M50 MATHEMATICA CENTER FOR STUDYING DISABILITY POLICY

REPORT

Volume 3: Tips for Conducting Analysis with the DAF16

May 18, 2018

Michelle Lee
Dawn Phelps
Jody Schimmel Hyde

Submitted to:
Social Security Administration
Office of Retirement and Disability Policy
Office of Program Development and Research
Washington, DC 20024-2796
Project Officers: Paul O'Leary and Debra Tidwell-Peters
Contract Number: SS00-16-60003

Submitted by:
Mathematica Policy Research
1100 1st Street, NE
12th Floor
Washington, DC 20002-4221
Telephone: (202) 484-9220
Facsimile: (202) 863-1763
Project Director: Cassandra Pickens Jewell
Reference Number: 50214.02.260.490.000

Suggested Citation: Lee, Michelle, Dawn Phelps, Jody Schimmel Hyde. "Disability Analysis File 2016 (DAF16) Documentation: Data from January 1994 through December 2016. Volume 3: Tips for Conducting Analysis with the DAF16" Washington, DC: Mathematica Policy Research, May 2018.

CONTENTS

TABLES

This page has been left blank for double-sided copying.

GLOSSARY

ADM	Awardee Data Mart
AIME	Average Indexed Monthly Earnings
B.E.S.T.	Benefits Entitlement Services Team
BFW	Benefits forgone due to work
BIC	Beneficiary Identification Code
BMF	Budget Month Factor
BOAN	Beneficiary's Own Account Number
BOND	Benefit Offset National Demonstration
CAN	Claim Account Number
CDR	Continuing Disability Review
CDRCF	CDR Control File
CER	Characteristics Extract Record 100% Field File
COLA	Cost-of-Living Adjustment
DAC	Disabled Adult Child
DAF	Disability Analysis File (previously known as TRF)
DBAD	Disabled Beneficiary and Dependents Extract
DCF	Disability Control File
DDS	Disability Determination Services
DER	Detailed Earnings Record
DI	Disability Insurance, also referred to as SSDI
DMG	Demographic component of the DAF
DSN	Dataset names

DWB	Disabled Widow Beneficiaries
EN	Employment Network (also called a TTW provider)
EPE	Extended Period of Eligibility
EVS	Enumeration Verification System
EXR	Expedited Reinstatement
FBR	Federal Benefit Rate
FCI	Federal Countable Income
FIPS	Federal Information Processing Standards (in reference to U.S. Census standardized codes for uniform identification of geographic entities)
FRA	Full Retirement Age
HI	Hospital Insurance (Medicare Part A)
HUN	Housed Under Number
ICD-9	International Classification of Diseases Coding Scheme
IPE	Individualized Plan for Employment, developed by SVR Agency
IRS	Internal Revenue Service
IRWE	Impairment-Related Work Expense
LAF	Ledger Account File
LAUS	Local Area Unemployment Statistics
LRF	Longitudinal Record Format
MBR	Master Beneficiary Record
MBR810	MBR extract, version number 810
MBR814	MBR extract, version number 814
MEF	Master Earnings File
MIE	Medical Improvement Expected

MO	Milestone + Outcomes payment system
MPR-EVS	Mathematica's EVS
NBS	National Beneficiary Survey
NSCF	National Survey of SSI Children and Families
NUMIDENT	Numerical Identification File
OIM	Office of Information Management
OO	Outcomes-Only payment system
PAN	Person's Account Number
PASS	Program to Achieve Self-Support
PHUS	Payment History Update System
PIA	Primary Insurance Amount
PIN	Personal Identification Number
POMS	SSA's Program Operations Manual System
Provider	Service provider under TTW (also called an EN)
REMICS	Revised Management Information Counts System
RIB	Retirement Insurance Benefits
RMA	Retrospective Monthly Accounting
RSA	Rehabilitation Services Administration
RSA-911	RSA Case Service Report
SAIPE	Small Area Income and Poverty Estimates
SAS	Statistical Analysis Software, used to produce the DAF
SCWF	Standalone Companion Work File
SER	Summary Earnings Record

SGA	Substantial Gainful Activity
SMI	Supplemental Medical Insurance (Medicare Part B)
SNAP	Supplemental Nutrition Assistance Program
SSN	Social Security Number
SSA	Social Security Administration
SSDI	Social Security Disability Insurance (also referred to as DI)
SSI	Supplemental Security Income
SSI-LF	SSI - Longitudinal File
SSR	Supplemental Security Record
STW	Suspension or termination of cash benefits for work
SVR Agency	State Vocational Rehabilitation Agency
T2	Title II, the SSDI Program
T16	Title XVI, the SSI Program
TANF	Temporary Assistance for Needy Families
TCNEI	Total countable non-earned income
TKT	DAF component containing data related to TTW participation
TRF	Ticket Research File, now called the DAF
TTW	Ticket to Work
TWP	Trial Work Period
VR	Federal/State Vocational Rehabilitation program
VRRMS	Vocational Rehabilitation Reimbursement Management System; data from this system is contained in the Payments component
YTD	Youth Transition Demonstration

OVERVIEW OF DAF DOCUMENTATION

The documentation for the DAF consists of the eleven volumes described below. Questions about these documents should be directed to ORDES.DAF@ssa.gov. All of these documents are available at https://www.ssa.gov/disabilityresearch/daf.html.

- **Volume 1: Getting Started with the DAF16.** Provides an overview of the structure and contents of the DAF and related linkable files.
- **Volume 2: Working with the DAF16.** Contains practical suggestions such as how to extract data and interpret blank or missing variables as well as more detailed information on DAF data marts and linkable files.
- **Volume 3: Tips for Conducting Analysis with the DAF16.** Contains suggestions for working with common research concepts in the DAF such as program participation, benefits paid versus benefits due, and constructed measures related to STW and BFW.
- **Volume 4: Lists of DAF16 Variables.** Contains lists of new, changed, and deleted variables, as well as lists of variables by DAF component and analytic category.
- **Volume 5: DAF Variable Detail Pages.** Contains specifications for each DAF variable, including name, definition, data format, identification of the DAF component to which it belongs, data source, availability, and (where applicable) SAS code used to construct the variable.
- **Volume 6: Validating the DAF16 Against Other Sources.** An explanation of validation methods as well as tables and charts comparing statistics computed from the DAF to SSA published statistics.
- **Volume 7: DAF16 Development History and Construction Methods.** Describes key changes in DAF construction methodology over time as well as a description of each step in the current year DAF construction process.
- **Volume 8: DAF16 Construction Workflow Charts and Task Tables.** Provides detailed information in both chart and table format on each step in the current year DAF construction process.
- **Volume 9: DAF16 Source File Descriptions.** Describes the administrative source files used to construct the DAF.
- **Volume 10: DAF Administrative Source File Documentation.** Contains documentation from SSA or other agencies on the administrative source files described in Volume 9.
- **Volume 11: DAF16 Construction Code and Data Mart Details**. Contains all SAS code used to construct the DAF.

The following table provides specific locations for common research-related questions and issues.

In order to …	Refer to …
Get started with a research task	Volume 2, "Working with the DAF16," for information about selecting beneficiaries using finder files versus selection criteria
Identify what's changed in the latest version of the DAF	Volume 1, "Getting Started with the DAF16"
View lists of DAF variables	Volume 4, "Lists of DAF16 Variables"
Understand individual variable definitions, specifications, and value ranges	Volume 5, "DAF16 Variable Detail Pages"
Understand the structure of the DAF data files at a high level	Volume 1, "Getting Started with the DAF16"
Identify variables for a specific research task	Volume 4, "Lists of DAF16 Variables," for a list of variables contained within each DAF file and by analytic category
Understand the beneficiaries for which the DAF does and does not contain data	Volume 1, "Getting Started with the DAF16"
Identify SSA administrative data sources for the DAF	Volume 9, "DAF16 Source File Descriptions"
Generate ideas for using the DAF more efficiently	Volume 1, "Getting Started with the DAF16" and Volume 2, "Working with the DAF16"
Find suggested ways to identify common research concepts in the DAF, such as calculating age of retirement, or disability title	Volume 3, "Tips for Conducting Analysis with the DAF16"
Understand what variables have changed in the most recent DAF	Volume 4, "Lists of DAF16 Variables"
Read about how information in the DAF is validated against other sources	Volume 6, "Validating the DAF16 Against Other Sources"

I. CALCULATING AGE-BASED MEASURES

A. Calculating age at a point in time

While birthdate information is available from the DAF, there is no age variable in the database, meaning that users must define it using date of birth relative to another point in time. As discussed in Volume 2, DOBBEST is the most accurate DAF variable for date of birth. Age can simply be calculated as reference date minus DOBBEST. One common method for performing this calculation is using a SAS INTCK statement such as *age=INTCK('Months',DOBBEST,ref date,'C')/12*. This statement will count the number of months between DOBBEST and the reference data and then divide by 12 to get an age in years and whole month fractions. This method is particularly useful because it can be used to compare age to FRA, which is calculated in two month intervals based on birth year.

Selection of a reference date or dates depends on the research design. In many cases, users just want to calculate age on a particular calendar date. In other cases, users want to calculate age at program entry. In this case, users would define the reference date from the variable they choose to use as the point of benefit start (discussed elsewhere in this volume). In other cases, users want to calculate age over a period such as a calendar year. This may be the case if one wants to create an extract of beneficiaries between the ages of 18 and 65 during the calendar year 2010. In this case, the user should calculate two age variables using two reference dates—January 1, 2010 and December 31, 2010—then keep all records that meet the age selection criteria at either time.

Users should keep in mind that the methods described for coding age at a point in time will code valid age values for deceased beneficiaries. To code age as missing for deceased beneficiaries, use the variable DODBEST. If it is blank or missing, no death date has yet been

recorded for the beneficiary. If it is populated and has a value that is earlier than the reference date, code age as missing.

B. Determining FRA

A beneficiary's FRA for purposes of Social Security benefits was previously 65 years, but amendments to the Social Security program in 1983 gradually increased FRA from 65 years, for those born in 1937 or earlier, to 67 years, for those born in or after 1960 (Table I.1). There are two variables on the DAF related to FRA. The first is FRA, the beneficiary's full retirement age as defined by the table below. The second is FRADATE, which is the date the beneficiary reaches FRA. Note that for beneficiaries born on January 1 of any year, SSA considers the previous year as their year of birth.

Table I.1. Retirement age, by year of birth

Year of Birth	Full Retirement Age
1937 or earlier	65
1938	65 + 2 months
1939	65 + 4 months
1940	65 + 6 months
1941	65 + 8 months
1942	65 + 10 months
1943-1954	66
1955	66 + 2 months
1956	66 + 4 months
1957	66 + 6 months
1958	66 + 8 months
1959	66 + 10 months
1960 and later	67

Note: If beneficiary was born on January 1 of any year, set Year of Birth as the previous year.

II. PROGRAM PARTICIPATION AND DISABILITY ONSET

To determine SSDI and SSI program participation at a given point in time, researchers often consult monthly payment status variables in the DAF along with one-time measures such as date of first payment receipt, first entitlement, and disability onset. Additionally, it is important that researchers use the appropriate variables for SSDI and SSI; variables are different for SSDI and SSI and are indicated throughout the chapter.

Note that in this chapter we are discussing program participation in terms of eligibility (i.e., entitlement to payment), not actual payment. There are many reasons why a beneficiary may have been eligible for but did not receive a payment in a given month, including recovery of an overpayment previously incurred and months prior to a favorable disability decision being made for which eligibility is retroactively determined. Similarly, it is possible for beneficiaries to have received payments in a month when no benefit payment was due in that month, such as when an underpayment for prior months has been retroactively determined but there is no current-month eligibility.

By description, there are numerous DAF variables that appear to delineate a beneficiary's program participation, but researchers should exercise caution and ensure that they fully understand these variables before basing analyses on them. For example, the DMG component variable DOEC (SSDI Most Recent Entitlement Date) is the most recent entitlement date to SSDI, but it remains populated even after benefits have terminated and so cannot be used to determine ongoing eligibility. However, it can generally be used with DOST (Date of Suspension or Termination) to identify the most recent period of eligibility. DOEI (Date of Initial Entitlement) gives the first date of entitlement to SSDI benefits. If there are multiple periods of eligibility these variables cannot be used in combination to identify the dates of all

periods of eligibility because all three—DOEC, DOST, and DOEI—are one-time only variables. In addition, DOEI may identify a non-disability-based period of entitlement, such as a person who initially receives Social Security early retirement benefits prior to receiving an SSDI award, or one who received auxiliary benefits, such as a dependent child, before receiving SSDI benefits.

Similar issues exist for comparable SSI variables such as ELG_RD (most recent date of SSI eligibility), SSIELIG_FIRST (earliest available SSI eligibility date) and CURSTAT (most recent SSI payment status). These issues are compounded by the fact that SSI data are maintained primarily in a multiple-record format of the SSR rather than the single-record format of the MBR for SSDI. Data are sometimes overwritten when new SSR records are created, which is why, for example, SSIELIG_FIRST is characterized as the earliest *available* date rather than the earliest actual date.

Absent compelling reasons to the contrary, researchers are therefore advised to follow the guidance outlined in this chapter to determine program participation. The variables SSDIyy and SSIyy have been constructed for researchers who only need to identify program participation on an annual basis. These binary flag variables indicate whether or not a beneficiary was in current pay status as defined below for any month in a given year.

A. Determining monthly participation based on payment status variables

Receipt of SSDI or SSI can vary from month to month as participants gain and lose their eligibility to benefits according to changing circumstances such as incarceration or a return to employment. In other words, once eligible for benefits, a participant will not remain eligible indefinitely. The DAF contains monthly variables that indicate whether a beneficiary was entitled to payments for a given month, regardless of whether or not that payment was actually received by the beneficiary in that month. These variables are LAFyymm for SSDI and

PSTAyymm for SSI. After reading this section, users interested in identifying beneficiaries during a certain period of time may want to consult the DAF code library, available at https://www.ssa.gov/disabilityresearch/daf.html#library, which provides sample code for that task.

1. SSDI beneficiaries

To determine program participation status based on eligibility for an SSDI beneficiary, use the SSDI payment status code variable (also known as the "ledger account file status" variable), LAFyymm, for a specific month or range of months. For example, to determine participation status for an SSDI beneficiary in 2003, use LAF0301 through LAF0312. Current pay status is indicated by a value beginning with C or E, with C being much more common (E only applies to beneficiaries of the Railroad Retirement Board). If a beneficiary was in current pay status for January through October 2003 then terminated from the SSDI program in November 2003 due to death (for example), the LAFyymm variables for the year 2003 would contain values as shown in Table II.1, with a LAF value of T1 referring to termination due to death. See the LAFyymm entry in Volume 5 for a full description of possible variable values.

