SURVEY OF CONSUMER FINANCES

EMPLOYER SPONSORED PENSION BENEFIT PLANS

PENSION ESTIMATION PROGRAM DOCUMENTATION

Richard T. Curtin

Survey Research Center

University of Michigan

TABLE OF CONTENTS

LIST OF TABLES

PROGRAM OVERVIEW

This program is designed to estimate the pension entitlements held by respondents in the Survey of Consumer Finances, based on the plan formulas and benefit provisions obtained from the linked sample of pension providers. The pension program uses systems of equations to represent each of the pension plans, including all benefit formulas and payment provisions. These equations, in turn, use as input the work and income histories of the respondents. When combined with assumptions needed for the estimation of the present value of future benefit flows, the program calculates the appropriate pension entitlements, and prepares output data files for subsequent analysis.

Program users have complete control over the input data and estimation parameters, providing the capability of a wide array of analytic strategies. Estimates can be made for any set of respondent employment characteristics, whether based on the actual survey data or simulated profiles. The pension entitlements due from any one or from every pension plan can be calculated. The program provides estimates of all types of retirement benefits—normal retirement, early retirement, late retirement, disability retirement, and surviving spouse benefits. The date of employment termination—which must be assumed for most participants since they were still working at the time of the survey—can be set or limited to any given age range or calendar dates. In addition to estimates based on a single quit date, the program can also produce an estimate of the benefit amount due at every age from the time of the survey until the participant reaches the maximum quit age allowed. These yearly estimates cover all types of retirement, disability and survivor benefits.

This document describes the program, and defines all user commands and analysis options. Although this document has been tailored for users of the University of Michigan's computer system, all of the program commands and options would be identical if installed on another computer system. Michigan's computer system is easily accessible through the MERIT network, with local access in many cities, including Washington DC. The programs can then be used on MTS or downloaded to another computer installation.

PROGRAM EXECUTION COMMANDS

The commands required to run the program are just two: The first loads the Pascal program object file SPFN:PP4.0, which contains the pension plan formulas and benefit requirements, and the second gives the user defined options for each replicate of the analysis desired. The user defined variables include the income and work history characteristics of the participants, and the behavioral and economic assumptions needed to estimate the present discounted value of pension entitlements. All of the user defined commands and options are specified on a single input line for each replicate analysis desired. The program is executed by issuing the following commands:

```
$RUN SPFN:PP4.0
RUNTYPE INDATA INPARM INPLAN OUTDATA OUTTAB OUTERR
( second replicate )
( third replicate )
(      etc.       )
$ENDFILE
```

The file SPFN:PP4.0 holds the executable Pascal (JB) object program. All of the user options are specified on the second, and subsequent input lines. The RUNTYPE specifies the type of analysis to be performed, and the input files INDATA, INPARM, and INPLAN specify the data cases, parameters, and pension plans to be used in the estimation process. The INPLAN file should only be specified when the program estimates are based on simulated data cases, and must be excluded when the program estimates are based on actual participant data. The output from the program is written to the files OUTDATA, OUTTAB, and OUTERR. The OUTDATA file contains the estimated entitlement amounts in a format suitable for input to analysis programs. The OUTTAB file has the results formatted for printed output, including table headings, and a summary of the input parameters. The OUTERR file gives information on all calculations that, while not necessarily incorrect, are flagged for potential problems. The output files must always be defined by the user, following standard MTS file name conventions.

The names used to designate the various files are not unique; any permissible file name can be used. The program identifies the files by the order in which they are listed on each command line. The first entry on each line must be the RUNTYPE, followed by the file names which

contain the input data (INDATA), the file containing the estimation assumptions (INPARM), the file containing the selected pension plans (INPLAN), the file to write the output data set (OUTDATA), the output tables (OUTTAB), and estimation problems (OUTERR). The order of the listed files must be followed exactly, and no file name can be omitted from the list, except for INPLAN which is only allowed for analysis based on simulated data cases.

Since replicate analyses will often be desired, for example to compare benefit entitlements under different assumptions about future economic conditions, one execution of the program will perform as many separate analyses as required.

The commands can be entered directly from the terminal keyboard, or entered into a file which can then be $Sourced or submitted as a batch job. With keyboard entry, the $Endfile command is required. When using a source file, the $Endfile command is not required if it would have been the last line in the file; if other commands follow (for example, to print the output), the $Endfile is required—it is needed by the program to signal the end of the input stream, and does not terminate processing of the $Source file.

