General Description Booklet

For the

2014 PUBLIC USE TAX FILE
Demographic File

Compiled by
Victoria Bryant
Individual Statistics Branch
Statistics of Income Division
Internal Revenue Service
June 2020

Contents

2014 STATISTICS OF INCOME PUBLIC USE TAX DEMOGRAPHIC FILE

<u>INTRODUCTION</u>

The Internal Revenue Service 2014 Public Use Tax Demographic File contains 100,860 records and provides demographic data corresponding to the 2014 Public Use Tax File. The 2014 Public Use Tax File was selected as part of the Statistics of Income program that was designed to represent national statistical information for the 149.6 million Form 1040, Form 1040A, and Form 1040EZ Federal Individual Income Tax Returns filed for Tax Year 2014. The Demographic File is designed to supplement the Public Use File tax data with selected demographic filing information.

The 2014 Public Use Supplemental File is offered as part of the 2014 Public Use Tax File which can be purchased through the Internal Revenue Service, Statistics of Income Division. Any questions concerning the cost and acquisition of the current Individual Tax File should be directed to:

Director
Statistics of Income Division RAS:S
Internal Revenue Service
P.O. Box 2608
Washington, DC 20013-2608

Telephone number: (202) 803-9400

2014 STATISTICS OF INCOME PUBLIC USE TAX DEMOGRAPHIC FILE

<u>DISCLOSURE AVOIDANCE PROCEDURES</u>

To preserve the character of the microdata file while also protecting the identity of individuals, we have modified the demographic information in the following ways.

First, demographic information are only provided for Public Use File records in strata 7 through 18.

Second, the age of dependents do not appear on returns reporting farm net income or loss or a first-time homebuyers credit.

Third, dependent age information is only presented for returns that had more than one hundred returns in the population for any combination of wages, marital status, age of primary, gender of primary, number of dependents, earnings splits on joint returns, and the presence of unemployment benefits.

2014 STATISTICS OF INCOME PUBLIC USE TAX DEMOGRAPHIC FILE

<u>CORE RECORD LAYOUT</u>

Code Fields

1. AGEDP1[1]	4. AGERANGE	7. RECID
2. AGEDP2[1]	5. EARNSPLIT	
3. AGEDP3[1]	6. GENDER	

[1]Counts have been limited depending on filing status. For joint and head of household returns the limit is 3, for single returns the limit is 2 and for married filing separately returns the limit is 1. These limits are applied sequentially by type of dependent, starting with children living at home, then children living away from home, then other dependents, then parents.

2014 STATISTICS OF INCOME PUBLIC USE TAX DEMOGRAPHIC FILE

<u>CODE DEFINITIONS</u>

1. AGEDP1   Age of Dependent 1

    Not present 0
    Under 5..1
    5 under 13 2
    13 under 17 3
    17 under 19 4
    19 under 24 5
    24 or older...6

Note: Age of dependent 1 is only shown for the youngest dependent on returns sampled at a rate of 0.07 percent and further restricted based on filing status, age of primary, wages, earnings split (for joint returns), and presence of farm income, unemployment benefits, and first-time homebuyer credit.

2. AGEDP2   Age of Dependent 2

    Not present 0
    Under 5..1
    5 under 13 2
    13 under 17 3
    17 under 19 4
    19 under 24 5
    24 or older...6

Note: Age of dependent 2 is only shown for the 2nd youngest dependent on returns sampled at a rate of 0.07 percent and further restricted based on filing status, number of dependents, age of 1st dependent, age of primary, wages, earnings split (for joint returns), and presence of farm income, unemployment benefits, and first-time homebuyer credit.

3. AGEDP3   Age of Dependent 3

Not present 0
Under 5..1
5 under 13 2
13 under 17 3
17 under 19 4
19 under 24 5
24 or older...6

Note: Age of dependent 3 is only shown for the 3rd youngest dependent on returns sampled at a rate of 0.07 percent and further restricted based on filing status, number of dependents, ages of 1st and 2nd dependent, age of primary, wages, earnings split (for joint returns), and presence of farm income, unemployment benefits, and first-time homebuyer credit.

4. AGERANGE Age range of Primary Filer

For Non-dependent returns (DSI=0):
Not present 0
Under 26...1
26 under 35 2
35 under 45 3
45 under 55 4
55 under 65 5
65 or older...6

For Dependent returns (DSI=1):
Not present 0
Under 18...1
18 under 26 2
26 or older...3

Note: Age range of Primary Filer is only shown for returns sampled at a rate of 0.07 percent

5. EARNSPLIT Earnings Split on Joint Returns

Not present 0
75 percent or more earned by primary 1
Less than 75 percent but more than 25 percent earned by primary 2
Less than 25 percent earned by primary 3

Note: Earning Splits are shown only for those returns sampled at a rate of 0.07 percent and report a filing status of Married Filing Jointly.