Table II.1. Illustration of monthly participation variables for SSDI beneficiaries

Variable	LAF 0301	LAF 0302	LAF 0303	LAF 0304	LAF 0305	LAF 0306	LAF 0307	LAF 0308	LAF 0309	LAF 0310	LAF 0311	LAF 0312
Value	C	C	C	C	C	C	C	C	C	C	T1	T1

Beginning in DAF16, the DMG file includes variables that allow users to identify the most recent LAF status of a beneficiary. LAF_MR indicates the most recent LAF code for the beneficiary, while the LAF_MR_DT indicates the month in which that most recent status occurred. This means that a user wanting to select SSDI beneficiaries in current pay status as of a

particular month should be able to do so using the DMG rather than looping through the *yymm* variables on each of the Annuals.

The STWDIyymm variable can also be used to establish eligibility in a month or range of months. Months in which beneficiaries have a status of "0" for this variable are months in which the beneficiary was eligible for payment. See Chapter VIII in this volume for additional information on how to use this variable.

2. SSI beneficiaries

To determine the participation status for an SSI beneficiary, use the SSI payment status code variable, PSTAyymm, for a specific month or range of months. For example, to determine participation status for an SSI beneficiary in 2003, use PSTA0301 through PSTA0312. Current pay status is indicated by a value of C01.[1] If a beneficiary was in current pay status for January thru October 2003 then suspended from the SSI program in November 2003 as a result of incarceration (for example, PSTA value of N22), the PSTAyymm variables for the year 2003 could contain values as shown in the table below. See the PSTAyymm entry in Volume 5 for a list of possible variable values.

Table II.2. Illustration of monthly participation variables for SSI beneficiaries

Variable	PSTA 0301	PSTA 0302	PSTA 0303	PSTA 0304	PSTA 0305	PSTA 0306	PSTA 0307	PSTA 0308	PSTA 0309	PSTA 0310	PSTA 0311	PSTA 0312
Value	C01	C01	C01	C01	C01	C01	C01	C01	C01	C01	N22	N22

The STWSSIyymm variable can also be used to establish eligibility in a month or range of months. Months in which beneficiaries have a status of "0" for this variable are months in which

[1] In addition, payment status codes of M01 and M02 indicate that eligibility has been manually determined. Although M01 and M02 are sometimes associated with ineligibility, this can be difficult to determine because of how manual computations display (or do not display) in the administrative data. Many researchers consider M01 and M02 payment status codes to indicate program participation analogously to C01 because these codes generally indicate manually computed eligibility rather than ineligibility.

the beneficiary was eligible for payment. See Chapter VIII in this volume for additional information on how to use this variable.

B. Determining initial benefit entitlement date

Conceptually, the initial benefit entitlement is the month for which the beneficiary is first due payments. This is often different than the first benefit payment date because when beneficiaries are first deemed eligible, they may receive retroactive benefits for some of the months since their disability onset date. For example, someone receiving benefit payment for the first time in September 2004 may also receive retroactive payments for the previous three months at the same time as the September benefit payment is made. In this case, the initial benefit entitlement date would be June 2004, while the first payment date would be September 2004. The payment status codes described above will reflect eligibility in a situation such as this one without regard to whether a payment was actually made in that month.

Two constructed variables on the DAF16 ADM–DIELIGDT for SSDI beneficiaries and SSIELIGDT for SSI beneficiaries–identify the first month of eligibility for benefits. These variables, however, are only available for beneficiaries on the ADM, i.e., only beneficiaries who receive their *first benefit payment* as a non-retiree adult in 1996 or later. Because beneficiaries who received their first payment before 1996 are excluded from the ADM, the ADM contains approximately 10 million fewer beneficiaries than the full DAF DMG.

In addition to data in the ADM, variables from the MBR in the DMG file can identify additional entitlement dates for SSDI beneficiaries, but should be used with caution. There are two variables that could be of interest: 1) the DOEI ("Date of Initial Entitlement") variable and 2) the earliest date stored in the variable series ENTDATn, which records entitlement dates for up to twelve possible periods of eligibility. In the vast majority of MBR records, DOEI and the earliest value from ENTDATn contain the same date. In instances when the dates do not match,

there does not seem to be a consistent logic for the difference, though they may make sense in particular analytical circumstances. For SSI beneficiaries, there is no comparable entitlement date in the administrative data. SSI eligibility can be determined by evaluating monthly SSI program status codes (PSTA) for current pay.

C. Identifying first benefit payment date

The first payment date for SSDI or SSI benefits is often used to determine when a beneficiary began participating in the disability program. As described above, in many cases for SSDI beneficiaries, this is different than the date of initial benefit eligibility. Two variables on the DAF16 ADM identify the first benefit payment date received as an adult under SSDI or SSI, if that date was in 1996 or later: DIAWARDDT and SSIAWARDDT, respectively.

The variable MINFRSTP (SSI Earliest Available Application, First SSI Payment Date) contains the initial SSI payment date for the earliest available entitlement period, regardless of adult or child status. However a word of caution regarding MINFRSTP—SSA periodically overwrites the data from which this variable is drawn (FIRST-PAY-DTE in the SSR) and sometimes the initial payment date for the first entitlement period is replaced by the initial payment date for a subsequent entitlement period. To illustrate, suppose a beneficiary has three entitlement periods on record in the SSR spanning a period of twenty-five years. In this example, the first entitlement period began twenty-five years ago, the second began twenty years ago, and the third began two years ago. The date of the initial payment for all three entitlement periods is stored in the SSR, and the earliest occurring initial payment date represents the initial payment date for the first entitlement record, twenty-five years ago. Later, the SSR data are updated, and perhaps because the earliest entitlement period began such a long time ago, the update routine overwrites the data for the earliest entitlement period. After the overwrite all that remains are the two most recent entitlement periods, with the initial payment date for the first entitlement period

occurring twenty years ago. In the absence of any other information, it appears that this beneficiary's initial payment date occurred twenty years ago, not twenty-five years ago. In actuality, the earliest initial payment date from the SSR reflects the earliest **available** payment date, which is not necessarily the date the SSI beneficiary first received a payment.

For SSDI beneficiaries there is no administrative variable that identifies first payment. Instead, the first SSDI payment for a beneficiary is determined by examining the LAFyymm and PAYDyymm variables for the first month, where the beneficiary is in current pay status and has a payment. Again, caution should be used. The first payment date under SSDI, could have been for a child auxiliary payment unrelated to a later period of adult disability.

D. Determining disability onset date

Researchers are sometimes interested in the date a beneficiary first had an onset of disability. In most cases, this is distinct from the date of initial benefit eligibility and the first payment date. To determine the earliest date of onset of a beneficiary's date of disability:

- for SSDI beneficiaries, use the first occurrence of DDOn (SSDI Disability Onset Date), where n ranges from 1 to 12; and
- for SSI beneficiaries, use SSIELIG_FIRST (SSI Earliest Eligibility Date (Constructed))—note that this variable contains the earliest available eligibility date; in some cases data on earlier periods of eligibility, particularly if they were far in the past, are overwritten.

This page has been left blank for double-sided copying.

III. BENEFIT AMOUNT DATA

Monthly benefit amount data are available separately for SSDI and SSI benefits. Categories of benefit variables include the amount due to the beneficiary, the amount paid, the federal or state dollar portion of the SSI benefit payment, and the amount due or paid for dependents. The amount paid may differ from the amount due as deductions may be applied to account for a variety of factors including Medicare premiums or previous overpayments. The benefit amount variables require special handling for accurate use and details are provided below.

A. Formats for amount variables

A variety of currency formats are in use for benefit and earnings amounts variables. SSDI benefit amount variables generally include cents, and use SAS formats such as "$$$$.c" or "$$$$.cc". SSI benefit amount variables do not include cents, and use SAS formats such as "$$$$$$" or "$$$". Other variables related to SSI earnings, such as IEA_BLINDyymm (SSI earned income amount – blind work expense) and IEA_PASSyymm (SSI earned income exclusion amount – PASS), do include cents, using SAS formats such as "$$$$.cc."

B. SSDI benefit due variables

There are three variables for SSDI benefits due:

- **MBAyymm–Federal SSDI Benefit Due.** Monthly Benefit Amount (MBA) is the benefit amount due to the beneficiary *before* rounding or adjustments, and can be thought of as the "base" amount due. The source variable is the SSA variable MBA from the MBR file.
- **DUEDyymm–SSDI Benefit Due.** DUED is the benefit amount due to the beneficiary *after* rounding to the nearest lower dollar but *before* adjustments are made. SSA refers to this as "amount credited" and it is the basic benefit amount field to use for analysis as other amount fields are affected by factors such as overpayments or deductions for Medicare. Mathematica uses DUED to determine a beneficiary's total benefit amount for a given month. The source variable is the SSA variable MBC (Monthly Benefit Credited) from the MBR file
- **DUEOyymm–SSDI Dependent Benefit Due.** This variable is similar to DUEDyym but represents the benefit amount due to the beneficiary's dependents who are entitled to cash benefits as a result of the beneficiary's entitlement, *after* rounding to the nearest lower dollar

but *before* adjustments are made. It is based on the same SSA source variable, MBC from the MBR file, and is calculated by summing only the benefit amounts due to dependents of the beneficiary. It excludes benefits due to the beneficiary but includes all benefits payable to dependents, including DACs and DWBs.

To accurately determine the monthly benefit amount due to an SSDI beneficiary, it is not enough to use only the benefit due variables—you must also use the payment status code (also known as the "ledger account file status," or LAF variable) for the corresponding month, LAFyymm. Our investigations of the SSDI benefit amount due variables and ensuing discussions with SSA staff revealed that the SSDI benefit due variables are often populated even when the beneficiary is not in current pay status. Such situations occur because SSA computer systems often store placeholder values in the benefit due variables for future months and these placeholders are not always removed from the variables when a beneficiary changes from current pay status. Therefore, to accurately determine the amount due for an SSDI beneficiary, SSA staff advise first determining whether a beneficiary is in current pay status in the month in question using the LAF variable for the corresponding month and year. LAF values of "C" or "E" in the first position indicate the beneficiary is in current pay status.

C. SSDI benefit paid variables

Benefit paid variables reflect payments made in a given month and are conceptually distinct from benefit due variables. For example, a positive value for a benefit paid variable in a given month does not mean the beneficiary is entitled to benefits for that month; the amount in PAYDyymm may reflect a delayed entitlement for a prior month. To determine entitlement for a specific month, use the benefit due variables described in the previous section (MBAyymm, DUEDyymm, DUEOyymm). Unlike benefit due variables, benefit paid variables are only populated when the beneficiary received a payment. There are three variables for SSDI benefits paid:

- **MBPyymm—Federal SSDI Benefit Paid:** MBP represents the benefit amount payable to the beneficiary. The source variable is the SSA variable MBO, which is based on the variable MBC that is used to generate DUEDyymm (see above), with adjustments such as Medicare premiums and previous under or overpayments. It may differ from the actual check amount paid to the beneficiary (PAYDyymm); the latter may include under or overpayments.[2]
- **PAYDyymm—PHUS SSDI Benefit Paid:** PAYD reports the actual amount paid to the beneficiary in a given month. It is derived from two SSA variables: Direct Pay (DP—the actual amount of Social Security benefits disbursed for a specific month) minus Medicare Premium (MD – Part A and Part B premiums) from the PHUS.
- **PAYOyymm—PHUS SSDI Dependent Benefit Paid:** PAYO reports the actual amount paid to the beneficiary's dependents. It is similar to PAYDyym and is based on the same SSA source variables, DP and MD, from the PHUS. It is calculated by summing only the check amounts paid for qualifying dependents of the beneficiary. It excludes check amounts for the beneficiary, as those check amounts are available in the PAYDyymm variable for the individual in question. It includes dependents, such as DACs and DWBs.

D. SSI benefit due variables

Unlike the SSDI variables, the SSI benefit due variables are populated only when a payment is actually made, so there is no need to determine whether a beneficiary was in current pay status before using these variables. Benefit due variables for SSI beneficiaries include the following:

- **FAMTyymm—Federal SSI Benefit Due.** FAMT reports the federal benefit amount for which the SSI beneficiary is eligible. FAMT can be retroactively adjusted to reflect under or overpayments as a result of changes in income or other factors such as living arrangement that affect benefit amount. The SSA source variable is FEDAMT (also called FAMT) from the SSI-LF.
- **SAMTyymm—State SSI Benefit Due.** SAMT reports any federally-administered state SSI benefit the beneficiary is eligible to receive. It can be retroactively adjusted to reflect under or overpayments as a result of changes in income or other factors such as living arrangement that affect benefit amount. The SSA source variable is SUPAMT (also called SAMT) from the SSI-LF.
- **DUESyymm—SSI Benefit Due.** DUES (the sum of FAMT and SAMT) reports the total federal and federally administered state benefit amount for which the SSI beneficiary is eligible. It can be retroactively adjusted to reflect under or overpayments as a result of changes in income or other factors such as living arrangement that affect benefit amount. The SSA source variables are FEDAMT and SUPAMT (also called FAMT and SAMT) from the SSI-LF.

[2] Researchers should note that, over time, MBPyymm is overwritten as retroactive adjustments are made.

E. SSI benefit paid variables

The SSI benefit actually paid may differ from the amount due because retroactive changes may be applied to the amount due to reflect changes in income or other factors that affect benefit amount, such as living arrangements or payment withheld to recover an overpayment. Similarly, the amount paid in a given month may differ from the amount due in that month, due to under or overpayments in prior months. There are three SSI benefit paid variables:

- **FPMTyymm—Federal SSI Benefit Paid:** FPMT reports the federal benefit amount paid to the SSI beneficiary. The SSA source variable is FEDPMT (also called FPMT) from the SSI-LF. Unlike FAMTyymm, FPMT is not retroactively adjusted.
- **SPMTyymm—State SSI Benefit Paid:** SPMT reports the federally-administered state SSI supplement paid to the SSI beneficiary. The source is the SSA variable STATPMT (also called SPMT) from the SSI-LF. Unlike SAMTyymm, it is not retroactively adjusted.
- **PAYSyymm—SSI Benefit Paid:** PAYS reports the total federal and state benefit amount paid to the SSI beneficiary. It is the sum of the SSA source variables FEDPMT and STATPMT (also called FPMT and SPMT) from the SSI-LF. Unlike DUESyymm, this variable is not retroactively adjusted.

