO FILE	NON- FILERS (NO IRTF MATCH)	ALL NOT REQUIRED TO FILE	NON- FILERS (NO IRTF MATCH)	ALL NOT REQUIRED TO FILE	NON- FILERS (NO IRTF MATCH)	ALL NOT REQUIRED TO FILE
GENDER										
MALE	45.1	45.9	61.0	60.2	5.6	7.0	14.1	17.8	39.5	41.4
FEMALE	54.9	54.1	39.0	39.8	94.4	93.0	85.9	82.2	60.5	58.6
ACTIVITY LAST WEEK										
WORKING	15.7	30.2	3.3	4.3	8.5	21.7	1.5	4.9	5.7	13.0
WITH JOB NOT AT WORK	0.6	1.3	0.5	0.3	0.1	0.8	0.0	0.0	0.2	0.5
LOOKING FOR WORK	5.9	5.2	0.0	0.2	2.8	2.6	0.0	0.0	2.0	2.3
KEEPING HOUSE	40.9	36.4	27.9	29.2	64.7	52.9	45.9	40.9	28.3	24.9
GOING TO SCHOOL	7.7	5.2	0.0	0.0	8.2	8.0	0.0	0.0	26.7	24.2
UNABLE TO WORK	10.6	6.6	6.1	4.4	5.6	4.1	13.6	10.4	7.8	5.4
RETIRED	2.4	1.5	49.3	51.0	0.4	0.4	28.7	34.7	17.3	18.6
OTHER	16.1	13.6	12.8	10.6	9.8	9.4	10.3	9.0	12.1	11.1
EDUCATION LEVEL										
EDUC <= 10TH GRADE	41.3	38.4	69.9	65.3	38.6	32.7	69.2	60.3	58.0	46.0
EDUC 11TH OR 12TH GRADE	44.4	44.8	26.7	28.8	47.6	50.8	27.3	31.6	32.1	38.1
EDUC 1-3 YRS COL	11.2	11.6	0.5	3.3	11.9	14.0	2.2	4.7	7.6	11.9
EDUC 4YR COLLAGE+	3.1	5.3	2.9	2.6	1.8	2.5	1.3	3.5	2.3	4.0
HOUSEHOLD STATUS										
HOUSEHOLD HEAD	96.4	95.2	100.0	100.0	90.3	89.1	100.0	99.9	54.3	57.0
HOUSEHOLD MEMBER	3.6	4.8	0.0	0.0	9.7	10.9	0.0	0.1	45.7	43.0
RACE										
BLACK	17.0	14.9	31.8	24.9	44.1	40.5	35.0	28.5	22.3	18.3
ASIAN, INDIAN, OTHER	8.4	7.8	3.8	5.0	3.4	3.1	2.7	2.5	4.2	3.8
HISPANIC	18.4	20.5	9.9	11.5	12.8	11.7	8.6	7.4	9.5	8.5
WHITE	56.1	56.8	54.5	58.5	39.6	44.7	53.8	61.7	64.1	69.4

TABLE 2 (continued)
 PERCENT OF ACTUAL NON-FILERS AND PERSONS NOT REQUIRED TO FILE:
 BY ASSORTED DEMOGRAPHIC AND ECONOMIC CHARACTERISTICS
 BY TYPE OF PERSON

