

Memo

To: Paul O'Leary and Debra Tidwell-Peters
From: Dawn Phelps and Jody Schimmel Hyde
Date: 6/30/2020
Subject: Changes over time in SSI Earnings in the Disability Analysis File

The purpose of this memo is to describe changes over time in the source of earnings for Supplemental Security Income (SSI) beneficiaries in the Social Security Administration's DAF, and the rationale for such changes.

History of SSI Earnings in the DAF

Prior to DAF16, SSI earnings in the DAF were derived from the Disability Control File (DCF). While working with the DAF15, SSA discovered that some earnings that were recorded in the SSR were not being included in the DCF. As a result, the earnings values in the DAF prior to DAF16 were underestimates of earnings values. An investigation at the time suggested that the DCF values for employment were about 5 percent lower than those from the SSR. In the DAF16, we updated the SSI earnings data to be sourced from the Supplemental Security Record (SSR).

Under the structure of the SSR, earnings are set in their first applicable month and then carried forward until new information of an earnings increase, decrease, or cessation is recorded. As we were building DAF18, we determined that the earnings data we constructed using the SSR overstated beneficiary earnings because we carried earnings forward in some cases where those earnings should have ceased. This error was present for SSI earnings in both the DAF16 and DAF17, though the error primarily affected the DAF16, as we discovered the problem shortly after producing the DAF17 and therefore removed the earning files from the DAF at that point. We have corrected the issue as of the DAF18 version released in July 2020.

To explain the issue, it helps to understand the SSR extract that we use to process records for the DAF. The SSR extract we receive is organized by earnings spell, rather than beneficiary like the DAF structure. Specifically, it includes an observation for every beneficiary, record establishment date, earnings date, and earnings type (including exclusions). A beneficiary can have multiple records established with multiple earnings types across multiple months. Each record includes variables that indicate the type of the earnings (wage, self-employment gain, self-employment loss, income-related work expense, pathway to self-sufficiency, and student exclusion), the frequency of earnings (continuous, one-time, or terminated), and the amount of the earnings or exclusion. Generally, the record also has the dates when specific earnings values start and end.

To include the data on the DAF, we need to transform this source data so that it is at the beneficiary level with earnings data contained in monthly variables. We first collapse the observations to the record establishment level, using the earnings start and end dates. We create monthly earnings and exclusion variables for each SSN and record establishment date combination. However, some of these earnings have a start date but no stop date. In those cases, earnings are carried forward.

To: Paul O'Leary and Debra Tidwell-Peters
From: Dawn Phelps and Jody Schimmel Hyde
Date: 6/30/2020
Page: 2

Mathematica

Until DAF18, we were not fully aware of the restrictions we needed to include on how far forward the continuous earnings (records with no end date) should be propagated. As a result, continuous earnings were carried forward longer than they should have been. For example, if a continuous value existed without an end date, we would carry that value forward indefinitely, replacing the value in months where one-time earnings were observed.¹ Additionally, the termination code on earnings values does not fully account for beneficiaries whose benefits are terminated, so we have used other data available in the SSI Longitudinal File to stop continuous earnings following termination of benefits. Without curtailing earnings in these cases, we were showing earnings that were almost double what is shown in SSA published statistics. As of DAF18 we have accounted for these scenarios to properly restrict the propagation of earnings values.

We then we collapse the resulting record establishment level file to the beneficiary level. To do this, we keep values for each monthly occurrence from the latest established record, eliminating values in months for which a beneficiary is in terminated status.

Alignment of SSI Earnings in the DAF to SSA Published Statistics

Our earnings based on the SSR are now close to SSA published statistics (Table 1). For example, in December 2018, we identify 386,801 SSI beneficiaries with earnings and average earnings of \$771, compared to 342,046 SSI beneficiaries with earnings and average earnings of \$783 in SSA published statistics (specifically, [Table 46 of the 2018 SSI Annual Statistical Report](#)). We attribute the differences between the DAF and SSA published statistics to be due to timing of the data pull; we assume that DAF data was pulled slightly later than information for SSA's tables, allowing more time for SSA to post earnings to the SSR that have been reported by beneficiaries or identified through earnings enforcement procedures.

We did not find any critical differences between the DAF and SSA published statistics in the share of beneficiaries with earnings or average earnings in December 2018 based on earnings or age groups (Table 1). In most cases, DAF identifies slightly more beneficiaries with earnings but about the same average earnings among earners, without a clear pattern of differences by earnings group or age.

For users who plan to work with the DAF, it is important to note that to align to SSA published statistics, we limited to beneficiaries whose SSA payment status is current pay (C01), eligible for federal and/or state benefits based on eligibility computation, but no payment is due based on the payment computation (E01), or non-pay - Claimant's chargeable income exceeds both the applicable SSI payment and his state's payment standard (N01). The number of beneficiaries with earnings without that restriction is about 10 percent higher (421,743 in December 2018), and with \$883 monthly earnings.