IV. VARIABLES RELATED TO BENEFICIARY INCOME

There are a variety of useful income variables in the DAF, though these data are generally less reliable than the official earnings data recorded by the IRS. For example, in the early years of the DAF, monthly income amounts of $201 occur frequently for SSDI beneficiaries because the TWP limit was $200 and SSA staff did not distinguish specific levels of earnings so long as they were above that amount. Yet, for both SSI and SSDI, DAF earnings variables offer a significant advantage over the IRS data in that they provide monthly rather than annual data.

Researchers working with DAF earnings data should pay particular attention to the limits of the data. For example, when SSDI beneficiaries have low levels of earnings, they are often not recorded or requested by SSA. In addition, SSDI earnings may not be present in the DCF variables described below when no benefit is due. These limitations are particularly evident in earlier years covered by the DAF. On the other hand, income information has long been collected from SSI beneficiaries because the SSI benefit amount varies according to income. For this reason, we begin by discussing SSI-related income variables.

A. Income information for SSI beneficiaries

The income received by SSI beneficiaries are generally self-reported and, where possible, verified by SSA staff against pay stubs. Accurate and timely information about SSI beneficiary income is important because the monthly SSI benefit paid is a calculated as a function of income received, with different forms of income being accounted for in different ways, per program rules. As such, the administrative data capturing SSI beneficiary income—and hence, the DAF—contain variables both for the total amount of income received and the amount of income actually counted in calculating the monthly SSI benefit amount.

Income received by SSI beneficiaries is recorded in the month in which the earnings are received. For example, a beneficiary who works in May and is paid on June 5th will have June earnings recorded in the SSI data. Often the initial income report is for projected income, and may be subsequently revised if the amount actually earned differs from the projection.

In addition to capturing countable and total income, the DAF contains two sets of SSI earnings variables. Researchers should be aware that neither set of income variables is of the same overall quality as the official data recorded by the IRS. However, the monthly level of the data combined with the fact that these variables comprise the actual basis for SSA adjudications of eligibility and payment may make them well suited for some analyses related to SSA program eligibility.

The first set of SSI earnings information comes from the CER100% Field File, snapshot files that are not revised as new information becomes available. Because SSI income variables are subject to frequent retrospective revision as new information is discovered, investigated, and verified, researchers should be aware that the CER source file does not reflect these revisions. Yet, for researchers interested what was known by SSA at the time these benefits were paid, these source variables are particularly useful and will align with data on benefits paid.

The second set of SSI earnings variables in the DAF comes from the SSR; the SSR captures retrospective updates and therefore reflects the earnings situation actually experienced by the beneficiary in the month, after SSA review. DAF16 is the first version of the DAF with SSI earnings data from the SSR. Prior versions of the DAF instead used data derived from the DCF. While the SSR and DCF values align in most cases, in certain instances, the DCF undercounted beneficiary income. As such, beginning in DAF16, income measures reflecting retroactive updates for SSI beneficiaries are from the SSR.

Table IV.1 identifies the SSI earnings variables from the CER100% file, the SSR, and the DCF. We provide the DCF variables for users more familiar with using them, either directly from the DCF or from earlier versions of the DAF.

Table IV.1. SSI earnings measures from various SSA administrative data files, as named in the DAF

	DAF variable name		
Earnings/Exclusion Description	**Source: CER100%**	**Source: DCF[1]**	**Source: SSR**
Earned income	EARNyymm		
Countable earned income			EICMyymm
Countable unearned income			UINCyymm
Wages/earnings	IEA_WAGEyymm	T16GRSAMTyymm	T16GRSAMTyymm
Self-employment	IEA_NESEyymm	T16NETAMTyymm[2]	T16SEAMTyymm
Net loss from self-employment	IEA_LOSSyymm		T16NETAMTyymm
Student earned income exclusion (SEIE)	IEA_SEIEyymm[3]	T16EXLAMTyymm	T16EXLAMTyymm
Blind work expense (Blind IRWE)	IEA_BLINDyymm	T16EXPAMTyymm[4]	T16BEXPAMTyymm
Disabled workers impairment-related work expenses (IRWE)	IEA_IRWEyymm		T16EXPAMTyymm
Program to Achieve Self Support (PASS)	IEA_PASSyymm	T16PASAMTyymm	T16PASAMTyymm
Unearned income from SSDI benefit	IUA_SSDIyymm		
Unearned income from workers' compensation	IUA_WCyymm		
Unearned income from TANF	IUA_TANFyymm		

[1]Note: SSI income variables from the DCF are no longer included on the DAF as of DAF16; we provide these variable names for users familiar with older versions of the file.

[2] The T16NETAMTyymm variable contained in the DCF was a composite measure of self-employment income (T16SEAMTyymm) and net loss from self-employment (T16NETAMTyymm); beneficiaries in the SSR may have one, though not both, of those values populated in a given month.

[3] This variable is available in 2005 onward.

[4] The T16EXPAMTyymm variable contained in the DCF was a composite measure of Blind work expenses (T16BEXPAMTyymm) and non-blind impairment-related work expenses (IRWE—T16EXPAMTyymm); beneficiaries in the SSR may have one, though not both, of those values populated in a given month.

Beginning in DAF16, we added an indicator variable developed to provide summary information about the earnings of SSI beneficiaries, derived from the SSR earnings variables. This measure was designed to be similar to the earnings indicator that was added in the previous version of DAF for SSDI beneficiaries. The SSI measure is called SSIERNLVLyymm and takes on the following values:

- 0=gross earnings are zero

- L=gross earnings are greater than 0 but less than the federal benefit rate (FBR)
- F=gross earnings are at or above FBR but lower than SGA, and earnings less exclusions are less than FBR
- B=gross earnings are at or above FBR but lower than SGA, and earnings less exclusions are greater than FBR
- A=gross earnings are at or above SGA but net earnings are less than FBR
- G=gross earnings are at or above SGA, and earnings less exclusions are greater than FBR but less than SGA
- S=earnings less exclusions are at or above SGA
- The values for SSIERNLVL are determined by summing the IEAMT for IETYP=W and S, then subtracting the IEAMT where IETYP=N to create total SSI Gross Earnings. We compare that value to the FBR and SGA thresholds for the year. We then sum the values of IEAMT, where IETYP=B, C, D, and T, to create total SSI Income Exclusions, subtracting that from SSI Gross Earnings to create earnings less exclusions. We then compare that second value to the FBR and SGA to complete the measure.

B. Income information for SSDI beneficiaries

Although earned income from working affect SSDI benefits, benefit amounts for SSDI are not subject to change as frequently as they are for SSI. Unlike SSI earnings, which are applied based in the month they are paid (regardless of when the work was performed), SSDI earnings have historically been applied based on the month in which they were earned. Using the earlier example, an SSDI beneficiary who works in May and is paid on June 5th will have May earnings recorded in the SSDI data. This changed in September 2016, as SSA enacts the provisions of the Bipartisan Budget Act of 2015.[3] Starting in September 2016, SSDI earnings will reflect a mix of when they were earned and when they were paid. Based on the change over time in SSDI earnings data, our recommendation is to use the SSI earnings data (and associated timing of

[3] Section 825 of the Bipartisan Budget Act of 2015 (PL 114-74) that was effective September 23, 2016 requires that when determining if an individual is engaged in substantial gainful activity to determine SSDI eligibility, SSA considers earnings to have been earned: (1) for the purposes of initial eligibility in the month in which such services were performed and (2) in all other cases, in the month such earnings were paid. Unless SSA can establish based on readily-available evidence the actual month such earnings were earned or the person (after having been denied benefits due to excess earnings) can demonstrate to SSA that such earnings were actually earned in a different month, the month they were paid will be used.

earnings) for beneficiaries receiving SSI only and concurrently receiving SSI and SSDI. Users should also assume that after September 23, 2016, SSDI earnings will reflect when they were paid, similar to the method used for SSI. Still, it is important to recognize that this will not be the case in all instances, meaning that annual earnings recorded in the DAF for SSDI beneficiaries will not necessarily match IRS earnings.

SSDI contains two sets of earnings variables, one that reflects the earnings received during SSDI entitlement, and another that is based on historical earnings prior to entitlement. The latter is necessary for determining eligibility, while the former is used by SSA in determining the monthly benefit amount.

Current monthly income information (during SSDI entitlement): SSDI income amounts are derived from the DCF and include:[4]

- T2GRSAMTyymm: SSDI Earnings Gross Amount (DCF)
- T2NETAMTyymm: SSDI Self-Employment Net Income Amount (DCF)
- ALLGAMTyymm: Alleged Earnings Amount (DCF—captured both for SSDI and SSI beneficiaries)

Work-related expenses and income exclusions are reported in:

- T2CDNAMTyymm: SSDI Special Condition Amount (DCF)
- T2EXPAMTyymm: SSDI Work Expense Amount (DCF)
- T2SBDYAMTyymm: SSDI Earnings Subsidy Amount (DCF)

The earnings variable contained on the DCF resulted from the advent of TTW, which required that SSA track SSDI beneficiaries' current income more closely in order to make payments to providers. These new variables are available in the DAF as early as January 2003.

[4] Note: the issue noted above in which some earnings from the SSR were not recorded in the DCF does not affect the DCF data for SSDI beneficiaries. The DCF remains the best SSA administrative source of earnings data for SSDI beneficiaries.

Some SSA field offices began using these new systems later than others, leading to inconsistencies in the early months of data availability. When possible, the amounts are verified by SSA staff against pay stubs.

Earnings recorded for SSDI beneficiaries may not always be reliable, particularly at lower levels of earnings, which do not affect benefit eligibility. They are reliable at higher earnings levels, after allowing for an administrative processing lag. Moreover, while they are not always consistent with IRS annual earnings data, they are monthly and therefore can be more suitable for analyses of SSA programs, where eligibility is calculated on a monthly basis.

Beginning in DAF15, we added a measure of SSDI earnings that takes into account the variables above to create a single summary measure of earnings relative to the TWP and SGA levels in each month. This new measure is called DIERNLVLyymm (SSDI Earning Level Indicator) and is described in Volume 5. To create this variable, we summed T2GRSAMT and T2NETAMT to create SSDI Gross Earnings. We then subtract the sum of T2CNDAMT, T2EXPAMT and T2BDYAMT from the SSDI Gross Earnings to create SSDI Countable Earnings. Using those values, we generated DIERNLVLyymm by comparing Gross and Countable earnings to the TWP and SGA limits in place during each month. The values of the newly constructed variable are:

- 0=gross earnings are zero
- L=gross earnings are greater than zero, but less than the TWP limit
- T=gross earnings are at or above the TWP limit, but less than SGA
- G=gross earnings are at or above SGA, but countable earnings are less than SGA
- S=countable earnings are at or above SGA

Earnings history information (prior to SSDI entitlement): To qualify for SSDI benefits, a worker must have had a work history with earnings above a certain level for a minimum number of quarters. This earnings history is then indexed to inflation and averaged across the

months covered by the period of the earnings history, a calculation captured in the Indexed Monthly Earnings (IMEn) variable. The highest thirty-five years of IME, in turn, are used to calculate a PIA, which is available on the DAF (PIAn). The monthly benefit amount due derives from and closely tracks the PIA. Historical earnings data are reported in:

- IMEn: SSDI Indexed Monthly Earnings
- PIEDn: SSDI PIA Effective Date

Because indexed monthly earnings are based on the highest 35 years of earnings, each of these variables may have up to 50 occurrences, where "n" represents the specific occurrence, e.g. IME1, IME2, PIED1, PIED2, etc. The PIEDn variable indicates the effective date represented by the IMEn variable. All of these variables are ascending such that the beneficiary's earliest PIA on the DAF is PIA1, their next PIA is PIA2, etc.

Only the 50 most recent occurrences of the variable are retained. More than 99 percent of beneficiaries have 38 or fewer occurrences. However, less than 1 percent of beneficiaries have more than 50 occurrences of these variables. For beneficiaries with more than 50 occurrences, the first occurrence (i.e., IME1, PIA1, PIED1, etc.) corresponds to the total number of occurrences minus 49. For example, the first occurrence for a beneficiary with 73 occurrences will correspond to the 24th occurrence overall and ascend from there.

This page has been left blank for double-sided copying.

V. IMPAIRMENT CODES

Impairment information for beneficiaries in SSDI and SSI is recorded by SSA only at select times—initially at disability award and updated later during periodic medical CDRs, generally scheduled to occur every three or seven years.[5] Primary and secondary impairment codes reflect the evidence used by SSA to determine whether a person is disabled at that time. However, SSA staff generally record only sufficient impairment information to justify disability benefits so if a primary impairment is sufficient to qualify a beneficiary for benefits, a secondary impairment may not be recorded.

When stored in the DAF, these periodic impairment codes are converted to monthly occurrences. The preferred variables to use are DXPRIBESTyymm and DXSECBESTyymm in the DAF Annual files. The impairment code is listed in each "yymm" occurrence in which the diagnosis was in effect. For example, for a beneficiary who became entitled to benefits in January 2002 with a disability due to a musculoskeletal disorder, the variable DXPRIBEST0201 would indicate the relevant code for the musculoskeletal disorder. If the diagnosis was revised during a CDR in January 2005 to a psychiatric disorder, the primary impairment code would be updated in the SSA administrative data, and the DAF variable DXPRIBEST0501 would contain the revised impairment code, while the intervening DAF variables—from DXPRIBEST0202 to DXPRIBEST0412—would contain the initial impairment code. For the months after January 2005, the subsequent variables, DXPRIBEST0502 and onwards, would reflect the revised impairment code, up to the time of the next CDR or the month when the beneficiary's

[5] The term "CDR" also has meaning in the context of a review to determine how an SSDI beneficiary's work affects their benefits. Such a review is called a work CDR as opposed to the medical CDRs under discussion here.

entitlement ceased. Converting the periodic codes to monthly occurrences permits a beneficiary's most recent impairments to be identified at any time while they are entitled to benefits.

A. Coding schemes and categorization

SSA records impairment information using four-digit SSA Impairment Codes, a system of diagnosis codes that has been in use since the mid-1980s. Prior to the advent of the SSA impairment codes currently in use, SSA used the ICD-9 coding scheme. In some cases, the ICD-9 codes are still present in the DAF data because medical data has not been updated, particularly for beneficiaries who have been receiving benefits for a long time. Both coding schemes contain many hundreds of potential impairments and it is therefore useful to group individual diagnosis codes into broader categories for analysis or discussion. There is not currently a single standard at SSA for grouping the diagnosis codes, however, SSA and Mathematica researchers have jointly developed a grouping scheme suitable for research projects that accommodates both the SSA Impairment Codes as well as the ICD-9 codes. This categorization is presented in Table V.1. After reading this section, users interested in consulting the DAF code library, available at https://www.ssa.gov/disabilityresearch/daf.html#library, which contains coding for categorizing based on primary impairment codes.