	MARRIED W KIDS, ALL AGES		MARRIED W KIDS, ALL AGES		UNMARRIED W KIDS, AGE < 62		UNMARRIED W KIDS, AGE >= 62		ALL RETURNS	
	NON- FILERS (NO IRTF MATCH)	ALL NOT REQUIRED TO FILE	NON- FILERS (NO IRTF MATCH)	ALL NOT REQUIRED TO FILE	NON- FILERS (NO IRTF MATCH)	ALL NOT REQUIRED TO FILE	NON- FILERS (NO IRTF MATCH)	ALL NOT REQUIRED TO FILE	NON- FILERS (NO IRTF MATCH)	ALL NOT REQUIRED TO FILE
LIVING QUARTERS										
HOUSE OR APARTMENT	89.9	89.3	92.7	94.1	93.2	92.3	96.3	95.9	92.1	92.7
OTHER	10.1	10.7	7.3	5.9	6.8	7.7	3.7	4.1	7.9	7.3
PUBLIC HOUSING ASSISTANCE										
NO ASSISTANCE	86.9	89.0	97.3	97.9	64.2	68.7	93.4	95.2	87.0	90.4
WITH ASSISTANCE	13.1	11.0	2.7	2.1	35.8	31.3	6.6	4.8	13.0	9.6
MEANS TESTED TRANSFERS										
NO TRANSFERS	35.5	46.9	66.9	77.9	17.7	26.6	68.7	75.6	67.8	75.2
WITH TRANSFERS	64.5	53.1	33.1	22.1	82.3	73.4	31.3	24.4	32.2	24.8
NON-MEANS TESTED TRANSFERS										
NO TRANSFERS	80.9	82.0	82.2	83.2	93.4	91.9	94.3	94.0	92.6	92.0
WITH TRANSFERS	19.1	18.0	17.8	16.8	6.6	8.1	5.7	6.0	7.4	8.0
SOCIAL SECURITY BENEFITS										
NO SOCIAL SECURITY	76.0	83.4	7.2	6.7	87.8	89.0	9.6	8.5	55.4	56.8
WITH SOCIAL SECURITY	24.0	16.6	92.8	93.3	12.2	11.0	90.4	91.5	44.6	43.2
COMPONENTS OF AGI										
NO EARNED INCOME	61.1	31.5	91.6	80.0	79.8	58.8	95.8	90.7	82.7	66.5
WITH EARNED INCOME	38.9	68.5	8.4	20.0	20.2	41.2	4.2	9.3	17.3	33.5
NO UNEARNED INCOME	87.7	80.9	70.6	66.4	90.6	86.8	72.0	62.5	74.3	67.9
WITH UNEARNED INCOME	12.3	19.1	29.4	33.6	8.4	13.2	28.0	37.5	25.7	32.1
NO TRANSFERS IN AGI	94.4	89.5	63.0	60.2	97.6	95.8	84.0	78.0	89.8	85.5
WITH TRANSFERS IN AGI	5.6	10.5	37.0	39.8	2.4	4.2	16.0	22.0	10.2	14.5
NO AGI	50.3	23.6	47.3	35.8	72.6	52.2	59.5	48.1	56.4	38.4
WITH AGI	49.7	76.4	52.7	64.2	27.4	47.8	40.5	51.9	43.6	61.6
ADDENDUM AGI < \$100	53.7	26.0	51.3	39.5	77.1	55.6	66.9	54.7	63.4	43.3

THE PROBIT EQUATION

One of the objectives of this paper is to provide a methodology for obtaining microdata information about non-filers that could be used in tax modeling. The approach taken here is to estimate a set of econometric probit equations that predict whether a person with income below the filing threshold is likely to be a non-filer. I ran separate, but identically specified, equations for each of the nine groups of people. (The appendix at the end of this paper shows the definition of each of the variables used in the probit equations.) To obtain consistent results for married couples, I ran the probit equation on only one spouse.¹⁷

Almost all of the independent variables used in the equations are dummy variables. The only continuous variable is AGI as a percent of the filing threshold. Note that I eliminated some variables from some equations because some values of certain variables were particularly sparse. For example, hardly any dependents had AFDC benefits. A parameter value of 0.000 shown in Table 3 indicates that either every observation had the same value for that variable or that I forced every observation to have the same value.

The results of the equations are seen in Table 3. A positive value for a parameter means the presence of that characteristic increases the likelihood that person will be a non-filer. For example, the parameters for people without earned income are positive for all subgroups. The results of the probit equations seen in Table 3 largely mirror the descriptive statistics shown in Table 2.

As a test, I applied the probit equations to the same set of observations used to create the equations. This produced, for each observation, a cumulative probability estimate. I then predicted whether each observation would be a non-filer based on whether the person's probability estimate was below the probability of being a non-filer as shown in column 3 of Table 1. For example, if the cumulative probability estimate for a dependent person was less than 64.8%, then I predicted this person would be a non-filer. Similarly, I predicted that married persons under 62 years of age and with children who had a probability estimate less than 36.1% were non-filers. This led to a 2-by-2 measure-of-association table (Table 4). The rows in Table 4 show the actual counts of filers and non-filers in the sample, while the columns show the predicted counts of filers and non-filers.¹⁸ Table 4 shows that the probit equations correctly predict whether a person is a filer or non-filer 76% of the time.