The files SPFN:INDATA, SPFN:INPARM, and SPFN:INPLAN are the program supplied default input files. Two versions of the INDATA file are provided: SPFN:INDATA.SCF, the default file for actual participant data derived from the Survey of Consumer Finances, and SPFN:INDATA.SIM, the default file of simulated participant characteristics. The file SPFN:INPLAN contains the pension plan IDs for all the pension plans included in the sample. These files are supplied for easy startup and learning exercises. Normally, users would generate their own versions of these files, so as to reset each of the data values and estimation parameters to the appropriate levels for the particular analysis desired. The default files SPFN:INPARM, SPFN:INDATA.SCF, SPFN:INDATA.SIM, and SPFN:INPLAN are permitted read access so that users can duplicate these files. The user's copy of the these file can then be edited to reflect the desired changes in the input data, pension plans, or estimation parameters. The SCF data base can be used to generate a new INDATA.SCF file, using OSIRIS, as described below.

RUNTYPE is the only option that does not represent a file name, but rather a numeric code

which selects the type of analysis to be performed. The values RUNTYPE can take are:

1.2 when the input is actual SCF respondent and spouse data, which is to be matched to the specific employer sponsored pension plans in which they participate, with all estimates based on an assumed quit date.

3.2 when the SCF respondent and spouse data is matched to the specific plans in which they participate, but no single quit date is assumed. Instead of one estimate, this program option produces the estimated pension entitlement that would accrue if the participant quit in 1983, or in any of the following years until reaching the maximum quit age set in the INPARM file. This output thus gives the profile over time of potential pension entitlements depending on the date of job termination.

4.2 when the input data is for simulated participant income and work histories. For each set of simulated participant characteristics, the benefit entitlement that would be due under every pension plan included in the INPLAN file is calculated. This option requires an assumed quit date.

If the RUNTYPE is entered incorrectly, the program will produce a error message indicating a "bad version ID." If the name of one or more of the input/output files does not exist, MTS will query for a replacement.

An example of the execution of the program, estimating one replicate for each of the three RUNTYPES, and using the default input files is as follows:

```
$RUN  SPFN:PP4.0
1.2 SPFN:INDATA.SCF SPFN:INPARM OUTDATA1 OUTTAB1 OUTERR1
3.2 SPFN:INDATA.SCF SPFN:INPARM OUTDATA2 OUTTAB2 OUTERR2
4.2 SPFN:INDATA.SIM SPFN:INPARM SPFN:INPLAN OUTDATA3 OUTTAB3 OUTERR3
$ENDFILE
```

In the following example, the program is executed with two replicates, each using the same data file, but two separate input parameter files. The data are written to separate output files for later analysis. The input data used is the actual participant data matched with the pension plans in which they participate. The purpose of the two replicates would be the comparison of entitlements under the two different sets of input assumptions. Of course, as many replicates as desired may be used, rather than just the two used in the following example:

```
$RUN  SPFN:PP4.0
1.2  SPFN:INDATA.SCF  MY.INPARM1  OUTDATA1  OUTTAB1  OUTERR1
1.2  SPFN:INDATA.SCF  MY.INPARM2  OUTDATA2  OUTTAB2  OUTERR2
$ENDFILE
```

When the same file is used as in the immediately prior replicate, the program can be signaled to use the same file by using an "*" instead of the file name. All of the file names can

represent temporary files, and if any of the output files are not wanted, the file name *dummy* can be used. In the example below, two replicates use the same input data file (a file created by the user), two different user defined input parameter files, two different temporary files to store the data output, one temporary file to write estimation problems for both replicates, and the "dummy" file name to suppress the table output.

```
$RUN SPFN:PP4.0
1.2 MY.INDATA MY.INPARM1 -OUTDATA1 *DUMMY*  -OUTERR1
1.2     *     MY.INPARM2 -OUTDATA2 *DUMMY*     *
$ENDFILE
```

Multiple data sets can be stored in one file, with the program instructed to read data starting at a specific line number. The MTS command

$CONTINUE WITH *DUMMY*

should be used to terminate each data set listing within the file (note that $Endfile will not work for this purpose). An example of the use of this option is shown below for the simulated analysis which uses the same input data file for all the replicates, but directs the program to read different line ranges in the file. Also note that the user file MY.PLAN can constrain any selected subset of the pension plans on which to base the simulated data estimates.

```
$RUN SPFN:PP4.0
4.2 MY.SIM(100) MY.INPARM MY.INPLAN OUTDATA1 OUTTAB1 OUTERR1
4.2 MY.SIM(200)    *    *    OUTDATA2     *       *
4.2 MY.SIM(300)    *    *    OUTDATA3     *       *
$ENDFILE
```

If the commands are entered into files to be $Sourced or run as batch jobs, other commands may follow, as shown in the following example, which sends the OUTTAB file to the printer:

```
$RUN SPFN:PP4.0
3.2 NEWDATA NEWPARM NEWOUTDA -PRINT NEWPROB
$ENDFILE
$COPY -PRINT *PRINT*
```

As can be seen from the above examples, the actual names of the files do not matter to the program; it is the RUNTYPE and the order of the given file names that the program uses to distinguish and identify each file.