Table V.1. Suggested impairment code groupings and labels

Group Number*	Group Label	Impairment Codes
1	Major affective disorders	2960-2969 3110-3119
2	Schizophrenia and psychoses	2950-2959 2980-2989
3	Anxiety and neurotic disorders	3000-3019 3080-3099
4	Other mental disorders	2900-2949 2990-2999 3030-3079 3100-3109 3120-3129 3138-3169 3195 only
5	Intellectual disability	3170-3194 3196-3199
6	Back disorders	7221-7249
7	Musculoskeletal system	7100-7200 7250-7399
8	Infectious & parasitic diseases	0110-0119 0450-0459 0930-1359 1380-1389
9	HIV/AIDS	0070-0079 0201-0449 0540-0559 0780-0789 1360-1369
10	Neoplasms	1400-2399
11	Endocrine/ nutritional	2400-2479 2500-2559 2630-2799
12	Blood/ blood-forming diseases	2800-2899
13	Severe visual impairment	3610-3699 3780-3789
14	Severe hearing impairment	3890-3899
15	Severe speech impairment	7840-7849
16	Nervous system	3200-3419 3430-3599 3860-3889
17	Circulatory system	3420-3429 3750-3759 3900-4599
18	Respiratory system	4600-4869 4910-5199 7690-7699
19	Digestive system	5200-5799
20	Genitourinary system	5800-6299
21	Skin/ subcutaneous tissue	6900-7099
22	Congenital anomalies	7400-7599
23	Injuries	8000-9599

*The assignment of group number is arbitrary, and serves as a suggestion for coding primary diagnosis variables.

B. How to determine statutory blindness within the limitations of the administrative data

Beneficiaries who are statutorily blind are accorded some special provisions, including employment supports and a higher threshold for considering whether earnings are above the level of SGA. The best way to identify beneficiaries who are subject to the provisions of statutory blindness is to use the BLINDDT variable. A valid date in BLINDDT indicates that a determination of statutory blindness was made at one time. If BLINDDT is blank, no onset of blindness has been recorded for the beneficiary. If it contains a valid date, compare it to the reference date, if appropriate. For example, if you need to know if a beneficiary was blind when they entered the SSDI rolls, and the initial date of program participation was June 2005, determine if the date in BLINDDT is before June 2005. In addition to BLINDDT, for SSI beneficiaries, the presence of data in IEA_BLINDyymm or a variable value beginning with the letter "B" in COMP_STA or TOA (Type of Action) also means that the beneficiary was, at some point, determined to be statutorily blind.

It is important to note that BLINDDT does not identify all beneficiaries who have blindness or low vision. A value of 3694 in DXPRIBESTyymm or DXSECBESTyymm definitively indicates statutory blindness in a given month. This diagnosis code was only introduced in 2012, however, and is not yet in widespread use. An impairment code of "3690" indicates blindness or low vision, and any diagnosis code that begins with "36" indicates a visual impairment, although none of those codes necessarily indicate *statutory* blindness. An analysis conducted in 2006 showed that a larger than expected number of beneficiaries with a valid date in BLINDDT had no corresponding impairment code for blindness in the primary or secondary impairment fields. This situation was thought to reflect the fact that many beneficiaries who are blind have a number of other significant disabilities that are sufficient to qualify the beneficiary for benefits,

and thus these other disabilities are recorded in the impairment code fields while the blindness is indicated by entering a date in the BLINDDT field. Another factor is that a beneficiary may become blind after their initial qualification for disability benefits, for instance, from the progression of diabetes. This would likely be reflected in BLINDDT, but might not be recorded in the impairment codes if it is not the primary or secondary reason for benefit eligibility. In addition, it is possible for a finding of statutory blindness to be revised or revoked, even if a beneficiary continues to be disabled as a result of blindness or low vision.

This page has been left blank for double-sided copying.

VI. IDENTIFYING SSI COUPLES

When two individuals of the opposite sex who are married or holding themselves out to the community as married, or when the same-sex couples are married, and both receive SSI and live together, they are considered one unit for benefit purposes.[6] As a couple, the thresholds for financial eligibility will be higher, and the income and resources of both members of the couple affect the cash benefits to which they are both entitled. This can mean that an individual who is working at a level that would make him ineligible for SSI as an individual may be eligible to receive SSI as a couple.[7]

Identifying SSI couples in the DAF is challenging, and not always possible with certainty, in part because being a member of an SSI couple is based on residence and so a beneficiary may go back and forth between individual and couple status. The SSR variable CUR-COMP, known as CURCOMP on the DAF, can be used to identify couple status. However, CURCOMP only identifies SSI couples approximately 60 percent of the time. Specifically, CURCOMP only identifies cases where the SSI couple came into existence as a result of successful application for SSI filed by the ineligible spouse of a beneficiary who is already receiving SSI. Because there are two other ways of that an SSI couple might come into existence, CURCOMP does not

[6] The "holding out" concept is similar but not identical to common-law marriage. See POMS SI 00502.145 for a full discussion. As of 2015, "same-sex couples will be recognized as married for purposes of determining entitlement to Social Security benefits or eligibility for SSI payments" (https://www.ssa.gov/people/same-sexcouples/).

[7] For example, the SSI Federal Benefit Rate (FBR, or maximum monthly federal payment) in 2016 is $733 for individuals and $1,100 for couples. An individual with between $733 and $1,100 in countable (i.e., after all exclusions and income disregards) income, from working or from some other source, is ineligible for SSI so long as he or she is not a member of a couple. However, if that individual is married to and living with another SSI beneficiary, he or she becomes eligible in that month so long as his or her spouse's countable income is not sufficient to add up to more than $1,100.

successfully identify all SSI couples.[8] In a small number of cases, CURCOMP also erroneously identifies beneficiaries as members of SSI couples.[9]

[8] The other two ways of establishing an SSI couple are by the joining, either through marriage, residency, or both, of two beneficiaries already eligible for SSI, and by the simultaneous application of two individuals neither of whom were previously eligible for SSI. CURCOMP cannot be used to identify SSI couple status in these two cases.

[9] These erroneous identifications include the one or two months at the beginning or end of a period of SSI couple status when a new SSI record is being established and when one member of the couple is incarcerated. Because there is no valid external benchmark, it is impossible to be certain how frequent these erroneous identifications occur. Our investigations suggest that they are quite infrequent.

VII. IDENTIFYING THE EXPECTATION OF MEDICAL IMPROVEMENT

Beginning in DAF15, we improved on the monthly measures of the expectation of medical improvement among beneficiaries by incorporating data from the DCF.[10] While the MIE related variables built from the DCF improve upon those previously available on the DAF and sourced from the 831 & 832/833 files, the improvement is limited to data after April 2008.[11] The implication of this is that the monthly DAF MIE variables through April 2008 will contain data sourced from the 831 & 832/833; those from May 2008 forward will be sourced from the DCF. The algorithm for the variables sourced from the 831 & 832/833 will remain unchanged from earlier rounds of the DAF.

This affects two variables in the DAF: MEDEXyymm (previously named MIEXyymm), which indicates the medical CDR diary type in effect in a given month and year, and FIRSTMIE (previously named FRST_MIE), which indicates the date in which a beneficiary was first in the MIE category.

The MEDEXyymm variable, as detailed in Volume 5, contains the following categories:

- P – Medical improvement is possible (MIP)
- N – Medical improvement not expected (MINE)
- E – MIE
- 0 – No medical improvement information

[10] Prior to DAF15, the expectation of medical improvement was drawn from the 831 & 832/833 files, which have rates of missing information for medical CDR diaries of about 20%. In 2016, we learned that more complete medical CDR diary information is available from the Medical table in the DCF from 2008 forward, so we are combining data drawn from the 831 & 832/833 files and the DCF Medical tables to provide the most complete data available for medical CDR diaries.

[11] In April 2008, SSA converted the historical data from one system to another (from VSAM to DB2) to populate the DCF Medical table. During that conversion, some historical dates were overlain making it impossible to determine the timing of the determinations.

This page has been left blank for double-sided copying.

VIII. CONSTRUCTED MEASURES RELATED TO SUSPENSION OR TERMINATION FOR WORK AND DOLLAR VALUE OF CASH BENEFITS FORGONE DUE TO WORK

Because the relationship between work and benefit receipt is a subject of intense interest, we have constructed two families of analytic variables to aid researchers. These variables consist of monthly indicators for cash benefits in suspense or termination for work (STW) and a constructed monthly measure of the dollar value of cash benefits forgone for work (BFW). The core STW and BFW measures are included on the Annuals files, and described in detail in Sections VIII.A through VIII.C below and the relevant variable detail pages in Volume 5. The algorithm for developing these measures has been refined over time, so we recommend using the STW variables from the most recent DAF file available. After reading this section, users interested in identifying beneficiaries in STW during a certain period of time may want to consult the DAF code library, available at https://www.ssa.gov/disabilityresearch/daf.html#library, which provides sample code for that task.

In addition to the core measures, the DAF16 also includes a standalone companion work file (SCWF), which contains alternative versions of the STW and BFW measures. The SCWF, first developed as part of DAF14 and substantially revised in DAF15, is described in detail in Section VIII.D. It contains alternate versions of STWSSI, STWCM, and BFWSSI_DRAFT, and BFWCM_DRAFT that more fully account for deemed income and the budget month calculation for SSI beneficiaries. Users interested in STW and BFW may want to consider these measures in addition to, or instead of, the comparable ones contained in the Annual files. As described in Section VIII.D, this file is relatively new and to our knowledge, has seen limited use thus far; we look forward to user feedback to refine it in future versions of the DAF version.

A. Overview of STW and BFW measures

The DAF contains three variables that provide monthly information on the non-payment of cash benefits following suspense or termination for work, each with names that start with "STW." These measures are described in Section VIII.B. There is one variable for SSI beneficiaries, one for SSDI beneficiaries, and also a combined variable that takes concurrent beneficiaries into account.[12] In many instances, the data in a given month are not definitive about whether a beneficiary whose benefits were suspended or terminated for work are, in fact, continuing to work at a level that would make them ineligible for benefits in the current month. Instead, they only tell us that benefits were suspended or terminated in the past because of work and that benefits have not been reinstated as of the current month.[13]

The second set of variables indicates the dollar value of BFW. These are described in Section VIII.C. As with STW, there are three such variables, one each for SSI and SSDI and also a combined variable for concurrent beneficiaries. Conceptually, the value of a BFW variable for a month represents the amount of benefits that the individual has forgone in the month because he or she worked at a level that caused the suspension, reduction, or termination of benefits. For SSDI beneficiaries, the STWDI value must indicate suspense or termination for work in order for the BFWDI value to be populated, reflecting the SSDI program rules. SSI beneficiaries however, may forgo benefits due to work even when they are still receiving a cash benefit, as described below. Hence, BFWSSI may be populated even if the beneficiary is not suspended or terminated for work per STWSSI. In certain instances, BFW values for SSI beneficiaries are "special missing" characters, described below.

[12] In the TRF10 and earlier versions of the database, these began with "LDW" for "left due to work."

[13] These measures cease to be populated when a beneficiary reaches FRA or dies.

Determination of the value of BFW requires counterfactual information that can only be estimated: how much the individual's benefits would have been in the current month if benefits had not been suspended, reduced, or terminated for work. For BFW, estimates are based on program rules and information on benefit amounts prior to suspense or termination for work. For SSDI, these rules are such that the estimates are likely exact in a large majority of cases, and very close in the remainder. The SSI rules and data are more complex, making it more difficult to accurately estimate BFW in a large share of cases.

SSA and Mathematica researchers collaboratively developed the complex algorithms required to create the STW and BFW variables. The inputs to these algorithms are DAF variables that describe program status, benefit payment, income, and other indicators of work. These elements are not complete in all records. In cases where one or more of these variables is absent, the algorithms use second-best alternatives. As a result, these variables are sometimes inexact. STW accurately reflects whether or not the beneficiary is in current pay status (that is, due a cash benefit) according to the latest information in the administrative records, but determination of the exact reason for benefit suspense or termination requires examination of several administrative data elements and is not always definitive.

B. STW indicators

In both the SSDI and SSI programs, a suspension or termination status implies a $0 cash benefit for that month. In the SSDI program, a suspension or termination for work requires the beneficiary to have engaged in SGA after completing the 9-month TWP and three grace period months. For SSI, however, an indication of suspension or termination for work does not indicate any specific minimum level of work or earnings. SSI suspensions and terminations are based on all sources of beneficiary income, including earnings from work, SSDI benefits (for concurrent beneficiaries), deemed spousal income, income from any other source, and in-kind support and

maintenance as monetized in SSI rules. Minimal earnings from work may therefore be enough to push a beneficiary into suspense or terminated status. Conversely, an SSI beneficiary may be suspended for excess income as a result of other unearned income alone, so the beneficiary is not considered suspended for work, regardless of the size of beneficiary earnings.

We developed three longitudinal indicator variables of non-payment status due to suspense or termination for work, included in each Annual file and populated on a monthly basis. The STWDI variables indicates how work has affected the status of SSDI benefits, the STWSSI variable does the same for SSI benefits, and the STWCM variable indicates how work has affected the combination of SSDI and SSI benefits. Each variable is named according to the year and month using the same *yymm* convention as elsewhere in the DAF (for example, STWDI1601, STWDI1602, and STWDI1603 are the STWDI values for the first three months of 2016).

1. The values of the STWDI and STWSSI indicators

The values for the STWDI and STWSSI variables show the beneficiary's work status in the specified program in any given month, but have somewhat different sets of values. The values for STWDI are as follows, all referring to SSDI status:

0 = in current pay status in this month

1 = suspended due to work in this month

2 = terminated due to work in this month

3 = presumed terminated due to work in this month: terminated due to work in an earlier month and never reinstated, and beneficiary is alive and has not attained FRA

8 = in suspense status in this month for a reason not determined to be work[14]

9 = in terminated status in this month for a reason not determined to be work

. (missing value) = beneficiary is not yet entitled to benefits, has died, or has attained FRA

[14] For previous users of STWDI, note that the value of 8 was first introduced in the DAF12.