Further, the total predicted number of non-filers from the sample (12,973), is approximately equal to the actual number of non-filers in the sample (13,725).

TABLE 3

ESTIMATES FROM PROBIT EQUATIONS PREDICTING PERSONS LIKELY TO BE NON-FILERS
BY TYPE OF PERSON

	DEPENDENTS ONLY		UNMARRIED NO KIDS, AGE < 62		UNMARRIED NO KIDS, AGE >=62		MARRIED NO KIDS, AGE < 62		MARRIED NO KIDS, AGE >=62	
	PROBIT ESTIMATE	STD ERROR	PROBIT ESTIMATE	STD ERROR	PROBIT ESTIMATE	STD ERROR	PROBIT ESTIMATE	STD ERROR	PROBIT ESTIMATE	STD ERROR
	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
INTERCEPT	0.7323	0.6279	0.4890	0.2830	1.7772	0.2648	0.9710	0.7372	1.9457	0.3863
GENDER										
MALE	-0.0070	0.0334	0.0268	0.0533	-0.0635	0.0604	0.0000	0.0000	0.0000	0.0000
EDUCATION LEVEL										
<= 10TH GRADE	1.0964	0.1098	0.5897	0.1130	0.5789	0.1118	-0.5424	0.2747	0.9191	0.1452
11TH OR 12TH GRADE	0.5900	0.1107	0.2417	0.1049	0.2234	0.1125	-0.4140	0.2669	0.4561	0.1476
1-3 YRS COL	0.1746	0.1176	0.0453	0.1087	0.0898	0.1307	-0.7199	0.3419	0.3028	0.1731
HOUSEHOLD STATUS										
HEAD	0.0000	0.0000	0.1094	0.0577	-0.2095	0.0723	0.0000	0.0000	0.0000	0.0000
RACE										
BLACK	0.1228	0.0504	0.2175	0.0730	0.3121	0.0929	0.5764	0.2441	0.0685	0.1171
ASIAN, INDIAN	0.1470	0.0694	0.2070	0.1154	0.1440	0.1907	0.6184	0.2927	0.0206	0.2244
HISPANIC	0.2532	0.0506	0.0586	0.0800	0.1045	0.1151	0.8322	0.2165	-0.0682	0.1219
LIVING QUARTERS										
HOUSE OR APT	-0.2365	0.0798	-0.2663	0.0904	-0.3477	0.0905	-0.1552	0.2070	-0.3358	0.1047
ACTIVITY LAST WEEK										
IN LABOR FORCE	-0.4220	0.0714	-0.0961	0.0801	-0.0742	0.1611	-0.5073	0.2154	-0.3210	0.1998
KEEPING HOUSE OR IN SCHOOL	-0.1307	0.0579	-0.0056	0.0794	0.1144	0.0805	0.0135	0.2185	0.0533	0.1573
UNABLE TO WORK	0.6875	0.1669	0.2595	0.1203	0.2089	0.1302	0.4684	0.2601	0.0535	0.1480
RETIRED	-0.1420	0.1352	-0.5181	0.2016	-0.0758	0.0785	-0.4119	0.2355	-0.2041	0.0900
PRESENCE OF EARNED INCOME										
NO EARNED INCOME	0.5069	0.0449	0.5297	0.0815	0.4147	0.1053	0.4388	0.2019	0.4424	0.0938
PRESENCE OF UNEARNED INCOME										
NO UNEARNED INCOME	-0.0920	0.0541	0.2055	0.0626	0.1397	0.0535	0.4733	0.1658	0.1191	0.0631
PRESENCE OF TAXABLE TRANSFERS										
NO TAXABLE TRANSFERS	0.1280	0.5743	0.0968	0.1159	-0.0266	0.0599	0.1371	0.2284	-0.2047	0.0670
PUBLIC HOUSING ASSISTANCE										
NO PUBLIC HOUSING	0.0482	0.0864	-0.1033	0.0918	-0.6259	0.0740	0.3421	0.3425	-0.6428	0.1700

TABLE 3 (continued)