INPUT FILES

Participant Data

INDATA represents the file name where participant input data is stored. The program expects the input data to be in free format (data values separated by blanks), with one participant/plan combination per line. If a respondent was covered by more than one pension plan, the data should appear on two or more separate lines, each associated with its own unique pension plan ID. And since pension entitlements were recorded for both the respondent and the spouse, a single household record is often the basis of multiple input lines. For each the respondent and the spouse, up to four pension plans were coded; thus one household could provide up to eight separate input records.

The program treats each line in the INDATA file as a separate case, and will only calculate estimates for those cases included in this file. The number of cases included in the INDATA can be as few as one or as many as the full sample of participants. Any subset of participants, in any order, can be selected for inclusion in the INDATA file.

The program option RUNTYPE determines the type and order of input data expected by the program. When the analysis uses the actual household data, the program expects each data line to begin with the IDs which identify the specific pension plan which covers the participant. When the program uses simulated participant data, these ID variables are not allowed since estimates are calculated for all pension plans listed in the INPLAN file.

In addition to the data necessary to calculate pension entitlements, the input data file can also contain additional variables that will be added to the output data set. These "pass through" variables can be entered in any format, but can not exceed 132 columns in total, counting blanks. This option facilitates the easy analysis of the output data file, since there is often no need to remerge the output data files with the original SCF data files. The additional variables are "passed through" to the output data set only when the RUNTYPE is designated for actual household data with a given quit date (RUNTYPE 1.2). The other RUNTYPEs ignore the presence of the "pass through" data in the input files.

Each line in the INDATA file must list the input variables in a fixed order, separated by blanks. The required order and variable definitions are given in Table 1. The first three variables are optional, depending on the RUNTYPE specified. These linking IDs are needed when actual participant data are matched with the specific pension plans under which they are covered (RUNTYPE 1.2 or 3.2), but are not needed when simulated data is used (RUNTYPE 4.2). The last variable listed is also optional, and is used to define the "pass through" variables.

Table 1

Participant Data
(INDATA File)

Order		Var Name	Definition
Act	Sim		
1		HHIDX	SCF household ID.
2		CODEID	Pension plan ID.
3		SEQ#	ID of pension plan/provider.
4	1	SPOUSEBD	Birth date of spouse, if married.
5	2	SEX	Sex of the participant.
6	3	BIRTHD	The participant's date of birth.
7	4	HIRED	The date the participant was hired.
8	5	QUITD	Date the participant terminates employment.
9	6	WKHRS	Average annual paid work hours.
10	7	WAGE	Annual dollar amount of salary and wages.
11	8	WAGEG	Differential wage growth rate.
12	9	CRVOL	Voluntary contributions.
13		PT	Pass through variables.

The three IDs are from the Survey of Consumer Finances data file; the program will detect and reject any cases that do not have valid IDs. The HHIDX is the ID for the SCF household, with the digit "X" appended to indicate the data record of the respondent (X=1), or the spouse (X=2). The CODEID is the ID of the pension plan under which the household participant is covered. The SEQ# is the ID given to each unique plan/provider combination.

The SPOUSEBD variable is defined as the marital status of the participant at the time of

retirement. If the participant is single at the time of retirement, this variable should be set equal to zero; if married, the variable should be set equal to the birth date of the spouse. The participant's date of birth, date of hire, and quit date are defined as BIRTHD, HIRED, and QUITD, respectively. All date variables are entered as real numbers, using decimals to indicate fractions of years. For example, April 1, 1965 would be entered as 1965.25, and July 1, 1983 would be entered as 1983.50. The sex of the participant was defined as one for males, and two for females.

The WKHRS variable is defined as the average annual number of paid work hours over the relevant horizon for pension calculations. The major impact of the WKHRS variable is in determining the accrual rate for service credit. This variable should be set to the average number of hours worked per year, counting all paid vacation, holiday, and sick time. For a full time worker, this would usually be 2,080 hours per year.

The WAGE variable is defined as the annual amount of salary and wages covered under the pension plan as of the time of the survey if the participant was still employed, or wage level in the final year of employment if the participant had already terminated employment but had not yet begun to receive retirement benefits. From this base level, two separate income growth factors are applied to calculate the level of WAGE over time. The first income growth factor is specified in the input parameter file. This parameter is defined as the overall average real rate of growth in wages, and is applied to all participant data. The second income growth factor, specified in the input data file, is designed to capture differences in growth rates across population groups. The input variable WAGEG is defined as the deviation from the overall mean for specific participants, given their age, income and employment histories, and other factors. Since the WAGEG variable can be constructed so as to sum to zero across all participants included in the analysis, the overall total rate of income growth can be fixed via the input parameter file.

