SOCIAL SECURITY ADMINISTRATION PUBLIC-USE MICRODATA SAMPLES

Brian Greenberg, Cherice Jefferies, and Jennifer Lee
Office of Research, Evaluation, and Statistics, Social Security Administration, 4-C-15 Operations Building,
6401 Security Boulevard, Baltimore, MD 21235

KEY WORDS: Public-Use Microdata Files; Social Security Administration; OASDI; SSI

ABSTRACT

Public-use microdata files are typically subsamples of survey or administrative data that are released to the public after application of procedures to prevent individual records from being linked to the person, business firm, or other entity to which they pertain. These files are valuable because they allow for a wide range of statistical analysis that cannot be met by tabular information prepared as part of a standard statistical publication program or by other special purpose tabulations. The Social Security Administration is making available two public-use microdata files drawn from Social Security program files. One file is based on administrative data for the Old Age, Survivors, and Disability Insurance (OASDI) program and the other on administrative data for the Supplemental Security Income (SSI) program. This paper discusses the creation of these public-use files, how to access them, their content, and future plans.

The Social Security Administration's Office of Research, Evaluation, and Statistics publishes extensive statistical information about the OASDI and SSI programs. Publications include background information about these programs as well as a wide range of tables and related summary tabulations. The two public-use microdata files provide data users additional analytic capabilities for research on these programs.

1.   Introduction

Social Security Administration's (SSA's) Office of Research, Evaluation, and Statistics (ORES), in the Office of Policy, publishes a wide range of statistical reports describing Social Security programs, available on the Office of Policy Web-site (www.socialsecurity.gov/policy/). They provide detailed statistical tables describing the

SSA programs and they also provide extensive program background information. A discussion of ORES statistical publications, is contained in Parker (2002). In addition to publishing statistical reports, the Office of Policy publishes a wide range of research studies, funds external research, and contributes to ongoing analysis of Social Security programs. Information on these activities can also be found on the Office of Policy Web-site.

Microdata files consist of data records that contain information about an individual person, business firm, or other entity. Public-use microdata samples are typically subsets of survey or administrative data that are released to the public with all personal identifiers removed and after application of procedures to prevent records from being linked to the individual person, business firm, or other entity to which they pertain. That is, all records are non-identifiable. Public-use microdata files are valuable because they can be used for statistical analysis and research studies that cannot be carried out using aggregate information. They provide researchers the ability to create summary information to meet specialized needs and allow them to investigate relations among variables that are not apparent based on aggregate information.

This paper describes two newly created public-use microdata files that have recently been released by ORES, which are based on Social Security administrative data files. These microdata files augment and enrich the range of products made available by the Office of Policy for the analysis of Social Security programs. These two public-use microdata files provide information about Social Security's two major programs – Old Age, Survivors, and Disability Insurance (OASDI) and Supplemental Security Income (SSI).

This paper describes the process of creating these files and provides information on how they can be accessed. It discusses the content of each file, discusses differences between them, and

addresses the reasons for differences. Also described are the sample design, estimation, confidentiality protection for these two files, and plans for the future. In addition, your comments and experiences on the use of these files are needed to assist in planning future releases, and contact information is provided.

2. Background Information on Social Security Programs – All available on the Office of Policy Web-site

OASDI is the familiar Social Security social insurance program into which participants make payroll contributions based on earnings. Benefits are paid to insured workers and eligible family members when they retire or become disabled and to the survivors of deceased workers. Extensive information about the OASDI program can be found in the following publications:

Annual Statistical Supplement to the Social Security Bulletin, and

 Social Security Handbook.

The SSI program is a means-tested income assistance program for disabled, blind, and aged persons who have limited income and resources regardless of their prior participation in the labor force. This program is administered in conjunction with the states. Extensive information about the SSI program can be found in the *Annual Statistical Supplement,* the *Social Security Handbook* and the following publications:

Understanding Supplemental Security Income,

SSI Annual Statistical Report,

Annual Report of the Supplemental Security Income Program, and

State Assistance Programs for SSI Recipients.