The values for STWSSI are as follows:

0 = in current pay status in this month

1 = suspended due to work in this month

2 = terminated due to work in this month

3 = presumed terminated due to work in this month: terminated due to work in an earlier month and never reinstated, and beneficiary is alive and has not attained FRA

4 = suspended due to excess income in this month, part of that income is earnings from work, but countable unearned income is sufficient to cause suspension without consideration of earned income

8 = in suspense status in this month for a reason not determined to be work

9 = in terminated status in this month for a reason not determined to be work

. (missing value) = beneficiary is not yet entitled to benefits, has died, or has attained FRA

There is one more possible value for the STWSSI variable than STWDI due to the greater complexity in determining suspension or termination for work in the SSI program. Because SSA administrative data do not distinguish between suspensions and terminations as a result of excess income from work versus excess income from other sources, STWSSI can take on a value of 4. For SSI beneficiaries with a suspension due to excess income, an STWSSI value of 1 indicates that the beneficiary was working during the months of suspension and that the earnings from work were necessary for suspension (even if they were not sufficient to induce suspension in the absence of unearned income). In contrast, an STWSSI value of 4 indicates that the beneficiary was working, but that the beneficiary would have been suspended for excess income in that month even without that earned income (that is, as a result of the unearned income alone).

When developing STWSSI, we first identify beneficiaries in suspense due to excess income (PSTA=N01). We then compare unearned income to the federal benefit rate (FBR); if unearned income is above the FBR, then we determine an SSI beneficiary is STWSSI=4 because the unearned income was sufficient to preclude a benefit payment. If unearned income is less than

the FBR, then the earned income was necessary to result in benefits being suspended, in which case STWSSI=1.

2. Combined STW indicator

The combined STW indicator was developed by taking into account the STW status for both the SSI and SSDI programs. As either the STWDI or STWSSI variables of themselves do not provide complete information for concurrent beneficiaries, the indicator is useful in any context where a DAF user wants to know the combined benefit status. If a beneficiary only receives benefits from SSDI or SSI, the STWCM indicator may also be used, reflecting their STW status in the relevant program. This variable was designed to support the evaluation of TTW, but it is useful in most other examinations of work and benefits as well. Under TTW, if a qualified service provider, called an EN, has accepted assignment of a Ticket from the beneficiary, the EN is entitled to an outcome payment if and only if the beneficiary is not receiving benefits from either SSDI or SSI; STWCM was designed to reflect EN eligibility for an outcome payment.

Construction of STWCM is based on the separate SSI and SSDI STW indicators and errs toward current pay status in cases where the two indicators are not the same. For example, if the separate indicators show suspense or termination for work in one program (1, 2 or 3), but current pay status (0) in the other program, STWCM indicates the beneficiary is "in current pay status." More generally, the STWCM indicator takes on the value of the lower of the two values for STWDI and STWSSI; for example, if STWDI=0 and STWSSI=1, then STWCM=0, and if STWDI=3 and STWSSI=2, then STWCM=2. The only exception to this rule is when STWSSI=4, which is treated as STWSSI=8 in setting the value of the corresponding STWCM variable. For example, if STWSSI=4 and STWDI=. (missing), 8, or 9, then STWCM=8. This exception was instituted because an STWSSI value of 4 does not indicate a loss of benefits

specifically due to income from work. Rather, the value of 4 indicates that the beneficiary is working, but would be in suspense for income status even if not working.

The STWCM values are as follows:

0 = in current pay status in this month for either SSDI or SSI.

1 = suspended due to work in this month for either SSDI or SSI and not in current pay in this month for the other program.

2 = terminated due to work in this month for either SSDI or SSI and not in current pay or suspended for work in this month for the other program.

3 = presumed terminated due to work in this month for either SSDI or SSI and not in current pay, suspended, or terminated for work in this month for the other program; but benefits terminated due to work in an earlier month in at least one program and never reinstated, and beneficiary is alive and has not attained FRA.

8 = in suspense status for reasons not determined to be work for either SSDI or SSI and not in current pay, suspended, or terminated for work in this month for the other program.

9 = in terminated status for reasons not determined to be work in either SSDI or SSI and not in current pay, suspended, or terminated for work in this month for the other program.

. (missing value) = beneficiary is not yet entitled to benefits, has died, or has reached FRA

Researchers should choose the STW indicator that is appropriate for their purposes taking into account how all three STW indicators are constructed and work together, and how SSDI and SSI interact.

3. Presumption of work after benefit termination

The STW measures take on a value of 3 in the months after SSA terminates benefits due to work, provided the beneficiary has not regained eligibility, reached FRA, or is known by SSA to be deceased. In these months, it is not possible to determine from the administrative data whether the former beneficiary is actually earning at a level that would preclude eligibility for benefits. Thus, an STW value of 3 indicates that work is the reason that benefits were terminated prior to that month, but does not imply that earnings during that month make the beneficiary ineligible for benefits during that month. For example, STWDI0606=3 means that benefits were terminated

for work prior to June 2006 and the beneficiary had not died or reached FRA, and had not returned to receiving cash benefits as of June 2006. But, it does not mean that the beneficiary is necessarily engaged in SGA or worked at any level in June 2006—he or she may be, but it is not possible to determine that from the administrative data. SSA has no reason to collect beneficiary work information after benefits have been terminated unless the beneficiary applies for reinstatement.

Researchers should not misinterpret a value of 3 as indicating that the former beneficiary is necessarily earning enough in that month to be ineligible for benefits. In many cases, we believe it is reasonable to presume continued work is the reason for not returning to benefits, though we have not verified this to be true in all cases. Even if the beneficiary is not engaged in SGA, a "3" STW value indicates that past work is the reason that SSA is not paying benefits to the individual in the current month.

4. Auxiliary SSDI beneficiaries

Special rules are applied to the DAF records of auxiliary Social Security beneficiaries with disabilities whose entitlement is based on their own medical eligibility and the earnings record of another person who is a primary Social Security beneficiary: DACs and DWBs.[15] For DWB, STW values that occur after the primary beneficiary is deceased apply to work of the DWB, not the primary beneficiary. For DAC, in many cases the primary beneficiary is retired or deceased, but in some cases the primary beneficiary is an SSDI worker beneficiary. In such cases, any evidence of suspension or termination on the record of a DAC might actually belong to the primary beneficiary, because a suspension or termination on the record of the primary beneficiary applies equally to auxiliary beneficiaries *even if* the auxiliary beneficiaries are not

[15] Approximately 10.7% and 2.6% of SSDI beneficiaries were entitled under DAC and DWB provisions, respectively, in December 2016 according to Chart 1 in the 2016 Annual Statistical Report on the SSDI Program.

working. Hence, if a DAC record is flagged as in suspense or termination following the above methods, further checks are done to ascertain whether the suspension or termination is due to the earnings of the auxiliary beneficiary.

SSDI records contain two identifying variables: 1) the CAN (the SSN of the primary beneficiary) and 2) the BOAN (the beneficiary's own SSN). For primary beneficiaries, the CAN is the same as the BOAN. For DAC beneficiaries, the CAN is the SSN of the primary beneficiary (i.e., the beneficiary's parent), and the BOAN is the DAC's own SSN. When constructing STW indicators, we therefore take note of the CAN when a primary beneficiary is flagged as suspended or terminated due to work. Later, if a DAC record is flagged as STW, we check to see if its CAN variable is the same as the CAN of a flagged primary beneficiary. If so, we assume that the suspension/termination information belongs to the primary beneficiary and not to the DAC, and the STWDI value for the DAC is set to 8 or 9 for suspension or termination, respectively.

5. SSI couples

STWSSI, and thus also STWCM, algorithms attempt to account for the difference in program rules associated with members of SSI couples using the CURCOMP variable described in Chapter VI above. An SSI beneficiary is a member of an SSI couple if he or she is married to and living with another SSI beneficiary. Under SSI program rules, SSI couples are treated as a single unit for the purposes of determining eligibility for an SSI payment and payment amount. In 2016, the SSI benefit amount for an individual was $733, for a couple it was $1,100.

Although CURCOMP is not a perfect indicator of SSI couple status, we determined in consultation with SSA that incorporation of monthly CURCOMP values into STWSSI construction was worthwhile. As a result, STWSSI values starting in DAF13 are derived differently from the values in prior versions of the database for beneficiaries now identified as

members of SSI couples. The effect of incorporating couples' status through CURCOMP into this calculation was in changing the relevant FBR; instead of comparing to the individual FBR ($733 in 2016), members of an SSI couple are compared to one-half of the couples FBR (or $550 each based on a couple FBR of $1,100 in 2016). This effectively reduced the FBR by 25 percent for beneficiaries identified to be part of a couple. As a result, incorporating CURCOMP made it more likely for members of an SSI couple to have unearned income that is sufficient for suspension of benefits due to excess income, resulting in a higher number of STWSSI=4 months, and fewer STWSSI=1 months.

SSA's rules for allocating non-SSI income of any sort to the two members of an SSI couple implies that income received by either member of the couple, alone, can affect both members' values for STW and BFW. That is, all earned and unearned income is attributed equally to both members of the couple. In fact, each member of a couple will generally have the same STWSSI status and it is not possible to determine which member of the couple did the work that resulted in any suspension or termination.[16]

We include CURCOMPyymm on the DAF so that users can determine for themselves which beneficiaries are potentially affected by this issue if they are interested in doing so. For more information about how CURCOMP was incorporated into STWSSI construction, please consult the SAS program STWSSI in Volume 11.

C. BFW indicators

We developed three sets of monthly variables capturing the dollar value of benefits forgone due to work, included in each Annual file and populated on a monthly basis. As with the STW

[16] There are exceptions to SSI couples having identical statuses. For example, one member of the couple may be incarcerated. In this case, CURCOMP continues to identify each beneficiary as a member of a couple during the months of incarceration, but the STWSSI values in those months are likely to be different.

indicators, there is one BFW indicator for SSDI, one for SSI, and a combined indicator. Each is named according to the year and month through the last month covered by the DAF as described above. Although BFWSSI incorporates CURCOMP for SSI couple status, it is subject to limitations that are described below, warranting the variable name BFWSSI_DRAFT. Because it relies on BFWSSI_DRAFT, the combined BFW indicator is called BFWCM_DRAFT.

1. BFWDI

BFWDI is a monthly estimate of the SSDI benefit amount forgone by an SSDI beneficiary and any associated auxiliaries because of work. After an SSDI beneficiary completes the TWP, the beneficiary and any auxiliaries are not entitled to a benefit payment in any month in which he or she engages in SGA after three grace period months. Therefore, for months in which the STWDI indicator is 1, 2, or 3, BFWDI is equal to the SSDI benefit and auxiliary amount for the most recent month in which the STWDI value was 0, i.e., when the beneficiary was last in current pay.

The benefit amounts used in BFWDI are contained in the DUEDyymm and DUEOyymm variables for the last month in current pay, increased for the annual COLAs that take place each year in January. To illustrate, these increases are in DAF-year dollars such that a beneficiary whose most recent STWDI value of 0 was in 1998 would have a 2001 BFWDI value that reflects COLAs between 1998 and 2001 applied to DUED and DUEO. The values in DUED and DUEO are used in the BFWDI calculation for as long as the STWDI value remains 1, 2, or 3, subject to the imputed COLA increase, until the beneficiary reaches FRA, dies, or becomes re-entitled to cash benefits (comes back into current pay status).[17]

[17] SSDI beneficiaries participating in the BOND are eligible to receive reduced SSDI benefits when they are working above SGA. In other words, for SSDI beneficiaries participating in BOND, benefit receipt is not binary. Because this is outside normal program rules and because BOND participants represent a small share of total beneficiaries, we did not separately change the BFWDI calculation for these participants.

2. BFWSSI_DRAFT

Unlike SSDI, SSI benefit receipt is not binary. Rather, an SSI beneficiary may receive any amount between $0 and the federal benefit rate (FBR), which is the maximum monthly federal benefit. Many types of income in addition to that received from work can affect SSI eligibility and reduce the amount due to a beneficiary. In addition to calculating this benefit reduction, there are several programmatic factors that complicate the calculation of BFWSSI, the most important of which are described below. Because of the limitations described below, DRAFT has been appended to the variable name (BFWSSI_DRAFT).

a. A refresher on the computation of monthly SSI benefits

In each month, SSI benefits are determined using a two-part assessment. The first step determines whether a benefit is due at all, while the second part determines the amount of the benefit. An SSI recipient is due benefits in a given month provided that the current month's income does not exceed the SSI federal benefit rate (FBR). In other words, current month income determines whether a benefit is paid. If income (either earned or unearned) exceeds FBR, then the SSI recipient does not receive a benefit. If income is under FBR, the benefit may be reduced, but the beneficiary is still in current pay status.

If a benefit is due, the next step is calculating the amount. As explained in POMS SI 02005, SSI payments are computed using a method known as RMA. Under RMA, income received in a given month generally affects the payment computation for two months later, with the following primary exceptions:

- When the income in a given month is sufficient to make a beneficiary ineligible due to excess income, the two-month look-back does not apply and the beneficiary is ineligible in that month.
- In the first month of overall eligibility or the first month of eligibility after a period of ineligibility for any reason, the two-month look-back does not apply. In that first month, the SSI payment is based on the income in that month; in the following month, assuming that eligibility continues, the payment is based on the same income as for the first month (one

month look-back); in the third month, assuming that eligibility continues, the normal two-month RMA look-back resumes and continues until another exception occurs. Thus, as long as eligibility continues, the income amounts used in the computation of benefits in each of the first three months are the same and the income in that first month is triple counted; that is, the income amounts used in the computation of benefits in each of the first three months after a period of ineligibility for any reason are identical with each other and with the income amounts in the first month in that string.

For any given month, the term budget month (BM) is used to refer to the month the income of which is used to calculate the given month's benefit amount. Depending on the circumstance, the budget month is either 0 (the current month), 1 (the month prior), or 2 (two months prior). In other words, the benefit computation for any given month's SSI benefit is based on the income in either the current month, the month before, or two months prior. While countable income below the FBR in the current month places the SSI recipient in current pay (PSTA=C01) status, the budget month, which determines the payment amount, is usually two months prior.[18] So, for example, an SSI recipient in current pay status (PSTA=C01) in February 2002 has countable income less than FBR will have a computed benefit based on income in December 2001, the relevant budget month in that case. When an SSI recipient is in nonpayment status due to excess income (PSTA=N01), however, the budget month is the current month. In other words, when countable income is greater than FBR, that month's income factors into the determination of both a benefit being paid and the benefit amount (which in N01 months is always $0).

b. Calculating BFWSSI

An issue in the construction of the BFW variable for SSI beneficiaries is how to interpret the conceptual intent of the BFW variable—to calculate benefits forgone due to work—when the

[18] Exceptions to this rule are noted above. This list of exceptions to RMA is not exhaustive, as explained in the previous section on POMS. However, these are the most significant exceptions and the only ones that are addressed in the construction of BFWSSI_DRAFT.

BM for the calculation of BFW is affected by earnings.[19] Specifically, the budget month in reality changes depends on whether or not a beneficiary had earnings, meaning that under the scenario of no earnings, the budget month would be different. In other words, the BFW calculation is based on a counterfactual: it is the difference between what a month's benefit *would* have been in the absence of earned income and the benefit actually paid in that month (if any). Under RMA rules, the BM under the BFW counterfactual might differ from the BM that was actually used for the same month, and the beneficiary might have had earned income in the counterfactual BM.

