

## Measuring Income from SSI Benefits with Administrative Universe Data<sup>1</sup>

**Angela Wyse**  
University of Chicago

**Gillian Meyer**  
University of Pennsylvania

**Derek Wu**  
University of Virginia

**Bruce D. Meyer**  
University of Chicago,  
NBER, and AEI

May 6, 2024

### I. Overview

This document describes our methodology for measuring income from Supplemental Security Income (SSI) benefits with universe data from the Social Security Administration (SSA). The SSI program provides income support to aged, blind, and disabled people who have little or no income. In 2020, the program paid \$56.3 billion in benefits to 8.7 million beneficiaries.<sup>2</sup>

We produce a cleaned dataset indicating federal and state SSI benefit payments for a single month of each year 2010-2014 and 2016. Our analyses suggest that SSA's extraction methods resulted in the incorrect attribution of zero benefit receipt to about 1.59-3.46 percent of true SSI recipients in these datasets. We propose scaling factors for researchers wishing to measure monthly and annual SSI receipt with these data. First, we propose scaling factors to remedy the incorrect attribution of zero benefit receipt to the small share of individuals described above. Second, because each year's dataset reflects receipt in a single month, we calculate a scaling factor to estimate annual rather than monthly receipt. Our cleaned and scaled datasets track closely with public aggregates, producing recipient counts within 0.7% of monthly published totals and 6.4% of annual published totals. We produce benefit amounts generally within 5.3% of monthly totals (except in 2011, where we underreport in comparison to monthly public aggregates by 11.8%). We produce benefit amounts generally within 4.9% of monthly totals, but we underreport in comparison to annual public aggregates by 10.4%.

---

<sup>1</sup> This memo is released to inform interested parties of research and to encourage discussion. Any views expressed are those of the authors and not those of the U.S. Census Bureau. The Census Bureau has reviewed this data product for unauthorized disclosure of confidential information and has approved the disclosure avoidance practices applied to this release, authorization number: CBDRB-FY2023-CES005-011. Authors can be contacted at the following email addresses: Angela Wyse ([awyse@uchicago.edu](mailto:awyse@uchicago.edu)); Gillian Meyer ([gplmeyer@wharton.upenn.edu](mailto:gplmeyer@wharton.upenn.edu)); Derek Wu ([derek.wu@virginia.edu](mailto:derek.wu@virginia.edu)); Bruce D. Meyer ([meyer1@uchicago.edu](mailto:meyer1@uchicago.edu)).

<sup>2</sup> SSI Annual Statistical Report, 2020

## II. Description of Administrative Data

Our primary administrative data source is an extract of SSA's Supplemental Security Income Record (SSR) database called the Address file. The Address file contains a unique record for each SSI beneficiary and some ineligible family members in its extraction month. Each observation indicates address, benefit amount (state and federal), date of birth, and eligibility status (e.g. disabled child, aged adult, ineligible parent, ineligible spouse, etc.). Records can be linked to other datasets using Protected Identification Keys (PIKs).

We draw on a second administrative data source to assess the coverage and completeness of the Address file. This second source, which we call the CPS-linked SSI file, contains a monthly history of SSI receipt for people surveyed in various years of the Current Population Survey (CPS). While they do not provide the universe of SSI recipients, these files are useful because they are constructed from the SSR directly, not the Address file, and therefore do not have the same shortcomings as the Address file.

The Address files available to Census researchers do not indicate which month of data we observe. We investigate this question by linking the Address file to the CPS-linked SSI files. Those analyses indicate that the 2010-2015 Address files reflect receipt in May in 2010, April in 2011-2014, and March in 2016.<sup>3</sup> We validate benefit amounts in the Address file by comparing them to the CPS-linked SSI files and find that they match in nearly all cases in the months listed above.

## III. Addressing Limitations of Cleaned Datasets

The cleaned datasets suffer from two shortcomings that impair their usefulness for calculating annual SSI benefit receipt. This section describes those shortcomings and our approach to addressing them.

### *1. Lack of information about months other than extraction month*

The Address file indicates receipt in a single month of each data year. We do not have receipt indicators or benefit amounts for people who received benefits only in other months of the year.

To address this issue, we propose scaling factors within cells based on age group and year. We use the CPS-linked SSI files to calculate these factors because they allow us to calculate both monthly and annual receipt. Scaling factors are calculated as the ratio of people in an age group cell who have receipt in any month of the year to those with receipt in the month covered by the Address file. We calculate scaling factors ranging from 1.014 to 1.140 across age groups and year.

### *2. False indications of zero benefit receipt in Address file*

In the broader SSR database, which is not available to Census researchers, an SSN may appear on multiple records due to marriage, divorce, or other changes in living arrangements. In selecting the unique record with current payment address, the Address file extraction process drops all other records for a given SSN, regardless of whether those records indicate a non-zero benefit amount.

---

<sup>3</sup> The 2015 file appears to contain information on receipt in the following year, 2016, despite its label.

We estimate that this extraction process causes about 1.59-3.46% of recipients to be incorrectly identified as non-recipients.

To address this issue, we once again calculate scaling factors within cells based on age group and year using the CPS-linked SSI file. We first calculate the number of SSI recipients in the CPS-linked SSI file for a given age group and year. We then link these individuals to the Address file and calculate the number with non-zero receipt in that file. The difference between these counts is the number of people falsely indicated as having zero benefits in the Address file. The scaling factor is the ratio of the first to second count. These scaling factors range from 1.001 to 1.055 across age groups and year.