The CRVOL variable is defined as the amount of voluntary contributions made by the participant, expressed as a percentage of the participant's annual wages. Since the majority of pension plans do not include any provisions for voluntary contributions, this variable equals zero

for most cases. For participants that do make voluntary contributions, the CRVOL variable defines the average yearly contribution rate over the time horizon used for pension calculations. For pension plans that require contributions from participants, the mandatory contribution rates are built into the pension program, and thus do not need to be specified in the input data file.

The default input data file INDATA.SCF is based on certain assumptions concerning the participants. First, the QUITD included is defined as the actual quit date if the participant had already terminated employment, and the respondent's own estimate of when he/she would quit if they were still employed at the time of the household survey. Although the default file includes the expected QUITD, this information can be independently controlled via the input parameter file. Second, the SPOUSEBD and WKHRS variables use the respondent's current marital status and current work hours as estimates of future marital status and average work hours. Current annual salary and wages were used to estimate WAGE if the participant was still employed, otherwise the annual wage in the final year of employment was used. The WAGEG variable was calculated from estimates of life time earnings profiles, and adjusted to reflect differences from the overall mean of 5.5%. The CRVOL variable used the current rate of voluntary contributions reported by the respondent as an estimate of the average yearly contribution rate.

Creating Participant/plan Input Records

Osiris can also be used to create a new INDATA file from the Survey of Consumer Finances data set. The required commands and data transformations are included in the public use file SPFN:OS.DATASCF. This file can be copied, and edited to reflect any desired change in the definition of the input variables, or in the sample of cases included.

If users prefer, other programs like SPSS, SAS, and MIDAS, can be used to create these participant/plan input records. If the subsequent analysis of the pension entitlement data will be done using one of these or some other software package, it would probably be more efficient to use the same software package to generate these input records. Any program that can output free format data records, which can be done by most software packages, can be used to generate the participant/plan input data file.

Estimation Parameters

INPARM represents the name of the file which contains the user defined parameters needed in the estimation process. The parameter settings must be given in a fixed order on each line. Each line must start with the parameter settings and may be followed by a comment. Only the parameter settings on each line are read by the program, the rest of the line is ignored. The required contents of the INPARM file are shown in Table 2.

Table 2

Estimation Parameters
(INPARM File)

Line	Name	Default	Definition
1	SURVEYD	1983.50	Date of survey observations
2	TODAYD	1983.50	Base date for present value calculations
3	CPIrate	0.00	Annual inflation rate
4	INTrate	4.00	Real interest rate
5	INCrate	1.50	Real wage and salary growth rate
6	SSWB	34700.00	SS wagebase (as of survey date)
7	SSWBrate	1.50	Real SS wagebase growth rate
8	SS	5874.00	SS benefit (base amount)
9	SSrate	1.50	Real SS benefit growth rate
10	QAGErng	16 80	Quit age range (min, max)
11	QDATErng	1925 2075	Quit date range (min, max)
12	SSPUrto	0.50 1.00 1.50	SS primary unreduced (min, mid, max)
13	SSJUrto	0.75 1.50 2.25	SS joint unreduced (min, mid, max)
14	SSPRrto	0.40 0.80 1.20	SS primary reduced (min, mid, max)
15	SSJRrto	0.60 1.20 1.80	SS joint Reduced (min, mid, max)
16	SSPDrto	0.50 1.00 1.50	SS primary disability (min, mid, max)
17	SSJDrto	0.75 1.50 2.25	SS joint disability (min, mid, max)
18	SSSSrto	0.50 1.00 1.50	SS surviving spouse (min, mid, max)
19	BENMAX	35000 90000	Maximum benefits (age55, age65)
20	SSCUTS	17350 34700	SS wage cutoffs (min, max)
21	PROFCON1	5.00 0.50	Contributions (profits, forfeitures)
22	PROFCON2	3.00 0.30	Contributions (profits, forfeitures)
23	PROFCON3	1.00 0.10	Contributions (profits, forfeitures)
24	NACONrate	5.00 5.00	NA contributions (other, mandatory)
25	NAVOLrng	0.00 25.00	NA CRVOL range (min, max)
26	NAMATCH	0.00 0.00	% NA matching (MMAN, MVOL)
27	ANNTYPE	3.00 50.00	Annuity type / % of payment continued
28	VESTYPE	0.00 5.00	New vesting (1=on/0=off) / Yrs to vest
29	MRAGE	1.00	Plan maximum retirement (1=on/0=off)
30	PENTYPE	0.00	Pension type (0=NRT/ERT/VDT/LRT,1=DRT,2=DST)

The first two parameters are dates used to set the program's time reference. The first date is used to set the time frame for all input variables. This date defines the base period data values to which all growth factors are applied to estimate future trends. The second date defines the year on which to base the discounted present value calculations. For most purposes, both these dates should be set to equal the actual date of the SCF household survey (1983.5). If the date for present value calculations is set to a date later than the survey date, this would result in estimates of the present value of entitlements for subsequent years.