ESTIMATES FROM PROBIT EQUATIONS PREDICTING PERSONS LIKELY TO BE NON-FILERS
BY TYPE OF PERSON

	DEPENDENTS ONLY		UNMARRIED NO KIDS, AGE < 62		UNMARRIED NO KIDS, AGE >=62		MARRIED NO KIDS, AGE < 62		MARRIED NO KIDS, AGE >=62	
	PROBIT ESTIMATE	STD ERROR	PROBIT ESTIMATE	STD ERROR	PROBIT ESTIMATE	STD ERROR	PROBIT ESTIMATE	STD ERROR	PROBIT ESTIMATE	STD ERROR
PRESENCE OF FOOD STAMPS										
NO FOOD STAMPS	-0.2088	0.0569	-0.1031	0.0734	-0.4368	0.1118	-0.3034	0.2114	-0.9412	0.2051
PRESENCE OF SOCIAL SECURITY										
NO SOCIAL SECURITY	-0.0588	0.0730	-0.1155	0.0813	-0.2043	0.1208	-0.1780	0.1791	-0.1680	0.1666
PRESENCE OF SSI										
NO SSI	-0.4089	0.1233	-0.4346	0.0945	-0.5456	0.1098	-0.6743	0.2551	-0.6727	0.1700
PRESENCE OF AFDC										
NO AFDC	0.0000	0.0000	-0.2119	0.1074	0.0000	0.0000	-0.7647	0.3177	0.0000	0.0000
PRESENCE OF "OTHER" BENEFITS										
NO "OTHER" BENEFITS	-0.3840	0.1083	0.1185	0.0807	0.1133	0.0879	0.1679	0.1659	0.0182	0.0881
AGI DIVIDED BY FILING THRESHOLD	-2.0285	0.1374	-1.2196	0.1134	-1.1628	0.0944	-0.8699	0.2805	-1.4177	0.1194
SAMPLE SIZE	8469		3544		4379		462		2590	

TABLE 3 (Continued)

ESTIMATES FROM PROBIT EQUATIONS PREDICTING PERSONS LIKELY TO BE NON-FILERS
BY TYPE OF PERSON

	MARRIED W KIDS, AGE < 62		MARRIED W KIDS, AGE >= 62		UNMARRIED W KIDS, AGE < 62		UNMARRIED W KIDS, AGE >= 62	
	PROBIT ESTIMATE	STD ERROR	PROBIT ESTIMATE	STD ERROR	PROBIT ESTIMATE	STD ERROR	PROBIT ESTIMATE	STD ERROR
INTERCEPT	-0.3651	0.4599	4.8799	1.4199	0.2442	0.3705	0.1617	0.8032
GENDER								
MALE	0.0000	0.0000	0.0000	0.0000	0.1672	0.1193	-0.4765	0.1692
EDUCATION LEVEL								
<= 10TH GRADE	0.1664	0.2146	-0.5725	0.5891	0.3784	0.1813	0.8919	0.3389
11TH OR 12TH GRADE	0.2584	0.2092	-0.5808	0.5904	0.1594	0.1758	0.4659	0.3408
1-3 YRS COL	0.1219	0.2402	-2.4300	1.1376	0.1689	0.1862	0.0943	0.4148
HOUSEHOLD STATUS								
HEAD	0.0000	0.0000	0.0000	0.0000	-0.0690	0.1033	0.0000	0.0000
RACE								
BLACK	0.1702	0.1557	0.3390	0.2722	0.0321	0.0737	0.5219	0.1700
ASIAN, INDIAN	0.1579	0.1775	-0.7021	0.4189	0.0876	0.1456	0.6209	0.3974
HISPANIC	0.1091	0.1192	0.0429	0.2908	-0.0714	0.0907	0.3161	0.2246
LIVING QUARTERS								
HOUSE OR APT	-0.0495	0.1506	0.0995	0.4362	0.1327	0.1192	-0.2261	0.3075
ACTIVITY LAST WEEK								
IN LABOR FORCE	0.0395	0.1364	0.5470	0.5628	-0.2528	0.1132	-0.6626	0.4453
KEEPING HOUSE OR IN SCHOOL	0.1641	0.1430	-0.6947	0.5050	0.0842	0.1013	-0.2316	0.2460
UNABLE TO WORK	0.0458	0.1928	0.1920	0.4932	0.1216	0.1962	-0.0070	0.3346
RETIRED	0.1335	0.3201	-0.3255	0.3555	-0.5402	0.4261	-0.7095	0.2445
PRESENCE OF EARNED INCOME								
NO EARNED INCOME	0.8002	0.1425	0.7567	0.3273	0.9030	0.0861	0.9598	0.2911
PRESENCE OF UNEARNED INCOME								
NO UNEARNED INCOME	0.3852	0.1332	-0.3639	0.2319	0.0569	0.0936	0.2927	0.1425
PRESENCE OF TAXABLE TRANSFERS								
NO TAXABLE TRANSFERS	0.4000	0.1781	-0.8199	0.2759	0.1512	0.1491	0.0027	0.1712
PUBLIC HOUSING ASSISTANCE								
NO PUBLIC HOUSING	-0.0082	0.1544	-1.3613	0.6994	-0.1331	0.0734	-0.3004	0.3870