3. OASDI Public-use Microdata File

Information necessary to administer the OASDI benefit program resides on SSA's Master Beneficiary Record (MBR) file. In particular, the file includes data used to determine program eligibility and entitled benefits amounts. In

December 2001, the MBR contained records for approximately 46 million individuals who were entitled to receive an OASDI benefit for that month.

The OASDI Public-Use Microdata File contains an extract of data fields from the MBR and consists of a 1- percent random, representative sample of persons entitled to receive an OASDI benefit payment for December 2001. Records on the underlying file for persons entitled to receive an OASDI benefit payment for December 2001 were ordered by state, by sex of beneficiary within state, by age within sex, and by type of benefit within age, and then a systematic random sample was selected. The File contains approximately 460,000 records, with 16 data fields on each record, and is available in a standard SAS file structure.

Since the sample for the OASDI Public-Use Microdata File is a representative 1-percent systematic random sample of an OASDI benefit payment records for December 2001, all records have weight equal to one hundred. Variance and standard errors can be approximated with the standard formulas used for simple random sampling.

4. SSI Public-Use Microdata File

Recipient information needed to administer the Federal SSI program resides on the Supplemental Security Record (SSR) file. In particular, this file contains data used to determine federally administered SSI benefit payments. In December 2001, the SSR contained records for approximately 6.4 million individuals who received a Federal SSI benefit for that month. This SSI Public-Use Microdata File represents only persons receiving a Federal SSI benefit payment in December 2001. Persons receiving a federally administered State supplementary payment but no Federal SSI benefits in December 2001 are not represented by this file and are not included on it.

The SSI Public-Use Microdata File contains an extract of data fields from the SSR and consists of a 5-percent random, representative sample of persons who received a Federal SSI benefit in December 2001. Records in the underlying SSR for persons who received a Federal SSI payment for December 2001 were ordered by state, then by disability diagnosis within state, then by age

within disability diagnosis, and then by sex within age, and then a 5-percent systematic random sample was selected. The file contains approximately 320,000 records, with 13 data fields on each record, and is available in a standard SAS file structure.

Since the sample for the SSI Public-Use Microdata File is a representative 5-percent systematic random sample of persons receiving a SSI payment in December 2001, all records have weight equal to twenty. Variance and standard errors can be approximated with the standard formulas used for simple random sampling.

5. Data fields on the OASDI and the SSI Public-Use Microdata Files

The data fields on the OASDI and SSI public-use microdata files are listed below with a brief description of each field to give potential users an idea of what type of information the file contains and what uses it may meet. More complete descriptions are provided in the Data Dictionary for the respective files, which describe the content of each field, method of presentation, and disclosure avoidance steps taken to provide confidentiality. Note that the information on each file refers to a single time period – December, 2001.

A. OASDI Public-Use Microdata File

YOB   Year of Birth of beneficiary – information provided by single year of birth

SEX   Sex of beneficiary

YOCE   Year of Current Entitlement – year beneficiary met the applicable requirements for the receipt of current type of benefits, including the filing of an application.

TOB   Type of Benefit – identifies the current type of benefit based on entitlement criteria. The major types of benefits are: retired workers, disabled workers, auxiliary benefits (spouses and children of retired and disabled workers), and adult and child survivors. Twenty-three detailed benefit types are specified in the file, including divorced spouses and minor, student and disabled adult children.

PIA   Primary Insurance Amount – the amount, related to a worker's earnings history, used as a base for computing benefits.

TPIA   Type of PIA – type of calculation, related to a worker's age and earnings history, used to determine the PIA.

MBC   Monthly Benefit Credited – monthly amount payable to a beneficiary, after any reduction for age, family maximum or other reasons, but before any deduction for Supplementary Medical Insurance premiums.

BROA   Benefit Reduction Based on Age of Beneficiary – indicates whether the monthly benefit payable was reduced due to entitlement before full retirement age.