For example, a beneficiary with $1,000 of countable earned income in August 2013 will have an actual BM of August 2013 in that month, because $1,000 is high enough to preclude benefit receipt in that month. Yet, the counterfactual BM (the budget month that would have been used in the absence of earnings) could have been either June or July 2013 (depending on the pattern of earnings in prior months). In those months, the beneficiary might have still had countable earned income. The intent of the BFW calculation, however, is to determine the difference between benefits received and benefits that *would* have been received in the absence of any earned income, regardless of *when* that income was earned. Hence, the BFWSSI_DRAFT variable ignores earned income in the counterfactual BM, regardless of when that month occurs. This means that even if the beneficiary had earnings in June or July 2013, the calculation of BFWSSI_DRAFT would ignore them.

[19] The construction of BFWSSI_DRAFT also accounts for SSA policy on nonrecurring income, described in POMS SI 02005.005.A.4. Prior to April 2005, it was possible for an SSI beneficiary to receive one-time income in the first month of eligibility after a period of ineligibility that, because of RMA, would reduce the SSI benefit in that first month, the second month, and the third month, assuming that eligibility continued. To address this issue, beginning in April 2005, SSA revised its policy on income received in the first month after a period of ineligibility, so that the non-recurring income would affect only the first month and not the two following months. This avoids triple-counting of non-recurring income.

c. Accounting for SSI couples

BFWSSI construction accounts for SSI couple status using CURCOMP. The same counterfactual principle described above applies to members of SSI couples, but as described above, suspense or termination for work is less frequent for couples after incorporating the couples indicator. In addition, the same caveat concerning the impossibility of determining which member of the couple earned the income that is equally attributed to each member of the couple in the administrative data applies to BFWSSI, i.e., a positive BFWSSI value for a member of an SSI couple does not necessarily mean that it was the individual's own (rather than his or her spouse's) work that resulted in forgone benefits.

d. Limitations of BFWSSI warranting the _DRAFT suffix

There are several known limitations of the BFWSSI_DRAFT variable as it is currently constructed. It is not possible to construct an entirely accurate BFW variable for SSI beneficiaries given the data currently in the DAF and the complexity of the rules for determining SSI benefit amounts. There are three program rules that we cannot address in the construction of BFW because of limitations of the data in the current DAF.

- **SSI couples.** Although BFWSSI_DRAFT accounts for SSI couple status using the CURCOMP variable described above, CURCOMP does not identify all beneficiaries who are members of SSI couples. Therefore, BFWSSI_DRAFT values of those beneficiaries who are members of SSI couples but not identified as such are systematically too high.
- **Deemed income.** The BFWSSI_DRAFT calculation does not incorporate information on deemed income. For BFWSSI_DRAFT purposes, in most cases this is income received by the spouse of an SSI beneficiary who lives in the same household and does not receive SSI. Because the counterfactual BFW calculation involves a reconstruction of what the benefit amount would have been in the absence of earnings from work, if there is deemed income, it should be deducted from the counterfactual benefit. The absence of deemed income in the calculation of BFWSSI_DRAFT means that BFWSSI_DRAFT for beneficiaries who have deemed income is systematically too high.
- **Proration.** As described in POMS SI 02005.008, benefits are prorated for the first month in which eligibility is obtained after a period of ineligibility when eligibility is reinstated on a date other than the first day of the month, such as when an individual returns to the U.S. after a period of absence of more than 30 days on a day other than the first of the month. It is not possible to calculate an accurate BFWSSI_DRAFT value in these cases because the

DAF does not contain the specific day on which eligibility is re-attained. As a result, the BFWSSI_DRAFT amount is too high if proration applies in the month for which the value is being calculated.

Of these three known limitations, proration is the least frequent and least significant, because it only affects one month in each instance. The limitations surrounding SSI couples and deemed income are more significant and it is important for researchers to be aware of these limitations in their research design and conclusions. In the next section, we describe a new file that contains new STW and BFW measures for SSI beneficiaries that attempts to better incorporate deemed income. We continue to look for ways to improve the measurement of SSI couple status.

In addition to these known limitations associated with specific SSI policies, certain anomalies in the DAF data also affect the BFWSSI_DRAFT computation. In rare instances, the regular formula for BFWSSI_DRAFT produces values that are incompatible with SSI policy, such as values less than zero or greater than the FBR. In these cases, we have assigned a "special missing" value of ".m" so that researchers are aware that no BFW value can be calculated for these SSI beneficiaries in those months given available data.

Researchers should also be aware that BFWSSI_DRAFT applies only to *federal* SSI benefits forgone due to work. The policies on state supplementation vary so much on a state-by-state basis that it is not feasible to incorporate anything other than federal SSI benefits in the calculation of the BFW counterfactual.

3. Values of the BFW indicators

The values of all three BFW indicators are expressed in dollars, with a minimum value of zero. BFWSSI_DRAFT variable values are constrained at the upper end by that year's individual FBR, or 50% of the couples FBR if the beneficiary is identified as a member of an SSI couple by the CURCOMP variable value in that month. The combined BFW variable, BFWCM_DRAFT,

is calculated simply by adding the value of BFWDI and BFWSSI_DRAFT together for each month. Note that this means that a beneficiary may have an STWCM value of 0 but a positive BFWCM value in any given month when benefits are suspended for work in one program but not the other. The combined BFW variable also contains the DRAFT suffix because it is derived in part from the BFW variable value for SSI beneficiaries and so is subject to the same limitations.

Because the STW variables presume a termination due to work (value of 3) in months after termination when SSA is no longer collecting data on the beneficiary, the BFW variable continues to be positive in such months as well. Beneficiaries whose benefits terminate due to work activity are eligible for a period of 60 months after benefit termination for an immediate return to cash benefits once they cease engaging in SGA under SSA's rules for expedited reinstatement (see POMS DI 13050). Therefore, it is reasonable to assume that the lack of reinstatement implies that these terminated beneficiaries continue to work, at least for the five years following the termination of benefits due to work, but researchers should be aware that this is an assumption and possible to be violated in some months.

None of the BFW values are adjusted for inflation. BFWDI values include imputed COLA increases from the last monthly benefit on record, but these are not adjustments for inflation. Rather, the COLA increases are imputed in order to approximate what benefits *would* have been had they not been suspended or terminated for work.

D. SCWF

The SCWF contains versions of BFWSSI and STWSSI that use modified algorithms that include information contained in an expanded set of administrative variables.[20] These variables and the resulting effects on STWSSI and BFWSSI include:

- **BMF (BMFyymm).** This variable allows us to more accurately specify the budget month for calculating the dollar value of benefits forgone due to work *in months when the beneficiary is in current pay status*. The change involving BMF is straightforward and applies to BFWSSI only.
- **FCI (FCIyymm).** This variable more completely accounts for all non-earned sources of income, including deemed income and in-kind support and maintenance (ISM). Modifying the BFWSSI computation algorithm to use FCI is complicated, but overcomes a known limitation in the current variable in accounting for deemed income. We use FCI to determine whether a beneficiary is suspended or terminated for work or for some other reason and, if for work, the dollar value of the benefits forgone due to work. As a result, incorporating FCI will imply changes to both STWSSI and BFWSSI.

These variables, along with updated STW and BFW variables (including those for SSI and the analogous combined indicators), described below, are available in the SCWF. This file is fully linkable to all other DAF files using SSN. It contains one record per beneficiary, with monthly indicators, for all years included in the DAF for the series of new variables; these cover the period from January 1994 to December 2016. More detail about the variables contained in the SCWF is provided in Table VIII.1.

We opted to offer these revised STW and BFW measures to users as part of the SCWF for two reasons. The first is that users familiar with the existing versions of STW and BFW may be interested in cross-version comparability in updating their existing analyses. Had we overwritten the algorithm with the version on the SCWF, this would not have been possible. The second is that these measures are still in the testing phase. While we conducted preliminary checks to

[20] In what follows, we omit the _DRAFT suffix off of the BFWSSI measure for clarity of exposition. The measure is identified with that suffix for the reasons documented in the previous section and continues to be so for the variables described here.

verify that the updated algorithms worked as expected, we recognize that oftentimes, those conducting research are more able to notice inconsistencies in new measures than is possible using these types of checks. As such, we encourage users to consider both measures in their analyses and welcome feedback on any discrepancies, unexpected findings, or general comments about these improvements. Please direct comments to ORDES.DAF@ssa.gov.

Table VIII.1. Variables contained in the SCWF

SCWF Variable	SCWF Variable Name	Notes
BMFyymm	Budget Month Factor	Source variable necessary to construct new BFWSSI variable.
FCIyymm	Federal Countable Income	Source variable necessary to construct new STWSSI and BFWSSI variables.
STWSSI_FCIyymm	SSI Suspense or Termination Due to Work, FCI Considered	STWSSI variable that incorporates FCI in its construction, as described above.
STWCM_FCIyymm	Combined Suspense or Termination Due to Work, FCI Considered	STWCM variable that uses STWSSI_FCIyymm instead of STWSSIyymm.
BFWSSI_DRAFT_BOTHyymm	SSI Estimated Benefits Forgone for Work, BMF and FCI Considered	BFWSSI variable that incorporates both BMF and FCI in its construction, building off of BFWSSI_DRAFT_BMFyymm. Because this variable incorporates FCI data, it relies on the SCWF variable STWSSI_FCIyymm.
BFWCM_DRAFT_BOTHyymm	Combined Estimated Benefits Forgone for Work, BMF and FCI Considered	BFWCM variable that incorporates BFWSSI_DRAFT_BOTH in its construction.

1. Incorporating BMF into the BFWSSI calculation

As described above, income in the budget month determines the amount of SSI benefit payment in a given month. In previous versions of the DAF, we estimated the budget month based on data about the beneficiary's status otherwise available in the DAF, following the rules outlined above. Because our method for determining budget month was only an estimation of the true value, it may have differed from the budget month used by SSA. With the BMF variable now available, we are able to substitute an administrative data source to determine the accurate budget month.

2. Incorporating FCI into STWSSI

Abstracting from the budget month for a moment, the amount due to an SSI beneficiary is determined by comparing the FBR to a beneficiary's income minus any exclusions. In other words, the federal amount due to an SSI recipient (FAMT) is the FBR, minus countable income. In building prior versions of the DAF, we have had access to EICM (countable earned income) and UINC (countable unearned income). In most cases, EICM+UINC is the countable income amount used for computation of the monthly benefit. A known limitation of that sum, however, has been that EICM+UINC does not include deemed income. The new variable available, FCI, incorporates EICM+UINC, but also includes deemed income in cases where it applies. Using FCI, then, should yield a better estimate the payment amount, that is, FAMT=FBR-FCI.

Replacing FCI to calculate income affects the STWSSI calculation. Up until this point, when PSTA=N01, we determined whether STWSSI is 1, 4, or 8 by first identifying whether there was countable earned income (EICM) in the current month. We use the current month because in N01 status, the budget month is the current month. If there was not (i.e., EICM=0), then we set STWSSI=8, meaning in suspense for reasons not determined to be work. If there was countable earned income in that month (EICM>0), we next compared the amount of the countable unearned income (UINC) in the month to the FBR. If the UINC was greater than the FBR, we set STWSSI=4 because the unearned income alone was sufficient to produce a suspension for excess income. If the UINC<=FBR, STWSSI=1 because EICM>0 was necessary to put the total countable income for that month over the FBR.

The addition of FCI allows us to overcome the exclusion of deemed income that was a result of using EICM and UINC alone. Using FCI allows us to generate TCNEI, including deemed spousal income, using the formula FCI – EICM = TCNEI. Put another way, TCNEI is the sum of UINC, deemed income, and income support and maintenance (ISM). In the decision logic for

STWSSI=1 and STWSSI=4 cases described above, we therefore now compare TCNEI (rather than UINC) to the FBR. When we first developed the SCWF in DAF14, we considered the share of beneficiaries in PSTA=N01 status in each month of 2012 (preceding the end of DAF14 enough to minimize the effects of data processing lags). Among beneficiaries in PSTA=N01, we found that approximately two out of three have some amount of deemed income (for example, 62 percent of those in N01 status in January 2012 had deemed income based on TCNEI> UINC). Among beneficiaries who were originally STWSSI=1 in months during 2012, more than nine in ten had deemed income, meaning that they were potential candidates for moving from STWSSI=1 to some other status (STWSSI_FCI=4 or 8) when we incorporated FCI. For example, in January 2012, 94 percent of those in STWSSI=1 had TCNEI>UINC.

In Table VIII.2, we compare the STWSSI measure contained in the DAF15 Annual files with STWSSI_FCI using data from January 2013; we did not do a similar comparison in DAF16, but do not have reason to believe the results would differ. We selected 2013 because it was sufficiently before the end of the DAF15 observation period to be confident that data reporting lags should contribute minimally to the findings. There are several things to note when comparing the two variables:

- In the majority of cases, STWSSI_FCI is identical to STWSSI.
- Approximately 2 percent of STWSSI=1 cases move to STWSSI_FCI=4 as a result of deemed income, and 1 percent of STWSSI=2 cases move to STWSSI_FCI=4.
- The biggest effect of the change in calculation is in STWSSI cases that were in 2 or 3 status; STWSSI=2 cases move to STWSSI_FCI=8 (6.1 percent) or STWSSI_FCI=9 (14.6 percent) and among STWSSI=3 cases, and 18.8 percent move to STWSSI_FCI=9. The reason this occurs is similar to the cases that move from 1 to 4; incorporating FCI means that earned income is no longer a factor in suspense or termination.21 Thus, the effect of incorporating FCI is in reducing the share of beneficiaries suspended or terminated for work, meaning that not incorporating this information had led to STWSSI being upward biased.

[21] The STWSSI algorithm is contained in Volume 11 and is quite complex; users seeking more information about the reasons for these changes can contact ORDES.DAF@ssa.gov.

Table VIII.2. A comparison of STWSSI and STWSSI_FCI in January 2013

	Total	STWSSI_FCI1301 (row percent)						
STWSS1301		0	1	2	3	4	8	9
0	5,279,383	100.0	0.0	0.0	0.0	0.0	0.0	0.0
1	72,941	0.0	97.6	0.0	0.1	1.8	0.5	0.0
2	495	0.0	0.2	79.0	1.6	0.2	5.5	13.7
3	50,786	0.0	0.0	0.0	81.4	0.0	0.0	18.6
4	32,119	0.0	0.3	0.0	0.01	99.5	0.05	0.0
8	1,186,352	0.0	0.0	0.0	0.15	0.0	99.85	0.0
9	3,606,593	0.0	0.0	0.0	0.15	0.0	0.0	99.85

Note: The STWSSI_FCI1201 values are the percentage of the total shown in the first column. These statistics were developed using the DAF16. We selected January 2013 to allow any lags in recording of variables used to generate STWSSI to be minimal.