¹ For example, suppose a continuous earnings value of \$356 started in July 2012, but a one-time value of \$110 appeared in January 2015. Our previous algorithm would have populated a value of earnings in \$356 in each month from July 2012 through December 2014, a value of \$110 in January 2015, then a value of \$356 again from February 2015 until a termination code was reached or another record was identified. In the current algorithm, \$356 would be populated from July 2012 through December 2014, \$110 in January 2015, and \$0 until another earnings start date is identified.

Remaining Concerns about SSI Earnings Data from the DCF

As part of our investigation to benchmark SSI earnings in the DAF to SSA published statistics, we also analyzed the earnings posted in the DCF. Our review finds that the number of SSI beneficiaries with earnings in the DCF are substantially underestimated. For example, information on SSI earnings from the DCF in December 2018 showed 182,383 SSI beneficiaries with earnings compared to 342,046 SSI beneficiaries with earnings in the published statistics. The difference between the number of earners relative to published statistics is substantially larger than we understood at the time that we made the shift from the DCF to the SSR in DAF16.

We also found that this issue with the DCF has become more problematic over time (Table 2) and is especially problematic for work incentives (not shown). In December 2018, the number of earners in the DCF was 53 percent of the number shown in published statistics, while it was 60 percent in December 2017, 65 percent in December 2016 and 71 percent in December 2015. We confirmed that these differences are not due to sampling for the DAF; the values in the DAF mirror those in the DCF extract used to build the DAF.

In 2018, the average earnings values were similar in the DCF, the SSR and the SSA published statistics. As shown in Table 2, the average earnings in the DCF in earlier years were substantially higher than in SSA's published statistics.

Table 1. SSI Beneficiary Earnings in December 2018, in SSA Published Statistics and the DAF18

	SSA Published Statistics		DAF181	
	Number of beneficiaries with earnings	Average Earnings (\$)	Number of beneficiaries with earnings	Average Earnings (\$)
Total	342,046	783	373,062	783
Earned income				
Wages	316,597	816	341,933	815
Self-employment income	27,377	436	35,581	372
Earnings (dollars)				
65 or less	52,708	39	56,736	37
66–99	13,032	82	15,558	82
100–199	32,411	142	36,561	143
200–299	26,801	239	29,503	242
300–399	22,878	339	24,742	342
400–499	20,504	438	22,447	441
500–599	19,603	535	20,556	540
600–699	17,695	636	19,091	641
700–799	14,900	739	16,322	742
800–899	14,921	835	16,248	839
900–999	10,978	937	12,370	942
1,000–1,099	11,961	1,031	12,643	1,035

	SSA Published Statistics		DAF181	
	Number of beneficiaries with earnings	Average Earnings (\$)	Number of beneficiaries with earnings	Average Earnings (\$)
1,100–1,199	7,691	1,138	8,729	1,141
1,200–1,299	8,616	1,232	8,779	1,236
1,300–1,399	5,830	1,338	6,323	1,341
1,400–1,499	5,433	1,436	5,700	1,440
1,500 or more	56,084	2,390	60,754	2,411
Age				
Under 18	3,050	833	3,532	816
18–21	31,289	700	28,965	650
22–25	47,743	798	50,373	789
26–29	52,239	850	55,905	855
30–39	79,752	812	87,654	814
40–49	44,565	824	51,827	819
50–59	49,016	778	62,537	756
60–64	19,886	704	26,324	686
65 or older	14,506	487	5,945	593
Work incentives				
Section 1619(a)	14,443	1,453	15,476	1,400
Section 1619(b)	94,708	1,631	99,106	1,659
Plan to achieve self-support (PASS) ^b	204	1,092	331	947
Impairment-related work expenses (IRWE)	2,942	852	2,907	512
Blind work expenses (BWE)	955	1,331	807	941

¹ DAF statistics are limited to beneficiaries whose payment status is C01, E01 and N01; we have not verified the selection criteria used in the SSA published statistics. Additionally, the DAF is limited to beneficiaries under SSA's Full Retirement Age (66 in 2018), so the overall samples in the published statistics and DAF18 are not strictly comparable. Before limiting to the DAF sample, we identified 386,801 beneficiaries with earnings in December 2018, and an average dollar amount of \$771.

Table 2. SSI Beneficiary Earnings in December, by Year and Source

	Number of Beneficiaries with Earnings ¹	Average Earnings (\$)
December 2018		
SSA published statistics	342,046	783
DAF18, SSR	373,062	783
DAF18, DCF	182,383	770
December 2017²		
SSA published statistics	342,217	715
DAF17, DCF (not included in the DAF)	203,582	789
December 2016²		
SSA published statistics	336,837	654
DAF16, DCF (not included in the DAF)	220,195	777
December 2015		
SSA published statistics	328,008	618
DAF15, DCF (included in the DAF)	232,551	734

¹ DAF statistics are limited to beneficiaries whose payment status is C01, E01 and N01; we have not verified the selection criteria used in the SSA published statistics.

² The DAF16 and DAF17 used SSR earnings for SSI, but because of the propagation issue identified above were not reliable for this analysis.