BRADS Benefit Reduction for Age of Deceased Spouse – indicates whether the widow(er)s monthly benefit was reduced due to the deceased spouse's entitlement before full retirement age.

DRI   Delayed Retirement Indicator – indicates whether the monthly benefit was increased due to nonreceipt of benefits after full retirement age.

DEI   Dual Entitlement Indicator – indicates whether the beneficiary is entitled to more than one type of benefit. Most dual entitlements are due to retired or disabled workers who are also entitled to a higher spousal or widow(er) benefit.

DEOP   Dual Entitlement to Other Program – identifies other type of OASDI program -- retirement, survivor, disability -- for dually entitled beneficiaries.

OTOB   Other Type of Benefit – identifies other type of benefit; for worker, spouse, widow(er), or child, for dually entitled beneficiaries.

DEBI   Dual Entitlement Beneficiary Indicator – provides additional information concerning spousal or survivor benefits for dually entitled individuals.

LEMBA Larger Excess Monthly Benefit Amount – indicates difference between the larger (usually auxiliary or survivor benefit) and the smaller (usually worker) benefit for dually entitled individuals.

SAMBA Smaller Actuarially Reduced Monthly Benefit Amount – indicates amount of the smaller (usually worker) benefit for dually entitled individuals.

B. SSI Public-Use Microdata File

STAT State of residence of recipient – 50 states and District of Columbia

YOB Year of birth – provided by single year of birth

SEX Sex of recipient

DIAG Primary diagnosis – indicates the category of primary diagnosis which is the basis for blindness/disability benefits for recipients age 64 or younger. Four broad categories are provided: diagnosis unspecified/aged, mental retardation, mental disorders other than retardation, and physical diseases/disorders.

DTEL Date of eligibility – the first year in which the recipient is eligible to receive payments.

PREL Program eligibility – aged or disabled.

HOUS Household composition – program eligibility category and household composition of recipient. Household composition is used for benefit computation and includes only household members whose presence affects the payment.

LIVA Living arrangement code – recipient's federal living arrangement for December 2001. This information is used to determine the recipient's SSI Federal Benefit Rate, which is the maximum allowable monthly Federal benefit if the person has no countable income.

FPMT Federal Payment – federal money amount paid to a recipient in December 2001. This amount may include back payments or payment reductions due to overpayment recovery.

FCNT Federal countable income – total amount of countable earned and unearned income that was charged in recipient's benefit payment computation for December 2001.

SSIN Social Security income – amount of unearned income from Social Security benefits (OASDI).

UEIN Unearned income other than Social Security income – amount of unearned income received from all sources excluding Social Security income.

ERIN Earned income – earned income a recipient receives through wages or earnings from self-employment.

6. Uses of These Files and Contrasts Between Them

The OASDI Public-Use Microdata File will be valuable for understanding the beneficiary population and to study the effects of current and proposed legislative and program provisions, particularly as they relate to benefit analysis. The variables included in the file allow analysts to study the relationships between factors that affect the benefit computation that must be considered in these studies.   For example, benefit amounts for retired workers are based on complex calculations related to early retirement, delayed retirement, and dual entitlement to other types of benefits.   The file may also be used to study how certain provisions have different effects on benefits for men and women or older and younger beneficiaries.

The SSI Public-Use Microdata File will be valuable for an understanding of the SSI population nationwide and in each state. One can use the file to obtain demographic and income characteristics of people who received a Federal SSI benefit payment in December 2001. In addition to limited income and resources, individuals under age 65 must be blind or disabled to be eligible for SSI. The DIAG field indicates the category of primary diagnosis for these recipients.   The SSIN field contains information on the individual's Social Security income (i.e., OASDI income) and is useful for identifying and analyzing the characteristics of SSI recipients who also receive OASDI benefits.

One major difference between these two files is that the SSI file provides state-of-residence information and the OASDI file does not. The SSI file provides the state of residence for persons receiving a Federal SSI benefit payment in December 2001 to allow for analysis by state. In contrast, benefit computations for the OASDI program are on the national level and are independent of state.