3. Incorporating FCI into the BFWSSI calculation

For the reasons described above, the BFWSSI estimate on the Annuals files is too high in months where there is deemed income, because EICM and UINC—used to measure total countable income—do not account for all income that the beneficiary might have received. To update BFWSSI, we initially thought we could perform an identical substitution based on the TCNEI formula above. This is not possible, however, because the FCI variable is based on countable income in *the budget month*, while EICM and UINC variables indicate the countable income received *in the current month*, regardless of budget month, payment status, or benefit amount. As described above, the budget month when PSTA=N01 (non-payment due to excess income) is the current month, but in C01 (current pay) months, the budget month is usually (though not always) two months prior.

The difficulty incorporating FCI into the BFWSSI calculation arises because BFWSSI is based on computing a counterfactual situation—how much *would* benefits have been without earnings? This counterfactual relies on the assumption that an SSI recipient had no earnings in any month that could possibly affect the benefit amount. Because the counterfactual relies on a situation that did not actually occur, the budget month may be different under the counterfactual

relative to the budget month used in the actual computation. The change in budget month under the counterfactual when FCI is only recorded for the actual budget month becomes problematic in cases where STW=1, 2, or 3. An example below illustrates this issue more clearly.

Incorporating the deemed income information contained in FCI is valuable enough to warrant a revised approach to calculating BFWSSI. In particular, when using FCI to estimate deemed income, we rely on actual budget month information (contained in BMF) rather than the estimated budget month that we constructed to determine benefits in the absence of any earnings. When we did not account for the sum of deemed income and ISM, we could say with certainty that BFWSSI is overestimated. Changing the budget month under the counterfactual would result in some error, though we do not have a strong indication that the error is systematically upward or downward.

Beginning in the SCWF in DAF15, we also impute a value for deemed income and ISM in the last two current pay months before an SSI beneficiary moves from current pay status to suspense.[22] By accounting for all months, this imputed value further reduces the likelihood that our estimate of BFWSSI is biased upward. This estimate is necessary because FCI while in suspense is based on the actual budget month (the suspense month) so for the two months prior to the suspense month, no information on deemed income is available. In DAF14, we did not attempt to populate values in the two months preceding suspense and instead relied on using the actual budget month.

We estimate deemed income in the months prior to suspense by using the average of deemed income in the two most recent current pay months where FCI is available (provided they are

[22] We refer to this as deemed income in the remainder of what follows for simplicity's sake, but our measure cannot distinguish between the two forms of income.

within the 12 months immediately preceding the suspense month) and the two earliest current pay months immediately succeeding the suspense month (provided they are within 12 months of the first suspense month).[23] It is important to note that this estimate of deemed income is based on the assumption that the deemed income in the months surrounding the suspense month are a reasonable expectation of what to expect in the months immediately preceding the suspense month. In the future, we may revisit the sensitivity of BFWSSI to this approximation by considering other alternatives, including deemed income only in the period prior to the suspension months or deemed income only during the suspension months.

4. An example illustrating the BFWSSI calculation using FCI

In this section, we show an example of the BFWSSI calculation under two schemes: the one we use in the DAF16 Annuals files, and the version used in the DAF16 SCWF. We present the same data for this sample beneficiary once for each of the schemes, with the final two columns of each table reflecting the relevant BFWSSI calculation. Though the data is the same in each case, not all of the columns are used in each calculation, and as such, we gray out columns that were not necessary for the relevant BFWSSI calculation. We highlight a common situation in which the switch from EICM+UINC to FCI and knowing the budget month becomes an important consideration.

In the example, the individual becomes an SSI beneficiary in December of 2006, entering in current pay status. The beneficiary moves from C01 to N01 in April 2007, and remains in N01 until June 2007. In March 2007, the actual budget month is January 2007 (two months before), but in April 2007, the budget month becomes April 2007. We show the beneficiary's monthly record through the end of 2007 for illustration purposes.

[23] We use data from the maximum number of months that has deemed income information available during this period, ranging from 1 to 4.

Table VIII.3 highlights the BFWSSI calculation in the core DAF16 Annuals files (BFWSSI_DRAFTyymm). In April 2007, the recipient enters N01 status. In this month, EICM=$625, higher than the FBR of $623 that year, meaning that the recipient was STWSSI=1. In N01 status, the budget month is the current month. This means that the FCI, EICM, and UINC values recorded would be from April 2007. Yet, under the counterfactual used for determining BFWSSI, we would be interested in what would have happened if not N01 (or more accurately, if the recipient had no earnings in any month that could have affected benefits). In that case, the budget month would have been February 2007. On the DAF15 Annuals files, BFWSSI would be based on UINC from February 2007 alone, which would have been pulled from the UINC amount in that month. In other words, the EICM of $625 in April 2007 made the recipient STWSSI=1, but the amount forgone is the amount that would have been paid in the absence of work, BFWSSI = FBR – (UINC in February 2007) = $623 – 50, which is $573.

Table VIII.4 highlights the addition of deemed income and the resulting calculation for BFWSSI used in the DAF16 SCWF (BFWSSI_DRAFT_BOTHyymm in DAF16). Because of the change in budget month when moving from C01 to N01 status in April 2007, it is impossible to calculate deemed income using FCI-(EICM+UINC) in February and March 2007 because FCI for those months was not recorded. As such, in the SCWF, we estimate deemed income in those months as described previously. In particular, we considered the deemed income amounts that were available in December 2006 and January 2007 (the two months of populated data in current pay months immediately prior to the suspense month) and in July 2007 and August 2007 (the two months of populated data in *current pay* (C01) months immediately following the suspense episode). Averaging those values ($125, $100, $100, $25) yields an estimated deemed income amount of $87.50, which was populated in the two months without deemed income data. By

Table VIII.3. An illustrative SSI recipient benefit and income scenario, BFWSSI_DRAFTyymm calculated using DAF16 Annuals method

Current Month	PSTA	STWSSI	Actual BM	Counter-factual BM	FBR	FCI (Actual BM)	FAMT (FBR-FCI in Actual BM)	EICM (Current Month)	UINC (Current Month)	Deemed + ISM (Current Month)1	TCNEI=FCI-EICM (Actual BM)	BFWSSI_DRAFT in DAF16 Annuals	
Dec-06	C01	0	Dec-06	Dec-06	603	200	403	0	75	125	200	EICM in actual BM	0
Jan-07	C01	0	Dec-06	Dec-06	623	200	423	0	25	100	200	EICM in actual BM	0
Feb-07	C01	0	Dec-06	Dec-06	623	200	423	200	50	*87.5*	200	EICM in actual BM	0
Mar-07	C01	0	Jan-07	Jan-07	623	125	498	400	100	*87.5*	125	EICM in actual BM	0
Apr-07	N01	1	Apr-07	*Feb-07*	623	725	0	625	50	50	100	FBR-UINC in Counterfactual BM	623-50=573
May-07	N01	1	May-07	*Mar-07*	623	850	0	800	25	25	50	FBR-UINC in Counterfactual BM	623-100=523
Jun-07	N01	1	Jun-07	*Apr-07*	623	675	0	625	0	50	50	FBR-UINC in Counterfactual BM	623-50=573
Jul-07	C01	0	Jul-07	Jul-07	623	540	83	400	40	100	140	EICM in actual BM	400
Aug-07	C01	0	Jul-07	Jul-07	623	540	83	400	40	25	140	EICM in actual BM	400
Sep-07	C01	0	Jul-07	Jul-07	623	540	83	200	50	50	140	EICM in actual BM	400
Oct-07	C01	0	Aug-07	Aug-07	623	465	158	0	100	75	140	EICM in actual BM	400
Nov-07	C01	0	Sep-07	Sep-07	623	300	323	200	50	50	340	EICM in actual BM	200
Dec-07	C01	0	Oct-07	Oct-07	623	175	448	200	50	75	465	EICM in actual BM	0

[1] Deemed income is estimated in the months shown in italics (using the method of averaging described in the text above). In other months, it is calculated by subtracting the current month's EICM and UINC from the FCI that corresponds to the current month (two months in the future in most C01 months and the current month in N01 months).

Table VIII.4. An illustrative SSI recipient benefit and income scenario, BFWSSI_DRAFT_BOTHyymm calculated using DAF16 SCWF method

Current Month	PSTA	STWSSI	Actual BM	Counter-factual BM	FBR	FCI (Actual BM)	FAMT (FBR-FCI in Actual BM)	EICM (Current Month)	UINC (Current Month)	Deemed + ISM (Current Month)[1]	TCNEI=FCI-EICM (Actual BM)	BFWSSI_DRAFT_BOTH in DAF16 SCWF
Dec-06	C01	0	Dec-06	Dec-06	603	200	403	0	75	125	200	EICM in actual BM 0
Jan-07	C01	0	Dec-06	Dec-06	623	200	423	0	25	100	200	EICM in actual BM 0
Feb-07	C01	0	Dec-06	Dec-06	623	200	423	200	50	*87.5*	200	EICM in actual BM 0
Mar-07	C01	0	Jan-07	Jan-07	623	125	498	400	100	*87.5*	125	EICM in actual BM 0
Apr-07	N01	1	Apr-07	*Feb-07*	623	725	0	625	50	50	100	FBR-(UINC+Deemed in Counterfactual BM) 623-(50+87.5)= 485.5
May-07	N01	1	May-07	*Mar-07*	623	850	0	800	25	25	50	FBR-(UINC+Deemed in Counterfactual BM) 623-(100+87.5)= 435.5
Jun-07	N01	1	Jun-07	*Apr-07*	623	675	0	625	0	50	50	FBR-(UINC+Deemed in Counterfactual BM) 623-(50+50)= 523
Jul-07	C01	0	Jul-07	Jul-07	623	540	83	400	40	100	140	EICM in actual BM 400
Aug-07	C01	0	Jul-07	Jul-07	623	540	83	400	40	25	140	EICM in actual BM 400
Sep-07	C01	0	Jul-07	Jul-07	623	540	83	200	50	50	140	EICM in actual BM 400
Oct-07	C01	0	Aug-07	Aug-07	623	465	158	0	100	75	140	EICM in actual BM 400
Nov-07	C01	0	Sep-07	Sep-07	623	300	323	200	50	50	340	EICM in actual BM 200
Dec-07	C01	0	Oct-07	Oct-07	623	175	448	200	50	75	465	EICM in actual BM 0

[1] Deemed income is estimated in the months shown in italics (using the method of averaging described in the text above). In other months, it is calculated by subtracting the current month's EICM and UINC from the FCI that corresponds to the current month (two months in the future in most C01 months and the current month in N01 months).

populating an estimated deemed income amount in the two months in C01 status prior to the first suspense month, we are able to revert to using the counterfactual budget month. In that case, looking to the values in February 2007 yields BFWSSI= FBR – (UINC + Deemed Income) = $623 – ($50 + $87.50) = $435.50.

Table VIII.5 summarizes the calculations described in each of the three scenarios to highlight where the three versions yield the same BFWSSI (current pay months), and in months that the calculations differ, the elements that go into making the calculation. We see that relative to the BFWSSI on the Annuals files, accounting for deemed income serves to reduce BFWSSI, as hypothesized.

Table VIII.5. A comparison of the BFWSSI calculations for the hypothetical beneficiary provided in Table VIII.3 using the Annuals and SCWF calculations

Current Month	BFWSSI_DRAFT on Annuals		BFWSSI_DRAFT_BOTH in SCWF	
Dec-06	EICM in actual BM (Dec-06)	0	Same as Annuals	0
Jan-07	EICM in actual BM (Dec-06)	0	Same as Annuals	0
Feb-07	EICM in actual BM (Dec-06)	0	Same as Annuals	0
Mar-07	EICM in actual BM (Jan-07)	0	Same as Annuals	0
Apr-07	FBR-UINC in Counterfactual BM (Feb-07)	573	FBR-(UINC+Deemed in Counterfactual BM (Feb-07))	485.5
May-07	FBR-UINC in Counterfactual BM (Mar-07)	523	FBR-(UINC+Deemed in Counterfactual BM (Mar-07))	435.5
Jun-07	FBR-UINC in Counterfactual BM (Apr-07)	573	FBR-(UINC+Deemed in Counterfactual BM (Apr-07))	523
Jul-07	EICM in actual BM (Jul-07)	400	Same as Annuals	400
Aug-07	EICM in actual BM (Jul-07)	400	Same as Annuals	400
Sep-07	EICM in actual BM (Jul-07)	400	Same as Annuals	400
Oct-07	EICM in actual BM (Aug-07)	400	Same as Annuals	400
Nov-07	EICM in actual BM (Sept-07)	200	Same as Annuals	200
Dec-07	EICM in actual BM (Oct-07)	0	Same as Annuals	0

5. Testing the effect of FCI on BFWSSI

As expected, the inclusion of deemed income appears to have reduced BFWSSI, as shown in Table VIII.6. In this table, we used data from DAF15 to first group SSI beneficiaries by their

BFWSSI_DRAFT value in January of 2013; the second column shows the number of beneficiaries in each group (i.e. based on the BFWSSI measure on the DAF15 Annual file).[24] The remaining three columns show how the new BFWSSI measure (BFWSSI_DRAFT_BOTH) in the same month compares; whether the value is less, the same, or higher than the BFWSSI_DRAFT value. In most cases, the BFWSSI value remains unchanged. To the extent there is a change, it most often leads to a reduction in BFWSSI. Though it is not possible to discern from the table, our inspections of the data show that most of the cases in the "less" group have a dollar value that is either in the same bin as the BFWSSI_DRAFT variable, or in the one just below. For these beneficiaries, the net effect of the new BFWSSI calculation is relatively small in dollar terms, at least within a single month. It is notable that 18.0 percent of those with BFWSSI_DRAFT between $600 and the FBR have a reduced BFWSSI under the new computation; most of that group moves to having no BFWSSI. This is similar to the share of beneficiaries who moved from STWSSI=3 to STWSSI_FCI=9, meaning that previously we would have calculated their BFWSSI as the FBR, and now we are categorizing them as off the rolls for a reason other than work, meaning they do not have BFWSSI.

[24] We did a similar comparison in DAF14 comparing to BFWSSI_DRAFT TO BFWSSI_DRAFT_BOTH and saw similar changes. As described at the start of the SCWF section, further exploration to consider the implications of changing the accounting of deemed income is warranted.

