

**General Description Booklet**

**For the**

**2015 PUBLIC USE TAX FILE  
Demographic File**

Compiled by  
Victoria Bryant  
Individual Statistics Branch  
Statistics of Income Division  
Internal Revenue Service  
September 2022

## Contents

INTRODUCTION .....	2
DISCLOSURE AVOIDANCE PROCEDURES .....	3
CORE RECORD LAYOUT .....	4
CODE DEFINITIONS.....	5
TECHNICAL DESCRIPTION OF THE FILE.....	8
DESCRIPTION OF FILE .....	9
WEIGHTED COUNTS FOR EACH CODE FIELD.....	11

# **2015 STATISTICS OF INCOME PUBLIC USE TAX DEMOGRAPHIC FILE**

## **INTRODUCTION**

The Internal Revenue Service 2015 Public Use Tax Demographic File contains 119,675 records and provides demographic data corresponding to the 2015 Public Use Tax File. The 2015 Public Use Tax File was selected as part of the Statistics of Income program that was designed to represent national statistical information for the 151.2 million Form 1040, Form 1040A, and Form 1040EZ Federal Individual Income Tax Returns filed for Tax Year 2015. The Demographic File is designed to supplement the Public Use File tax data with selected demographic filing information.

The 2015 Public Use Supplemental File is offered as part of the 2015 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

## 2015 STATISTICS OF INCOME PUBLIC USE TAX DEMOGRAPHIC FILE

### **DISCLOSURE AVOIDANCE PROCEDURES**

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 13.

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.

## 2015 STATISTICS OF INCOME PUBLIC USE TAX DEMOGRAPHIC FILE

### CORE RECORD LAYOUT

#### Code Fields

- |                        |              |          |
|------------------------|--------------|----------|
| 1. AGEDP1 <sup>1</sup> | 4. AGERANGE  | 7. RECID |
| 2. AGEDP2 <sup>1</sup> | 5. EARNSPLIT |          |
| 3. AGEDP3 <sup>1</sup> | 6. GENDER    |          |

<sup>1</sup>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.

2015 STATISTICS OF INCOME PUBLIC USE TAX DEMOGRAPHIC FILE

**CODE DEFINITIONS**

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 2<sup>nd</sup> youngest dependent on returns sampled at a rate of 0.07 percent and further restricted based on filing status, number of dependents, age of 1<sup>st</sup> dependent, age of primary, wages, earnings split (for joint returns), and presence of farm income, unemployment benefits, and first-time homebuyer credit.*

**2015 STATISTICS OF INCOME PUBLIC USE TAX DEMOGRAPHIC FILE**

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 3<sup>rd</sup> youngest dependent on returns sampled at a rate of 0.07 percent and further restricted based on filing status, number of dependents, ages of 1<sup>st</sup> and 2<sup>nd</sup> 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*

**2015 STATISTICS OF INCOME PUBLIC USE TAX DEMOGRAPHIC FILE**

**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.



## **2015 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 119,675 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.

## 2015 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 2015 PUF.

#### **Domain of the File**

The data in this file are categorical demographic information reported on unaudited Individual Income Tax Returns found on the 2015 PUF for records in strata 7 through 13, 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 2015 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 2015 Public Use Tax File**

Data should be joined to the 2015 PUF using a one-to-one match on RECID.