The benefit amount on the OASDI file represents the amount a person is entitled to receive for December, 2001 based on earnings history and entitlement criteria. This benefit includes the 2.6 percent cost-of-living adjustment effective for December 2001 and was payable in January 2002. The payment amount on the SSI file represents the amount a person actually received in December 2001. This amount may include back payments or payment reductions due to overpayment recovery. The SSI program also had a 2.6 percent cost-off-living adjustment effective with January 2002 payments.

7.   Confidentiality Protection and Data Quality

On these public-use microdata files individual identifiers have been removed from all records and other distinguishing characteristics have been modified to prevent identification of persons to whom a record pertains. Records are sequenced in random order. Topcodes and bottomcodes are employed for all numeric fields to avoid showing extreme field values on a data record. Values beyond the topcode or bottomcode are replaced by the average of the values in excess of the respective topcode or bottomcode. Topcode and bottomcode values were derived at the national level and the replacement values are derived and applied at the state level for the SSI file and at the national level for the OASDI file. Values shown for categorical fields are groupings of detailed information from the internal file and numeric variables are rounded to the nearest $5. All topcode and bottomcode values, replacement values, and information on detail groupings are provided to users as part of file documentation.

A Disclosure Review Board (DRB) reviewed these proposed Public-Use Microdata Files prior to release. The DRB consisted of staff from ORES who are familiar with the underlying program files and their uses and with confidentiality requirements and staff from the Census Bureau and the National Center for Health Statistics to serve as external confidentiality experts. Part of the DRB preparations included completion of the *Checklist on Disclosure Potential of Proposed Data Releases,* prepared by Interagency Confidentiality and Data Access Group. The DRB review determined that information on these files is suitable for release.

Because information on the internal MBR and SSR record files is used to administer the OASDI and SSI programs, data quality on both files is extremely high as data are validated and kept current. For detailed information about the MBR and SSR as well as other SSA programmatic data files, see Panis (2000).

8.   Data Access and Future Plans – We Welcome Your Comments

These two files are available in SAS format and accessed directly from the Office of Policy Web-site (www.socialsecurity.gov/policy/), which also contains supporting documentation. In addition, these two files will be available in the Census Bureau's public-use microdata access system, FERRET.

The files discussed here represent SSA's first effort to produce large public-use microdata files from the SSR and the MBR. Since these files include only a small fraction of records from these administrative files, the opportunity exists to select non-overlapping samples with which to create alternate public-use microdata files. In particular, for the OASDI program, a file is planned with all the benefit data shown on this file along with earnings histories. This new file would allow users to better understand the relationship between earnings and benefit amounts for various classes of beneficiaries.

It is important that users of these files let us know their experiences in using them and what they would like to see on subsequent releases of public-use files based on the MBR and SSR, so that future releases can respond to user needs. Comments, suggestions and experiences concerning these files can be sent to the e-mail addresses provided on our web-site.

REFERENCES

Panis, Contantijn et al. 2000. *SSA Program Data User's Manual.*, Baltimore: Social Security Administration. Office of Research, Evaluation, and Statistics.

Parker, Robert P. 2002. Statistical Reports from the Social Security Administration. *Business Economics*. 37(1):64-5

Social Security Administration. 2000. *State Assistance Programs for SSI Recipients*. Washington, D.C.: U.S. Government Printing Office.

Social Security Administration. 2001. *SSI Annual StatisticalReport* . Washington, D.C.: U.S. Government Printing Office.

Social Security Administration. 2001. *Social Security Handbook*, 14th ed. Washington, D.C.: U.S. Government Printing Office.

Social Security Administration. 2002. *Annual Statistical Supplement, 2002,* to the *Social Security Bulletin*. Washington, D.C.: U.S. Government Printing Office.

Social Security Administration. 2003. *Annual Report of the Supplemental Security Income Program*. Washington, D.C.: U.S. Government Printing Office.

Social Security Administration. 2003. *Understanding Supplemental Security Income*. Washington, D.C.: U.S. Government Printing Office.