TABLE 3 (Continued)

ESTIMATES FROM PROBIT EQUATIONS PREDICTING PERSONS LIKELY TO BE NON-FILERS
BY TYPE OF PERSON

	MARRIED W KIDS, AGE < 62		MARRIED W KIDS, AGE >= 62		UNMARRIED W KIDS, AGE < 62		UNMARRIED W KIDS, AGE >= 62	
	PROBIT ESTIMATE	STD ERROR	PROBIT ESTIMATE	STD ERROR	PROBIT ESTIMATE	STD ERROR	PROBIT ESTIMATE	STD ERROR
PRESENCE OF FOOD STAMPS								
NO FOOD STAMPS	0.0295	0.1170	-0.8264	0.3585	-0.1031	0.0846	0.1796	0.2131
PRESENCE OF SOCIAL SECURITY								
NO SOCIAL SECURITY	-0.4712	0.1552	-0.1318	0.4855	-0.0573	0.1042	-0.1649	0.2354
PRESENCE OF SSI								
NO SSI	-0.4057	0.1855	-1.1241	0.4757	-0.3349	0.1320	-0.6717	0.2538
PRESENCE OF AFDC								
NO AFDC	-0.3949	0.1326	-0.2289	0.4773	-0.1808	0.0855	0.0000	0.0000
PRESENCE OF "OTHER" TRANSFERS								
NO "OTHER" TRANSFERS	0.0323	0.1263	-0.1430	0.2602	-0.1034	0.1063	0.1227	0.2814
AGI DIVIDED BY								
FILING THRESHOLD	-0.5652	0.2029	-2.2047	0.5024	-0.9825	0.1484	-0.9898	0.3107
SAMPLE SIZE	940		233		2413		692	

Table 4

CROSS-TABULATION OF ACTUAL AND PREDICTED FILERS AND NON-FILERS
 (Values are sample number of people, percent in parenthesis)

Actual	Predicted from Probit Equation		
	Non-filer	Filer	Total
Non-filer	10422 (43.93)	3303 (13.92)	13725 (57.86)
Filer	2551 (10.75)	7446 (31.39)	9997 (42.14)
Total	12973 (54.69)	10749 (45.31)	23722 (100.0)

Chi-squared statistic with 1 degree of freedom is 5933.

CONCLUSIONS AND FURTHER RESEARCH

This study, as it should, raises more questions and suggests additional avenues of research that I or others may pursue. How sensitive are my results to alternative working assumptions? How does the population examined here compare to the population of "low income" persons who are required to file? How are the results presented here likely to have changed since 1990? What, if any, is the relationship between poverty, the tax filing requirement, and tax liability in the U.S.? What is the overlap between the Earned Income Tax Credit and other means-tested government transfer programs?

One area I specifically avoided in this paper was an estimate of the actual number of non-filers. This question has its own estimating issues; and many of these issues are outside the context of this paper.

APPENDIX - DEFINITION OF VARIABLES

The definition of variables used in the probit equation are presented below. The CPS variable name is shown in parentheses. The first variable, INON, is the dependent variable. All others are independent variables.