All growth factors are annual percentage rates of change, and are defined as the average rate over the relevant horizon for pension calculations. Setting the nominal rate of inflation along with the real rates of change in the other variables, gives users the ability to set the overall level of nominal rates of change while preserving real growth rates. In setting these values, the most important consideration involves the comparative size of the selected real growth rates for wages and interest rates. The former rate is used to increase amounts over time, while the latter is used to discount amounts back to the present. Although it is usually assumed that the former is lower that the latter, the selected size of the difference will have an important impact on the estimated results.

All growth rates are entered as average annual percentage amounts—for example, an annual growth rate of 3.5% is entered as 3.5. A zero real rate of growth yields a nominal growth rate equal to the rate of inflation. Growth rates may be set to either positive, zero, or negative values for the rate of inflation, income, and the Social Security wagebase and benefit amounts; the real interest rate, however, must be set to a value of greater than zero. This combination of optional settings allows the easy manipulation of the estimation assumptions, in a format that focuses the selection on both the nominal and real rates. As part of the program output, all the nominal rates are calculated and displayed in the file OUTTAB.

The inflation rate is intended as a common yardstick for assessing changes in all dollar values over time. The inflation rate selected represents the average rate of price increases over the entire relevant horizon for pension calculations. The real interest rate variable is defined as

the appropriate rate of discount to use in present value calculations. The real growth rate of income is defined as the average annual increase in salaries and wages received by plan participants, adjusted for inflation. Thus, all three variables are similar in that they signify general concepts, and are not identified with any specific price or wage index, or with a particular interest rate. Since an estimate for an extended future time period is often needed—say over the next 10 to 50 years—these rates should be chosen carefully. A needlessly high inflation rate, holding constant real growth rates, will only result in needlessly high nominal dollar values as the estimation time period lengthens.

The program options concerning the permissible quit age and date ranges are distinctive parameters in that they permit the censoring of the input data. When the maximum specified quit age is 65, the program will automatically cause all pension estimates to be based on a quit age of no higher than 65. If any of the input data records include a QUITD later than age 65, the data record will automatically be revised to reflect this age limit. Similarly, a limit on the quit date of 1983.5 would cause all pension estimates to be based on quitting as of the date of the survey. The minimum quit age/date option can be used to adjust the input data to restrict the minimum QUITD to age 55, or age 62 for example. The combined use of minimums and maximums can adjust all quit dates so that everyone retires at a certain age or on a certain date. For example, if all participants were assumed to retire at age 65, both the age minimum and maximum would be set equal to this age. Or if all pension entitlements were to be calculated as though all participants terminated employment as of the date of the survey, the minimum and maximum dates would be set equal to the survey date. Various combinations of quit age and date ranges can be selected to reflect any desired restrictions on the time of employment termination. This program option allows greater flexibility in data analysis since separate input data files are not needed to analyze differences in quit dates.

The Social Security variables define the level of the wagebase and benefit levels as of the date of the survey, as well as the real annual rates of growth in these amounts over time. The Social Security wagebase variable is defined as the maximum wage and salary income subject to

Social Security taxation. The benefit amount is defined as the Social Security benefit payable to a single person, without dependents, retiring at age 65, whose work and income history warrant the "average" benefit level. The other types of Social Security benefits are then estimated as proportions of this base amount, after being adjusted for growth over time.

Several sets of adjustments are used to account for the different types of Social Security benefits included in pension plan provisions. First, the base amount is adjusted for starting benefits at age 62 rather than at 65, what the pension coding instrument refers to as the reduced versus unreduced amount. The second adjustment is for the presence or absence of a spouse, what the coding instrument terms the primary versus joint amount. Other proportional adjustments are used to estimate disability benefits and surviving spouse benefits. In addition, all benefit amounts are adjusted to reflect differences in work and income histories, giving separate estimates for the minimum, the middle, and the maximum benefit amounts. Note that for all types of Social Security benefits, the program only takes account of the presence or absence of one dependent, namely the spouse if married.