#### **IV. Benchmarking Cleaned and Scaled SSI Receipt**

Our cleaned and scaled estimates from universe data benchmark well against public aggregates at the monthly level for recipients, where we fall within 0.7% of estimates published in SSI monthly reports. Our monthly benefit amount estimates are slightly less accurate, falling for most years within 5.3% of public aggregates. Our annual estimates, too, benchmark well against public aggregates but are less precise than our monthly estimates, falling within 6.4% of published estimates for recipient counts and, for most years, within 4.9% of benefit amounts.

Notably, we significantly overestimate both monthly and annual benefit amounts in 2011 by 11.8% and 10.4% respectively.

## V. Tables

<b>Benchmarking Cleaned and Scaled SSI Receipt against Public Aggregates</b>					
<b>From Public Data</b>					
<b>Year</b>	<b>Recipients</b>	<b>Monthly</b>		<b>Annual</b>	
		<b>Benefits (Millions)</b>	<b>Recipients</b>	<b>Benefits (millions)</b>	
2010	7,800,015	4,205	9,176,526	48,195	
2011	8,014,930	4,313	9,306,902	49,520	
2012	8,185,900	4,554	9,179,222	52,075	
2013	8,331,703	4,718	9,267,240	53,900	
2014	8,414,517	4,822	9,259,225	54,693	
2016	8,334,922	4,752	9,177,799	54,799	
<b>From Universe Data</b>					
<b>Year</b>	<b>Recipients</b>	<b>Monthly</b>		<b>Annual</b>	
		<b>Benefits (Millions)</b>	<b>Recipients</b>	<b>Benefits (millions)</b>	
2010	7,744,000	4,205	8,590,000	48,190	
2011	7,994,000	4,822	8,875,000	54,690	
2012	8,192,000	4,313	9,059,000	49,520	
2013	8,288,000	4,554	9,161,000	52,080	
2014	8,379,000	4,718	9,029,000	53,900	
2016	8,296,000	4,752	8,803,000	54,800	
<b>Universe/Public Ratio</b>					
<b>Year</b>	<b>Recipients</b>	<b>Monthly</b>		<b>Annual</b>	
		<b>Benefits (Millions)</b>	<b>Recipients</b>	<b>Benefits (millions)</b>	
2010	0.9928	1.0000	0.9361	0.9999	
2011	0.9974	1.1180	0.9536	1.1044	
2012	1.0007	0.9471	0.9869	0.9509	
2013	0.9948	0.9652	0.9885	0.9662	
2014	0.9958	0.9784	0.9751	0.9855	
2016	0.9953	1.0000	0.9592	1.0000	

Sources: SSI Annual Statistical Reports 2010-2016 (Tables 2 and 3), SSI Monthly Statistics 2010-2016 (Table 1), 2010-2015 SSR Files

**Validating SSI Universe Benefit Receipt Month and Amount**

<i>Share of CPS-SSR Extract Recipients with Universe Receipt in Month</i>						
Month	2010	2011	2012	2013	2014	2016
1	0.9369	0.9211	0.9310	0.9316	0.9483	0.9469
2	0.9459	0.9298	0.9310	0.9402	0.9569	0.9558
3	0.9550	0.9386	0.9397	0.9487	0.9655	0.9646
4	0.9640	0.9561	0.9569	0.9658	0.9828	0.9558
5	0.9641	0.9474	0.9483	0.9487	0.9655	0.9469
6	0.9550	0.9386	0.9397	0.9402	0.9569	0.9381
7	0.9459	0.9298	0.9310	0.9402	0.9483	0.9292
8	0.9459	0.9211	0.9224	0.9316	0.9483	0.9204
9	0.9369	0.9211	0.9224	0.9316	0.9397	0.9115
10	0.9369	0.9123	0.9224	0.9231	0.9310	0.9115
11	0.9279	0.9123	0.9138	0.9145	0.9224	0.9027
12	0.9189	0.9035	0.9052	0.9060	0.9138	0.8938

*Share with exactly matching month benefit (in month with maximum agreement share for a given year)*

1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
--------	--------	--------	--------	--------	--------	--------

**Note:** Top panel of table displays unweighted share of individuals with positive benefit receipt in a given month/year according to the CPS SSI extracts who appear in the SSI universe file for the respective year with non-zero benefit receipt. Bottom panel takes month with maximum agreement share and displays the share of individuals with exactly matching monthly benefit amounts.

**Source:** 2010-2015 SSR Files, CPS-SSR Extract files (1978-2020)

**SSI Scaling Factors by Age Group and Type**

Year	Scaling Factor A	Scaling Factor B
Ages 0-17		
2010	1.123	1.013
2011	1.139	1.006
2012	1.107	1.005
2013	1.121	1.002
2014	1.114	1.004
2016	1.077	1.001
Ages 18-64		
2010	1.132	1.022
2011	1.137	1.039
2012	1.140	1.047
2013	1.125	1.036
2014	1.090	1.034
2016	1.077	1.055
Ages 65+		
2010	1.051	1.004
2011	1.032	1.009
2012	1.026	1.006
2013	1.050	1.008
2014	1.028	1.011
2016	1.014	1.012

**Note:** Scaling factors are weighted ratios of recipients in SSI universe data and in CPS-ASEC-Admin file. Scaling factor A accounts for the fact that we see only a single month of benefit receipt. Scaling factor B accounts for false zeroes. Annual beneficiary scaling factor is the product of A and B. Monthly beneficiary scale is B. Annual amount scaling factor is 12\*B.

**Source:** 2010-2015 SSR files, CPS-SSR extracts (1978-2020)