INON

- 0 = NON-FILER
- 1 = FILER

GENDER (A-SEX)

- 1 = Male
- 2 = Female

EDUCATION LEVEL - HIGHEST GRADE ATTENDED (A-HGA)

- 1 = 3 years of high school or less (A-HGA <=11)
- 2 = 1 year of college (A-HGA =12 or =13)
- 3 = 4 years of college (A-HGA >=14 and <=16)
- 4 = 5 or more years of college (A-HGA =17 or =18)

HOUSEHOLD STATUS (HHDREL)

- 1 = Head (HHDREL =1)
- 2 = Non-Head (HHDREL >=2)

RACE (A-RACE and A-REORGN)

- 1 = Black (A-RACE =2)
- 2 = Indian, Asian, or Other (A-RACE >=3)
- 3 = White Hispanic (A-RACE =1 and A-REORGN <=7)
- 4 = White (A-RACE =1 and A-REORGN >=8)

LIVING QUARTERS (H-LIVQRT)

- 1 = House, apartment, flat (H-LIVQRT =1)
- 2 = Other (H-LIVQRT >=2)

ACTIVITY LAST WEEK (A-MAJACT)

- 1 = In the labor force (A-MAJACT >=1 and <=3)
- 2 = Keeping house or going to school (A-MAJACT =4 or =5)
- 3 = Unable to work (A-MAJACT =6)
- 4 = Retired (A-MAJACT =7)
- 5 = Other (A-MAJACT =0 or =8)

PRESENCE OF EARNED INCOME (WSAL-VAL, SEMP-VAL, and FRSE-VAL)

- 1 = No earned income
- 2 = With earned income

PRESENCE OF UNEARNED INCOME (INT-VAL, DIV-VAL, RNT-VAL, ALM-VAL (a))

- 1 = No unearned income
- 2 = With unearned income

PRESENCE OF TAXABLE TRANSFERS (RTM-VAL, UC-VAL, SRVS-VAL, DSAB-VAL (b))

- 1 = No taxable transfers
- 2 = With taxable transfers

PUBLIC HOUSING ASSISTANCE (HPUBLIC and HLORENT)

- 1 = No public housing assistance (HPUBLIC and HLORENT ^=1)
- 2 = With public housing assistance (HPUBLIC or HLORENT =1)

APPENDIX - DEFINITION OF VARIABLES (continued)

PRESENCE OF FOOD STAMPS (HFOODSP)
1 = No food stamps assistance (HFOODSP ^=1)
2 = With food stamps assistance (HFOODSP =1)
PRESENCE OF SOCIAL SECURITY (SS-VAL (a))
1 = No Social Security
2 = With Social Security
PRESENCE OF SSI (SSI-VAL)
1 = No Supplemental Security income
2 = With Supplemental Security income
PRESENCE OF AFDC (PAW-VAL (a))
1 = No public assistance including AFDC
2 = With public assistance including AFDC
PRESENCE OF "OTHER" BENEFITS (c)
1 = No other non-taxable, non-means tested transfer income
2 = With other non-taxable, non-means tested transfer income
AGI DIVIDED BY FILING THRESHOLD
= MAX(0., AGI / THRESHOLD)

-
- a = May include some money listed in other income fields such as survivor's income, disability income and/or other income.
b = Survivors income, disability income, and other income are allocated between taxable and non-taxable, depending on the values of accompanying codes.
c = All other transfer income includes veteran's benefits, non-taxable workers compensation, educational assistance and financial assistance, plus unallocated portions of survivors' income, disability income, and other income.

ENDNOTES

1. For a description of the CPS sample used in this study, see Current Population Survey, March 1991 Tape Technical Documentation, (Bureau of the Census, 1991). The Census Bureau also maintains an informative web page, see www.bls.census.gov/cps/cpsmain.htm.
2. For a discussion of non-compliant non-filers, see Brian Erard and Chin-Chin Ho, "Searching for Ghosts: Who are the Nonfilers and How Much Tax Do They Owe?" Unpublished paper presented at the Allied Social Science Associations Meetings, (1995). See also, Internal Revenue Service, Federal Tax Compliance Research: Individual Income Tax Gap Estimates for 1985, 1988, and 1992. Publication 1415. Washington DC, 1996.
3. For a description of the SOI sample, see Statistics of Income division of the Internal Revenue Service, Individual Income Tax Returns. Note that tax return data does not include information on certain non-taxable sources of income such as AFDC, SSI, food stamps, and certain other government transfer programs. Tax data is not strictly limited to information reported on individual tax returns. OTA augments the SOI sample with additional information from both tax and non-tax sources. For example, OTA exactly links each tax return on the SOI sample to a file of "information returns" held by the IRS. The most important type of information return is the W-2, which provides additional information on wages, employment taxes, and retirement plan participation. A second important information return is the SSA-RRB 1099. Here, amounts of Social Security and Railroad Retirement benefits are exactly linked to individuals on the SOI sample. As an example of non-tax data, the year of birth of taxpayers on the SOI sample is obtained from an exact match with Social Security Administration data.
4. Some low-income workers who do not reside with their children may be eligible for the new "childless" Earned Income Tax Credit. For a description of the Earned Income Tax Credit eligibility rules for any given tax year, see the Internal Revenue Service's annual publication, Publication 17, Your Federal Income Tax.
5. James Cilke, The Treasury Individual Income Tax Simulation Model. (Department of the Treasury, Office of Tax Analysis, 1994).