The default setting for the differential used for Social Security benefits beginning at age 62 is 80% of the age 65 amount, or a reduction of 6.67% per year. For married participants, the benefit was increased by 50%. The maximum benefit levels are set 50% higher than the base amount, and the minimum 50% below the base. Of course, all of these proportions can be reset to any desired level, including 1.0, which would neutralize the distinctions. With any given set of proportions, the dollar values of the benefit amounts are determined by the initial dollar value and growth rate set in the INPARM file.

The INPARM also includes a setting for the dollar levels used for determining whether the minimum, midpoint, or maximum benefit amounts are used. The values chosen for the cutoffs increase over time, according to to the value chosen for the real growth rate in the Social Security wagebase and the nominal inflation rate. The adjusted cutoffs are compared with the growth adjusted wages to determine in which of the three groups the respondent is included. This option can be neutralized by setting the minimum cutoff value to zero, and the maximum to a number

larger than the highest QWAGE. This would have the effect of assigning everyone the "middle" level benefit.

The program estimates the dollar value of Social Security benefits by using the base benefit amount and real growth rates given in the input parameter file, adjusted by a proportional factor. The program selects the appropriate proportional factor, whose value is specified in the INPARM file, based on the pension plan provisions that govern the specific benefit formula being used, as well as the retirement age, income history, and marital status of the participants.

The input parameter BENMAX controls the maximum benefit that the program will allow before truncation. The maximum is entered in two steps, the first applies to the maximum amount that can be received at age 55, the second applies to the maximum annual pension benefit that can be received at age 65. If the latter maximum exceeds the former, the maximum is interpolated for retirement ages between 55 and 65. Of course, both could be set to a very large value to avoid the use of this truncation parameter.

For defined contribution pension plans that base contributions on measures of profitability, the input parameter PROFCON can be used to set the average rate of contributions over the time period used for pension calculations. Since some pension plans incorporate more than one formula that is based on contributions linked to profits, three separate values can be assigned. Pension plans with only one profit based contribution formula would use the variable PROFCON1. If a second profit based formula was included in the same plan, PROFCON2 would also be used. Similarly, PROFCON3 would only be used for the third profit based formula included in any single pension plan. Each of the variables is defined in terms of the proportion of annual wages that these contributions represent. In addition, many defined contribution plans also provide for the allocation of forfeitures to current plan participants. The amount of contributions from forfeitures can be set in the INPARM file. As with the profit link contributions, three separate rates can be set, each defined in terms of the proportion of annual wages the contributions represent.

A few defined contribution plans provided insufficient information to determine the basis and amount of employer contributions. For these plans, the variable NACON can be used to set the

rate of contributions expressed as a proportion of annual wages. A few other plans had mandatory contributions, but failed to specify the required contribution rate—usually because the mandatory contribution had been discontinued at some distant point in the past. For these plans, the unascertained mandatory contribution rate can be set, using the second value on the NACON line, defined as a proportion of annual wages. For plans did not specify the permissible range of voluntary contributions, the minimum and maximum rates of voluntary contributions are determined by the NAVOLrng variable in the INPARM file. The minimums and maximums are expressed as proportions of annual wages. Lastly, a few plans did not specify the rate at which they would match mandatory or voluntary contributions. The matching rates can be set via the NAMATCH variables in the INPARM file. The rates are given in percentage terms, with dollar-for-dollar matching expressed as 100.

The INPARM variable ANNTYPE selects the type of pension annuity the program calculates. When ANNTYPE is set to 1, all pension benefits are calculated as single life annuities; when set equal to 2, all benefits are calculated as two-life annuities; and when set equal to 3, benefit annuities are calculated as either single or joint life annuities depending on whether the participant is married—as given in the INDATA file. If joint life annuities are selected, the variable CONTPMT gives the proportion of the initial annuity payment continued after the first death. All annuity calculations are based on yearly survival probabilities, which are determined by the sex and birth date of the participant and spouse given in the INDATA files.

To allow for changes in pension plan provisions since these data were collected, two additional options are provided. The variable VESTYPE in the INPARM file controls the vesting formulas. When this variable is set equal to 0 the program uses the original plan provisions; when set equal to 1, the program substitutes a "cliff" formula, defined by the VESTYRS variable. The VESTYRS variable gives the number of years until full vesting occurs. The second option allows users to determine whether the plan specified mandatory retirement age is used. When the variable MRAGE is set equal to 1 the plan provisions are used, and when set equal to 0, no mandatory retirement age is used in the calculations.