6. GENDER   Gender of Primary filer

Primary Filer is Male 1
Primary Filer is Female 2

7. RECID   RETURN ID
A unique number assigned for each sampled tax return during processing.
This number should be used for a one-to-one match to the 2014 Public Use File.

2014 STATISTICS OF INCOME PUBLIC USE TAX DEMOGRAPHIC FILE

TECHNICAL DESCRIPTION OF THE FILE

     The Public Use Tax Supplemental File is available in CSV format. This comma separated file contains 100,861 records where the first record consisting of only field names. Field names have been ordered by code fields in the same order as presented on page 4.

2014 STATISTICS OF INCOME PUBLIC USE TAX DEMOGRAPHIC FILE

DESCRIPTION OF FILE

This section describes the domain of the supplemental file and method for joining to the 2014 PUF.

Domain of the File

The data in this file are categorical demographic information reported on unaudited Individual Income Tax Returns found on the 2014 PUF for records in strata 7 through 18, or those returns sampled at a 0.07 percent rate. For additional information on filing distribution across strata, see pages 27 through 31 of the 2014 Public Use Tax File General Description Booklet. Table A reports the total count of returns selected for inclusion in the PUF and the Supplemental File by sample strata.

This supplemental file provides additional demographic information by adding variables for age (in ranges) for primary taxpayers and dependents, gender for primary taxpayers, and earnings splits (in ranges) for joint filers. The age ranges differ for primary taxpayers who are not dependents of another taxpayer, primary taxpayers who are dependents of another taxpayer, and dependents.   Age variables for dependents provide the age range for each dependent, starting with the youngest, up to the cap on the number of dependents for the marital status; these variables are not associated with specific dependent types. Earning for earning splits include wages and self-employment income reported on Schedule SE, for primary filers this is equal to E30400 minus E30500, for secondary filers this is equal to E30500.

Data Joining to 2014 Public Use Tax File

Data should be joined to the 2014 PUF using a one-to-one match on RECID.

2014 STATISTICS OF INCOME PUBLIC USE TAX DEMOGRAPHIC FILE

Table A.—Number of Individual Income Tax Returns in the Public-use file and Demographic File by Sampling Strata for 2014

	Number of returns	
Description of the sample strata	Public-use sample counts	Demographic counts
Grand total	178,130	100,860
Form 1040 returns only with adjusted gros income or expanded income of $200,000 and over, with no income tax after credits and no additional tax for tax preferences, total	0	0
Form 1040 returns only with combined Schedule C (business or profession) total receipts of $50,000,000 and over, total	26	0
Other Returns, total	178,104	100,860

Description of the sample strata	Degree of interest [1]	Form 1040, with Form 1116 or Form 2555		Form 1040, with Schedule C but without Form 1116 or Form 2555		Form 1040, with Schedule F but without Scheudle C, Form 1116 or Form 2555		All other forms		Public-use sample counts	Demographic counts
		Public-use sample counts	Supplemental counts	Public-use sample counts	Supplemental counts	Public-use sample counts	Supplemental counts	Public-use sample counts	Supplemental counts		
	(1)	(2)	(3)	(5)	(6)	(8)	(9)	(11)	(12)		
Total		34,377	3,824	32,007	15,744	4,440	837	101,561	80,455	178,104	100,860
Indexed Negative Income [2]											
$10,000,000 or more	All	35	-	125	-	14	-	128	-	302	-
$5,000,000 under $10,000,000	All	67	-	192	-	22	-	225	-	506	-
$2,000,000 under $5,000,000	All	306	-	669	-	105	-	866	-	1,946	-
$1,000,000 under $2,000,000	All	641	-	1,293	-	242	-	1,759	-	3,935	-
$500,000 under $1,000,000	All	505	-	1,051	-	202	-	1,312	-	3,070	-
$250,000 under $500,000	All	308	-	643	-	131	-	801	-	1,883	-
$120,000 under $250,000	All	30	30	89	89	17	17	124	124	260	260
$60,000 under $120,000	All	37	37	99	99	20	20	172	172	328	328
Under $60,000	All	29	29	269	269	14	14	373	373	685	685
Indexed Positive Income [3]										-	-
Under $30,000	1	-		-		-		24,211	24,211	24,211	24,211
Under $30,000	2	178	178	2,678	2,678	48	48	18,038	18,038	20,942	20,942
Under $30,000	3-4	193	193	4,005	4,005	71	71	4,714	4,714	8,983	8,983
$30,000 under $60,000	1-2	434	434	1,351	1,351	100	100	14,994	14,994	16,879	16,879
$30,000 under $60,000	3-4	509	509	2,606	2,606	170	170	4,743	4,743	8,028	8,028
$60,000 under $120,000	1-3	817	817	1,550	1,550	140	140	8,124	8,124	10,631	10,631
$60,000 under $120,000	4	595	595	1,781	1,781	119	119	2,364	2,364	4,859	4,859
$120,000 under $250,000	1-3	256	256	287	287	59	59	1,023	1,023	1,625	1,625
$120,000 under $250,000	4	746	746	1,029	1,029	79	79	1,575	1,575	3,429	3,429
$250,000 under $500,000	All	4,808	-	3,750	-	578	-	5,385	-	14,521	-
$500,000 under $1,000,000	All	7,612	-	3,796	-	808	-	4,834	-	17,050	-
$1,000,000 under $2,000,000	All	12,279	-	4,168	-	932	-	4,976	-	22,355	-
$2,000,000 under $5,000,000	All	5,611	-	1,226	-	235	-	1,541	-	8,613	-
$5,000,000 under $10,000,000	All	1,391	-	216	-	37	-	267	-	1,911	-
$10,000,000 or more	All	907	-	101	-	15	-	129	-	1,152	-