6. For a recent example, see Daniel Feenberg, Andrew Mitrusi, and James Poterba, "Distributional Effects of Adopting a National Sales Tax." NBER Working Paper Series no. 5885. National Bureau of Economic Research, 1997.
7. As provided under Section 6103 of the Internal Revenue Code, the Census Bureau annually receives from the Internal Revenue Service, a micro-level data file of selected information from all individual tax returns posted on IRS's Individual Returns Transactions File (IRTF) by the end of August. However, only infrequently will the Census Bureau link these tax returns to persons on the Current Population Survey (CPS). The last previous public use file of this type occurred when the March 1973 CPS file was linked to tax returns.
8. In the process of matching married CPS persons to tax returns, the Census Bureau attempted to match to both the primary and secondary SSNs. Husbands and wives matching to the same tax return had the same tax information linked to their individual records on the CPS. The absence of a secondary SSN on tax returns was a common problem for joint returns. In the case where a joint tax return with only the primary SSN present and that SSN matched to a married CPS individual, then the return was assumed to have matched to both spouses on the CPS.
9. The Office of Tax Analysis as well as a few other government agencies have received copies of the file. I am not sure whether the Census Bureau has made this file available to the public.
10. Employees of the Census Bureau have published several technical papers regarding masking confidential microdata. In particular, see Jay Kim and William Winkler, "Masking Microdata Files." in Proceedings of the Survey Research Methods Section, American Statistical Association, 1995.
11. J. Scott Turner, "Program Documentation for the UNI-TRIM Federal Income Taxes Module (FEDTAX)." May, 1976.
12. I ignored the possibility of someone on the CPS with a married-filing-separate or a qualified widow/widower filing status. I have not explicitly tried to measure the household maintenance test for purposes of determining whether a person can claim a head-of-house filing status. However, the effect of these assumptions on my results are likely to be negligible. It simply meant that I may have used an incorrect threshold amount for determining whether a person is required to file a tax return. So, for example, if I mis-identified a person as

having a head-of-house filing status instead of single, I would presume the person was not required to file if his income from taxable sources was under \$6800 instead of \$5300.

13. The relationship test includes some unmarried people who are related to someone in the household except the primary householder. The definition of a child broadly includes grandchildren and foster children. The support test does not include certain education benefits.
14. Being eligible for an additional standard deduction because of blindness does not affect the filing requirement. Further, personal exemptions besides those allowed for self and spouse do not affect the filing requirement. So, some people are required to file a Federal tax return, even if they have no tax liability.
15. There are a number of other circumstances that mandate a person file a tax return. These circumstances cannot be determined from CPS information. In particular, a person is required to file if he owes a special tax such as: a) the Alternative Minimum Tax, b) a lump-sum tax or penalty tax on a retirement plan distribution, c) Social Security taxes on tips that were not reported to an employer, or d) a recapture tax from a previously claimed credit. In addition, a person must file if he received an Advanced Earned Income Tax Credit payment, or if gross business income (business income before expenses) exceeds the regular gross income thresholds, or if the person worked for a qualified church-controlled organization that elected to exempt wage payments from Social Security taxes. For a more complete list of these circumstances that applied in 1990, see Internal Revenue Service, Publication 17, Your Federal Income Tax (1990).
16. Recall that data for this study was collected before people without children were eligible to receive the EITC. Further, approximately 96% of all EITC recipients would have a reason to file even in the absence of the EITC. See Statement of John Karl Scholz, Deputy Assistant Secretary (Tax Analysis), Department of the Treasury, Before the Committee on Ways and Means, United States House of Representative. May 8, 1997.
17. I did not want to create a situation where one spouse was likely to be a filer and the other spouse a non-filer. Recall that most of the dependent variables in the equation are shared variables and would be the same for either spouse. The spouse included in the probit equation was simply the first spouse I encountered.