## 2015 STATISTICS OF INCOME PUBLIC USE TAX DEMOGRAPHIC FILE

**Table A.—Number of Individual Income Tax Returns in the Population and Sample by Sampling Strata for 2015**

Description of the sample strata									Number of returns	
									Public-use sample counts [1]	Sample counts
									207,692	119,675
Grand total									0	0
Form 1040 returns only with adjusted gross income or expanded income of \$200,000 and over, with no income tax after credits and no additional tax for tax prefer									27	0
Form 1040 returns only with combined Schedule C (business or profession) total receipts of \$50,000,000 and over, total									207,665	119,675
Other Returns, total										
	Number of Returns by type of form attached									
	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			
Description of the sample strata	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)	(4)	(5)	(7)	(8)	(10)	(11)		
Total.....	50,217	7,045	39,066	20,583	4,390	1,257	113,992	90,790	207,665	119,675
Indexed Negative Income [2]										
\$10,000,000 or more.....	58	-	122	-	14	-	129	-	323	-
\$5,000,000 under \$10,000,000....	68	-	159	-	23	-	227	-	477	-
\$2,000,000 under \$5,000,000.....	298	-	672	-	108	-	858	-	1,936	-
\$1,000,000 under \$2,000,000.....	625	-	1,336	-	249	-	1,811	-	4,021	-
\$500,000 under \$1,000,000.....	470	-	1,032	-	213	-	1,315	-	3,030	-
\$250,000 under \$500,000.....	289	-	637	-	128	-	800	-	1,854	-
\$120,000 under \$250,000.....	36	36	93	93	19	19	124	124	272	272
\$60,000 under \$120,000.....	35	35	112	112	19	19	156	156	322	322
Under \$60,000.....	28	28	251	251	15	15	346	346	640	640
Indexed Positive Income										
Under \$30,000.....	371	371	6,844	6,844	127	127	48,129	48,129	55,471	55,471
\$30,000 under \$60,000.....	922	922	4,092	4,092	253	253	19,970	19,970	25,237	25,237
\$60,000 under \$120,000.....	1,386	1,386	3,441	3,441	286	286	10,764	10,764	15,877	15,877
\$120,000 under \$250,000.....	4,267	4,267	5,750	5,750	538	538	11,301	11,301	21,856	21,856
\$250,000 under \$500,000.....	4,454	-	3,619	-	502	-	4,991	-	13,566	-
\$500,000 under \$1,000,000.....	7,087	-	3,728	-	630	-	4,371	-	15,816	-
\$1,000,000 under \$2,000,000.....	14,084	-	4,854	-	899	-	5,739	-	25,576	-
\$2,000,000 under \$5,000,000.....	5,499	-	1,206	-	188	-	1,582	-	8,475	-
\$5,000,000 under \$10,000,000....	1,429	-	238	-	32	-	275	-	1,974	-
\$10,000,000 or more.....	8,811	-	880	-	147	-	1,104	-	10,942	-

[1] This sample count excludes the 4 aggregated records capturing extremely high/low income returns.

[2] Positive and Negative Income classes are divided by a Chain-Type Price Index for the Gross Domestic Product of 1.5874 to represent a base year of 1991.

**2015 STATISTICS OF INCOME PUBLIC USE TAX DEMOGRAPHIC FILE**

**Weighted Sum for Each Code Field**

<b>Codes and Values</b>	<b>2015 Full SOI Individual Sample</b>	<b>2015 Public Use Sample</b>	<b>Full Sample less Public Use</b>	<b>Percentage Difference between Full and Public use</b>
<b>AGEDP1</b>				
0	99,128,287	100,924,257	-1,795,970	-1.81%
1	13,158,025	12,276,699	881,326	6.70%
2	18,626,501	17,166,333	1,460,168	7.84%
3	7,719,946	6,724,879	995,067	12.89%
4	3,416,718	2,998,461	418,257	12.24%
5	4,599,165	3,842,279	756,886	16.46%
6	3,844,620	3,557,274	287,346	7.47%
<b>AGEDP2</b>				
0	123,199,216	124,149,435	-950,219	-0.77%
1	2,147,986	1,979,024	168,962	7.87%
2	11,489,949	10,535,128	954,821	8.31%
3	5,790,489	5,088,318	702,171	12.13%
4	2,571,052	2,202,105	368,947	14.35%
5	3,302,301	2,224,240	1,078,061	32.65%
6	1,992,270	1,311,931	680,339	34.15%
<b>AGEDP3</b>				
0	140,768,233	140,017,810	750,423	0.53%
1	101,099	83,975	17,124	16.94%
2	3,258,669	2,855,614	403,055	12.37%
3	2,595,202	2,267,820	327,382	12.61%
4	1,199,295	1,030,582	168,713	14.07%
5	1,326,115	711,006	615,109	46.38%
6	1,244,650	523,374	721,276	57.95%
<b>AGERANGE</b>				
<b>DSI=0</b>				
0	286,436	303,925	-17,489	-6.11%
1	16,613,774	16,641,953	-28,179	-0.17%
2	26,681,540	26,638,676	42,864	0.16%
3	24,983,876	24,698,493	285,383	1.14%
4	25,227,793	24,134,643	1,093,150	4.33%
5	23,034,204	22,184,854	849,350	3.69%
6	24,377,002	23,632,784	744,218	3.05%

## 2015 STATISTICS OF INCOME PUBLIC USE TAX DEMOGRAPHIC FILE

### Weighted Sum for Each Code Field

Codes and Values	2015 Full SOI Individual Sample	2015 Public Use Sample	Full Sample less Public Use	Percentage Difference between Full and Public use
AGERANGE				
DSI=1				
0	39,031	40,297	-1,266	-3.24%
1	1,925,073	1,907,312	17,761	0.92%
2	7,095,409	7,091,415	3,994	0.06%
3	229,124	215,830	13,294	5.80%
EARNSPPLIT				
0	103,169,077	102,557,960	611,117	0.59%
1	21,891,567	20,529,190	1,362,377	6.22%
2	17,875,625	17,240,887	634,738	3.55%
3	7,556,993	7,162,145	394,848	5.22%
GENDER [1]				
1	93,989,333	91,303,295	2,686,038	2.86%
2	56,503,930	56,186,887	317,043	0.56%

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