The final parameter in the INPARM file selects the type of pension benefit to calculate. In most instances the PENTYPE variable should be set equal to 0, allowing the plan provisions to determine the type of retirement benefit based on age, years of service, and so forth. For the special cases of disability and survivor benefits, this parameter allows estimation of benefits under those circumstances. When the PENTYPE variable is set equal to 1, the program calculates disability retirement benefits, using the quit date as the date of disablement. When the PENTYPE variable is set equal to 2, the program calculates survivor benefits, using the quit date as the date of the death of the participant. The PENTYPE option does not affect RUNTYPE 3.2, since this RUNTYPE is designed to generate estimates of all pension types.

It may well be true that no one set of input estimation parameters will ever be adequate for all participants under all retirement circumstance. To accommodate this, the program will estimate only those cases included in the INDATA file for each replicate, so that subgroups of participants can be identified and matched to different sets of input parameters. The multiple subgroups can be defined in the same INDATA file through the use of the $continue with *dummy* command. And with the use of the "*" option, all the output can be written to a single file. In this fashion, the program estimates can be tailored to the requirements of almost all situations, even the extreme of defining the input parameters on a case by case basis.

Pension Plans

When analysis is based on simulated data cases, the program will generate estimates under every pension plans listed in the INPLAN file. The default file SPFN:INPLAN includes all of the pension plans included in the sample. Any subset of plans can be included in this file. Using both the INDATA and INPLAN files allows the user to limit estimates to just one participant under one pension plan, to as many simulated participant histories as desired under all of the plans included in the study.

The IDs included in the file INPLAN are the CODEID and the SEQ#. Each line in the file must begin with the number "9" in column 1 and any unique four digit index starting in column 2. Separated by blanks, the CODEID and the SEQ# follow. All defined benefit pension plans have

CODEID numbers under 3000, defined contribution plans have IDs numbered in the 3000s, and plans that combine both features have IDs starting at 5000.

Valid Ranges

The program checks all input data and parameter settings for valid ranges. The check for valid ranges will only detect gross errors in the input data. The program allows for a wide range of data, including levels that would be thought by many to be extreme values. This check in no way insures the reasonableness or accuracy of the input data and parameters. For example, the valid date ranges often cover 100 years or more, and only values outside this extended period are flagged by the program as invalid. When the program encounters an invalid data value or parameter setting, that value is automatically recoded to a default value that lies within the valid range. When this occurs, the program notes this problem in the OUTERR file. The valid ranges for the the parameter settings are given in Table 3, and the valid ranges for the input participant data are given in Table 4. Both tables include the recoded values the program automatically assigns to input data falling outside of the valid range.

OUTPUT FILES

Data Files

OUTDATA represents the name of the file where the program will write output in free format, with one participant/plan combination per line. As with the input data file, each line represents the estimated pension entitlements from a single plan. The total household pension entitlement is thus the sum of entitlements due under all plans in which the respondent or spouse participates. And for estimates of household pension wealth holding, estimates for the respondent and the spouse must also be combined. Two special utility programs, described below, are provided to convert these records. One program aggregates participant/plan data into one record per person, the other program aggregates the participant/plan data into household records.

The output data file is always emptied when the pension program is invoked. When using

Table 3

Valid Ranges for Input Parameter Settings

Variables	Valid Range		Outside Range Program Recode Value
	Minimum	Maximum	
SURVEYD	1980.00	2000.00	1983.50
TODAYD	1980.00	2000.00	1983.50
CPIrate	−25.00	25.00	0.00
INTrate	0.00	20.00	1.00
INCrate	−20.00	20.00	0.00
SSWB	1.00	60000.00	1.00
SSWBrate	−20.00	20.00	0.00
SS	1.00	30000.00	1.00
SSrate	−20.00	20.00	0.00
QAGE MIN	16.00	80.00	16.00
QAGE MAX	16.00	80.00	80.00
QDATE MIN	1925.00	2075.00	1925.00
QDATE MAX	1925.00	2075.00	2075.00
SSPUrto MIN	0.00	5.00	1.00
SSPUrto MID	0.00	5.00	1.00
SSPUrto MAX	0.00	5.00	1.00
SSJUrto MIN	0.00	5.00	1.00
SSJUrto MID	0.00	5.00	1.00
SSJUrto MAX	0.00	5.00	1.00
SSPRrto MIN	0.00	5.00	1.00
SSPRrto MID	0.00	5.00	1.00
SSPRrto MAX	0.00	5.00	1.00
SSJRrto MIN	0.00	5.00	1.00
SSJRrto MID	0.00	5.00	1.00
SSJRrto MAX	0.00	5.00	1.00
SSPDrto MIN	0.00	5.00	1.00
SSPDrto MID	0.00	5.00	1.00
SSPDrto MAX	0.00	5.00	1.00
SSJDrto MIN	0.00	5.00	1.00
SSJDrto MID	0.00	5.00	1.00
SSJDrto MAX	0.00	5.00	1.00
SSSSrto MIN	0.00	5.00	1.00
SSSSrto MID	0.00	5.00	1.00
SSSSrto MAX	0.00	500000.00	500000.00
BENMAX 55	0.00	500000.00	500000.00
BENMAX 65	0.00	500000.00	0.00
SSCUT MIN	0.00	500000.00	500000.00
SSCUT MAX	0.00	30.00	5.00
PROFVOL1	0.00	10.00	0.50
FORTVOL1	0.00	30.00	5.00
PROFVOL2	0.00	10.00	0.50
FORTVOL2	0.00	30.00	5.00
PROFVOL3			