TABLE 2
 PERCENT OF ACTUAL NON-FILERS AND PERSONS NOT REQUIRED TO FILE:
 BY ASSORTED DEMOGRAPHIC AND ECONOMIC CHARACTERISTICS
 BY TYPE OF PERSON

	DEPENDENTS ONLY		UNMARRIED NO KIDS, AGE < 62		UNMARRIED NO KIDS, AGE >=62		MARRIED NO KIDS, AGE < 62		MARRIED NO KIDS, AGE >=62	
	NON- FILERS (NO IRTF MATCH)	ALL NOT REQUIRED TO FILE	NON- FILERS (NO IRTF MATCH)	ALL NOT REQUIRED TO FILE	NON- FILERS (NO IRTF MATCH)	ALL NOT REQUIRED TO FILE	NON- FILERS (NO IRTF MATCH)	ALL NOT REQUIRED TO FILE	NON- FILERS (NO IRTF MATCH)	ALL NOT REQUIRED TO FILE
GENDER										
MALE	50.9	50.6	53.4	53.3	21.3	21.7	37.5	37.7	52.5	52.9
FEMALE	49.1	49.4	46.6	46.7	78.7	78.3	62.5	62.3	47.5	47.1
ACTIVITY LAST WEEK										
WORKING	4.8	9.5	16.2	31.3	1.8	3.0	7.1	21.1	1.2	3.6
WITH JOB NOT AT WORK	0.1	0.3	0.7	0.7	0.1	0.2	0.4	1.3	0.0	0.3
LOOKING FOR WORK	2.4	2.8	4.8	5.4	0.0	0.0	2.6	3.6	0.0	0.1
KEEPING HOUSE	9.1	7.8	19.4	11.7	41.0	38.5	43.1	37.2	31.7	29.7
GOING TO SCHOOL	65.0	63.9	17.5	23.4	0.0	0.0	3.5	3.0	0.1	0.1
UNABLE TO WORK	4.6	3.2	19.1	9.7	7.2	5.9	15.5	8.0	6.4	4.6
RETIRED	2.6	2.2	1.8	1.6	39.4	42.0	8.3	10.7	51.1	53.3
OTHER	11.3	10.3	20.5	16.3	10.4	10.3	19.6	15.1	9.6	8.4
EDUCATION LEVEL										
EDUC <= 10TH GRADE	65.9	54.4	41.9	26.3	61.1	52.7	51.6	38.5	61.1	47.4
EDUC 11TH OR 12TH GRADE	26.2	31.8	38.5	41.1	30.9	35.9	34.8	45.4	31.1	40.1
EDUC 1-3 YEARS COLLEGE	6.3	11.3	15.7	25.5	5.2	7.4	7.1	9.0	5.8	8.1
EDUC 4+ YEARS COLLEGE	1.5	2.5	4.0	7.0	2.7	3.9	6.5	7.0	2.0	4.4
HOUSEHOLD STATUS										
HOUSEHOLD HEAD	0.0	0.0	53.4	40.1	85.0	86.3	93.5	93.1	98.0	98.8
HOUSEHOLD MEMBER	100.0	100.0	46.6	59.9	15.0	13.7	6.5	6.9	2.0	1.2
RACE										
BLACK	21.9	19.9	29.5	22.0	15.5	12.1	15.7	10.9	10.4	7.9
ASIAN, INDIAN, OTHER	5.9	5.3	4.4	3.7	1.7	1.5	5.1	4.9	2.3	1.8
HISPANIC	11.8	10.0	9.5	8.6	4.2	3.5	11.8	9.8	5.2	4.0
WHITE	60.5	64.8	56.6	65.7	78.5	82.9	67.3	74.4	82.2	86.3

18. Table 4 shows that the sample of non-filers makes up 57.9% of the total sample not required to file a tax return. However, Table 1 shows that 55.5% of the population under investigation is a non-filer. This difference is caused by two factors. First, for the probit equations, one spouse from each married couple was dropped. Second, Table 1, reflects population estimates, where each observation is adjusted by its relative weight in the population. Table 4 uses unweighted observations.