Note: The header spanning "Number of Returns by type of form attached" covers the Form 1040 and All other forms columns.

[1] Each population member is assigned a degree of interest based on how useful it is for tax modeling purposes. Degree of interest ranges from one (1) to four (4), with a one being assigned to returns that are the least interesting, and a four being assigned to those that are the most interesting. 'All' refers to income classes for which returns with all four degrees of interest are assigned.

[2] Positive and Negative Income classes are divided by a Chain-Type Price Index for the Gross Domestic Product of 1.5403 to represent a base year of 1991.

Weighted Sum for Each Code Field

Codes and Values	2014 Full SOI Individual Sample	2014 Public Use Sample	Full Sample less Public Use	Percentage Difference between Full and Public use
AGEDP1				
0	97,004,245	98,299,040	-1,294,795	-1.33%
1	13,315,282	12,350,875	964,407	7.24%
2	18,570,722	16,988,177	1,582,545	8.52%
3	7,699,892	6,666,920	1,032,972	13.42%
4	3,402,537	2,920,619	481,918	14.16%
5	4,699,129	3,934,536	764,593	16.27%
6	3,914,771	3,463,908	450,863	11.52%
AGEDP2				
0	121,076,870	121,283,085	-206,215	-0.17%
1	2,135,237	1,924,728	210,509	9.86%
2	11,668,107	10,678,422	989,685	8.48%
3	5,724,959	5,102,292	622,667	10.88%
4	2,598,310	2,118,258	480,052	18.48%
5	3,313,244	2,220,464	1,092,780	32.98%
6	2,089,852	1,296,825	793,027	37.95%
AGEDP3				
0	138,792,472	137,130,019	1,662,453	1.20%
1	103,043	83,494	19,549	18.97%
2	3,393,401	2,991,270	402,131	11.85%
3	2,584,587	2,193,338	391,249	15.14%
4	1,165,262	995,010	170,252	14.61%
5	1,337,257	746,280	590,977	44.19%
6	1,230,555	484,664	745,891	60.61%
AGERANGE				
DSI=0				
0	236,529	235,543	986	0.42%
1	16,525,863	16,519,903	5,960	0.04%
2	26,445,133	26,362,007	83,126	0.31%
3	24,702,410	24,006,660	695,750	2.82%
4	25,607,442	24,384,928	1,222,514	4.77%
5	22,450,584	21,263,379	1,187,205	5.29%
6	23,594,071	22,841,112	752,959	3.19%

2014 STATISTICS OF INCOME PUBLIC USE TAX DEMOGRAPHIC FILE

Weighted Sum for Each Code Field

Codes and Values	2014 Full SOI Individual Sample	2014 Public Use Sample	Full Sample less Public Use	Percentage Difference between Full and Public use
AGERANGE				
DSI=1				
0	46,161	47,011	-850	-1.84%
1	1,785,331	1,801,181	-15,850	-0.89%
2	6,946,806	6,901,921	44,885	0.65%
3	266,246	260,428	5,818	2.19%
EARNSPLIT				
0	101,299,402	100,007,317	1,292,085	1.28%
1	22,050,715	20,473,646	1,577,069	7.15%
2	17,762,367	17,105,320	657,047	3.70%
3	7,494,094	7,037,792	456,302	6.09%
GENDER [1]				
1	93,710,336	90,276,564	3,433,772	3.66%
2	54,896,242	54,347,511	548,731	1.00%

[1] This is the Gender of the primary filer.