Table VIII.6. A comparison of the new and old BFWSSI measures, January 2013

BFWSSI_DRAFT1301	SSI Beneficiaries with BFWSSI_DRAFT in bin	BFWSSI_DRAFT_BOTH1301 compared to BFWSSI_DRAFT1301 (row percent)		
		Less	Same	More
$0[1]	18,770,883	NA	99.9	0.0
Under $10	10,724	8.5	90.7	0.8
$10 to $24	19,417	7.9	91.4	0.8
$25 to $49	32,351	9.8	89.7	0.4
$50 to $99	41,736	6.2	93.5	0.5
$100 to $199	53,965	6.1	93.3	0.5
$200 to $299	34,121	5.1	94.5	0.4
$300 to $399	25,873	3.6	96.1	0.3
$400 to $499	24,164	3.4	95.9	0.7
$500 to $599	11,761	4.6	95.1	0.4
$600 and above	64,504	16.9	83.1	0.0

Note: BFWSSI_DRAFT is the variable in the DAF16 Annual file, BFWSSI_DRAFT_BOTH is the revised version in the DAF16 SCWF that takes FCI and BMF into account. Note that in 2013, FBR was $710.

[1] The "same" value in this row includes cases that have missing or zero BFWSSI_DRAFT_BOTH; users of the SCWF can assume that when BFWSSI_DRAFT_BOTH is missing and STWSSI_FCI is non-missing, that the beneficiary had no BFWSSI_DRAFT_BOTH in the month.

IX. RACE AND ETHNICITY CATEGORIES

As of DAF15, all measures related to race and ethnicity derived from SSA administrative data have been removed from the file. Race, as contained on the RSA-911 files, has been retained, as it is collected from VR participants through the RSA data systems. SSA published statistics no longer include race data; the practice ended in 2009 for OASDI and 2002 for SSI.[25] Because administering programs does not require knowledge of a participant's race or ethnicity, the agency no longer requires individuals to provide such information. The fact that the information is collected only on a voluntary basis means that those who report the information may be a self-selected and non-representative sample.

The decision to omit race and ethnicity measures from the DAF was not taken lightly, as we recognize many users may be interested in the relative outcomes for certain groups. Yet, in addition to current problems with the collection of systematic race information across all applicants, the way in which SSA has collected race data has changed over time.[26] As a result, the data available in the DAF was significantly limited in its ability to measure race and ethnicity concepts, and was unable to produce useful comparisons across time.

[25] Martin, P. P. (2016). Why Researchers Now Rely on Surveys for Race Data on OASDI and SSI Programs: A Comparison of Four Major Surveys. *Research and Statistics Note No. 2016-01*.

[26] Scott, C. G. (2000). Identifying the Race or Ethnicity of SSI Recipients. *Social Security Bulletin, 62*(4), 9—20.

This page has been left blank for double-sided copying.

X. TICKET EVENT DATES

The SSA administrative data provide a history of events for the Tickets received by a beneficiaries in the TTW program. Historically, Ticket events began when a beneficiary was selected for inclusion in the TTW program, after which a Ticket was mailed to the beneficiary. Physical possession of a Ticket has never been a factor in Ticket eligibility, and between December 2012 and March 2015, SSA did not automatically mail Tickets—instead, they were only mailed upon request. Mailings to new beneficiaries resumed in April 2015, along with mailings to those celebrating their first, second, or third benefit anniversaries. Catch-up mailings were also sent to all beneficiaries who had not been mailed a ticket during the hiatus.

The DAF16 incorporates the data from multiple SSA source files to populate the TKTMAILDTn, reflecting the various ticket mailing schemes that have been in place over time. The initial mailings, those through 2012, exist on one DCF table, TKT. Additionally, mailings that occurred during the "hiatus" when SSA did not automatically send tickets are recorded in the TKT table. Mailings occurring after the hiatus, or those in 2015 or after, are carried on another DCF table, TKTSENT.[27] Anniversary mailings are not reflected in the TKTMAILDTn variable, but can be found in the TKTSENT table by using the variable TKTREMLDDTn with TKTMAILTYPn (which indicates the type of mailing). Beneficiaries who do not appear on the TKTSENT table, those that have not had a new initial, catch-up, or anniversary mailing, but do appear on the TKT table will have a TKTMAILDTn that reflects the mail date appearing on the underlying data from the TKT table. If that date occurs between 2013 and March 2015 then the

[27] TKTMAILDTn is populated with either the mail date on the TKT table or an initial or catch-up mail date from the TKTSENT table. When both are populated, the mail dates from TKTSENT taking precedent over those from TKT.

TKTMAILDTn value reflects the date the bene was selected for participation in the TTW program although no physical ticket was mailed.

Once a beneficiary is eligible for the TTW program, he can choose to assign his Ticket to a provider of employment services. The Ticket can, though does not need to, be later unassigned and re-assigned to another provider. The Ticket remains active until the participant stops receiving benefits for reasons other than work, such as medical recovery, retirement, death, or not meeting timely progress requirements stipulated by the TTW program. If one of those events occurs, the Ticket is terminated. If a beneficiary has a subsequent period of disability eligibility, he can receive another Ticket, and event dates for each subsequent Ticket are also tracked. For example, if a beneficiary receives disability benefits for cancer, recovers, and then later has a new period of disability benefits due to an injury, there would be two periods of disability and two Tickets, with the first Ticket showing a terminated status and the second Ticket an active status.

A. Sequence of Ticket event dates and Ticket records in DAF

This section provides more information about Ticket event information, including some potential pitfalls, explains how the data are stored in the DAF, and illustrates Ticket events more fully.

Though not always the case, Ticket events are expected to proceed in the following logical sequence: (1) beneficiary becomes eligible for the TTW; (2) Ticket is mailed;[28] (3) the beneficiary assigns Ticket to provider; (4) beneficiary unassigns their ticket (after completing services or prior to moving to a different provider); (5) beneficiary reassigns their ticket and then

[28] The practice of automatically mailing a physical Ticket was discontinued from June 2011 through April 2015. Between those dates, physical Tickets were only mailed to beneficiaries who requested them and every beneficiary who met the criteria for TTW was considered to have been mailed a Ticket at the same time as he or she was selected for the Ticket program (Ticket Mailing Date is equal to Ticket Selection Date).

(6) the Ticket is terminated. We present examples of how recorded data may deviate from this model in what follows.

During construction of the DAF, dates for all Ticket entries are numbered and re-organized within a reference system, based on Ticket assignments. Each numbered Ticket assignment entry consists of the mail date, assignment date, unassignment date (often blank), and termination date, using the following variables, each of which can have up to 18 occurrences:

- TKTMAILDDT1 to TKTMAILDDT18
- TKTASGNDDT1 to TKTASGNDDT18
- TKTUNASGDT1 to TKTUNASGDT18
- TKTTERMDDT1 to TKTTERMDDT18

Most participants have fewer than 18 Ticket assignment entries and the number of entries for each participant is reflected in the variable NOE (Number of Entries). Therefore a DAF record where NOE = 1 would use the following variables to store the SSA administrative Ticket data:

- TKTMAILDDT1
- TKTASGNDDT1
- TKTUNASGDT1
- TKTTERMDDT1

The remaining occurrences, 2-18, for each of the four variables would be blank. A record where NOE = 2 would use the following DAF variables:

- TKTMAILDDT1 and TKTMAILDDT2
- TKTASGNDDT1 and TKTASGNDDT2
- TKTUNASGDT1 and TKTUNASGDT2
- TKTTERMDDT1 and TKTTERMDDT2

The remaining occurrences, 3-18, for each of the four variables would be blank.

B. Examples of conceptual Ticket assignment information and corresponding DAF records

In this section, we present hypothetical examples of Ticket assignment data. Through these examples, we demonstrate that while the majority of cases will contain the expected series of data, sometimes anomalies occur. We do not attempt to clean records with seemingly contradictory information. Thus, researchers should expect that a small number of date values will be out of alignment with others and therefore sequences of events should be carefully examined before use for research purposes.

Example 1

Example 1 is a conceptual example of the SSA administrative data for the Ticket event dates for a participant, showing a Ticket was mailed in February 2005, assigned in July 2005, and terminated—perhaps due to medical recovery—in February 2007.

Table X.1a Example 1 conceptual Ticket assignment information

Ticket 1			
Mail Date	**Assign Date**	**Status Code**	**Status Effective Date**
Feb 05	Jul 05	T	Feb 07

In the DAF, the corresponding variables for Example 1 would appear as shown below; all 18 occurrences for each of the four Ticket date variables are present in a single wide record. Note that in this example there would be just one Ticket assignment entry, with data in just the first occurrence of three variables. The first unassignment date would be blank, as would all variables for occurrences 2-18.

Table X.1b Example 1 corresponding DAF variables

NOE	TKTMAIL DDT1	TKTMAIL DDT2 thru TKTMAIL DDT30	TKTASGN DDT1	TKTASGN DDT2 thru TKTASGN DDT30	TKTUNAS GDT1	TKTUNAS GDT2 thru TKTUNAS GDT30	TKTTERM DDT1	TKTTERMDDT2 thru TKTTERMDDT30
1	Feb 05		Jul 05				Feb 07	

Example 2

Sometimes a participant re-assigns a Ticket to another provider, in which case the cycle includes two intermediate steps, unassignment, and reassignment (Example 2).

Table X.2a Example 2 conceptual Ticket assignment information

Ticket 1					
Mail Date	Assign Date	Unassign Date	Re-assign Date	Status Code	Status Effective Date
Feb 05	Jul 05	Aug 05	Nov 05	T	Feb 07

Note that this example does not reveal whether the beneficiary assigned their Ticket to different providers or to the same provider twice, but the DAF does contain this information (though not shown here) in the DUNSn variable on the Ticket Component. In the DAF, as shown below, there would be two Ticket assignment entries. Both entries have the same mail date and termination dates, but unique assignment dates. The assignment date for the first entry is the date the Ticket was originally assigned while the assignment date for the second entry is the date the Ticket was reassigned.

Table X.2b Example 2 corresponding DAF variables

NOE	TKTMAILDDT1	TKTMAILDDT2	TKTMAILDDT3 thru TKTMAILDDT30	TKTASGNDDT1	TKTASGNDDT2	TKTASGNDDT3 thru TKTASGNDDT30	TKTUNASGDT1	TKTUNASGDT2	TKTUNASGDT3 thru TKTUNASGDT30	TKTTERMDDT1	TKTTERMDDT2	TKTTERMDDT3 thru TKTTERMDDT30
2	Feb 05	Feb 05		Jul 05	Nov 05		Aug 05			Feb 07	Feb 07	

Example 3

A beneficiary with multiple periods of eligibility may receive more than one Ticket (Example 3). Note that the first Ticket is terminated before the cycle begins again for the second

Ticket and that the current status of the second Ticket is assigned, as indicated by the "I" for Status Code.

Table X.3a Example 3 conceptual Ticket assignment information

Ticket 1				Ticket 2			
Mail Date	Assign Date	Status Code	Status Effective Date	Mail Date	Assign Date	Status Code	Status Effective Date
Feb 05	Jul 05	T	Feb 07	Feb 08	Aug 08	I	Aug 08

For this example also, there would be two populated occurrences of each variable in the DAF, with the rest blank.

Table X.3b Example 3 corresponding DAF variables

NOE	TKTMAILDDT1	TKTMAILDDT2	TKTMAILDDT3 thru TKTMAILDDT30	TKTASGNDDT1	TKTASGNDDT2	TKTASGNDDT3 thru TKTASGNDDT30	TKTUNASGDT1	TKTUNASGDT2	TKTUNASGDT3 thru TKTUNASGDT30	TKTTERMDDT1	TKTTERMDDT2	TKTTERMDDT3 thru TKTTERMDDT30
2	Feb 05	Feb 08		Jul 05	Aug 08					Feb 07		

Example 4

Occasionally, the administrative data that documents the Ticket events may exhibit some anomalies (Example 4), in which the first assignment, status code, and date fields contain no data. Furthermore, the value for the second assignment date field occurs prior to the second mail date. The likely scenario here is that the second assignment date is actually the first assignment date and possibly likewise for the second status code and status date.

Table X.4a Example 4 conceptual Ticket assignment information

Ticket 1				Ticket 2			
Mail Date	Assign Date	Status Code	Status Effective Date	Mail Date	Assign Date	Status Code	Status Effective Date
Feb 05				Apr 07	Jul 06	T	Jan 08

In the DAF, such a case would look like the data below: note that although it is likely that the second assignment date should pair with the first mail date, we cannot be sure of this and therefore *we do not attempt to apply corrections to the data*.

Table X.4b Example 4 corresponding DAF variables

NOE	TKTMAILDDT1	TKTMAILDDT2	TKTMAILDDT3 thru TKTMAILDDT30	TKTASGNDDT1	TKTASGNDDT2	TKTASGNDDT3 thru TKTASGNDDT30	TKTUNASGDT1	TKTUNASGDT2	TKTUNASGDT3 thru TKTUNASGDT30	TKTTERMDDT1	TKTTERMDDT2	TKTTERMDDT3 thru TKTTERMDDT30
2	Feb 05	Apr 07			Jul 06						Jan 08	

Example 5

In Example 5, the first assignment date occurs after the second mail date; in this case it is likely that the first Ticket was never assigned and that the first assignment date should actually be the second assignment date.

Table X.5a Example 5 conceptual Ticket assignment information

Ticket 1				Ticket 2			
Mail Date	Assign Date	Status Code	Status Effective Date	Mail Date	Assign Date	Status Code	Status Effective Date
Feb 05	Jul 08			Mar 07			

In the DAF, this data anomaly would look like the data below. Because we cannot be sure that the first assignment date actually belongs with the first Ticket, we do not attempt to apply corrections to the data.

Table X.5b Example 5 corresponding DAF variables

NOE	TKTMAILDDT1	TKTMAILDDT2	TKTMAILDDT3 thru TKTMAILDDT30	TKTASGNDDT1	TKTASGNDDT2	TKTASGNDDT3 thru TKTASGNDDT30	TKTUNASGDT1	TKTUNASGDT2	TKTUNASGDT3 thru TKTUNASGDT30	TKTTERMDDT1	TKTTERMDDT2	TKTTERMDDT3 thru TKTTERMDDT30
2	Feb 05	Mar 07		Jul 08								

This page has been left blank for double-sided copying.

www.mathematica-mpr.com

Improving public well-being by conducting high quality, objective research and data collection

PRINCETON, NJ ■ ANN ARBOR, MI ■ CAMBRIDGE, MA ■ CHICAGO, IL ■ OAKLAND, CA ■ SEATTLE, WA ■ TUCSON, AZ ■ WASHINGTON, DC ■ WOODLAWN, MD