Table 3 (continued)

Variables	Valid Range		Outside Range Program Recode Value
	Minimum	Maximum	
FORTVOL3	0.00	10.00	0.50
NACON	0.00	30.00	5.00
NAMAN	0.00	30.00	5.00
NAVOLrng MIN	0.00	30.00	0.00
NAVOLrng MAX	0.00	30.00	0.00
NAMATCH MAN	0.00	200.00	100.00
NAMATCH VOL	0.00	200.00	100.00
ANNTYPE	1.00	3.00	3.00
CONTPMT	0.00	100.00	50.00
VESTYPE	0.00	1.00	0.00
VESTYRS	0.00	25.00	5.00
MRAGE	0.00	1.00	0.00
PENTYPE	0.00	2.00	0.00

Table 4

Valid Ranges for Participant Data Variables

Variables	Valid Range		Outside Range Program Recode Value
	Minimum	Maximum	
SPOUSEBD	0.00	2000.00	BIRTHD
SEX	1.00	2.00	1.00
BIRTHD	1880.00	1970.00	1950.00
HIRED	BIRTHD+16	BIRTHD+65	BIRTHD+20
QUITD	HIRED	BIRTHD+80	BIRTHD+65
WKHRS	1.00	2080.00	2080.00
WAGE	1.00	1000000.00	1.00
WAGEG	−20.00	20.00	0.00
CRVOL	0.00	30.00	0.00

replicates, one output data file for each replicate is usually desired. If the output data from the second or succeeding replicates should be appended to the end of the first data output file specified,

use the symbol "*" instead of the file name.

The type and order of variables that are written into the OUTDATA file are defined by the RUNTYPE specified. For analyses based on actual participant data with given quit dates (RUNTYPE 1.2), the output variables are shown in Table 5. Note these output variables represent one participant/plan combination, and that the combined person or household record will contain up to eight repetitions of this output variable set. The output variables from analyses of simulated data (RUNTYPE 4.2) are given in Table 6. Since each run allows up to five simulated data cases, this basic set is repeated up to five times on each data output line. When the program calculates the pension entitlements due at all possible future quit dates (RUNTYPE 3.2), the output variables are as shown in Table 7. Note that, unlike the above output data sets written one case per line, this data set is written in blocks, giving all the data for a particular case in a series of contiguous lines. When printed, the output will occupy one page per case. The output file from this RUNTYPE is very large, especially when generated for the full sample of cases.

All the data entries in the OUTDATA file are numeric. While most of the output variables are coded in natural units (dollars and years), the variable that represents the type of pension for defined benefit plans is coded 1 for normal retirement, 2 for late retirement, 3 for early retirement, 4 for vested deferred, 5 for disability, and 6 for survivor benefits. For defined contribution plans, the variable is coded 7. For those pension plans that had a combination of defined benefit and defined contribution provisions, the code 11 was used for normal retirement, 12 for late retirement, 13 for late retirement, 14 for vested deferred, 15 for disability, 16 for survivor benefits, and 19 if only the contribution formulas were used. If the participant did not qualify for any type of retirement benefit, this variable was set equal to 9. The program assigns the missing data code of -1 to any plan that generates errors in calculations when combined with the input data and assumptions.

OUTTAB represents the name of the output file to which the program will display a summary of the estimation parameters that were used in that particular replication, and text headings for the output data file. This file is suitable for obtaining a printed copy of the estimated

pension benefits for each case. Even if this type of printed output is not needed—if the OUTDATA file is read directly by an analysis program—the file name must still be specified, even if assigned to "dummy"

OUTERR represents the file name to write the program produced summary of any suspect dollar figure calculations in any portion of the estimation procedures. This file name must be present. The file contains short messages about potential problems, indicating the case IDs and a list of the data values. That a case has been flagged for inclusion in the problem file does not necessarily mean that the calculated value for the participant is incorrect, but that in the calculation process there were some suspect intermediate results.

Table 5

Output Data File for Entitlement Estimates
Based on Actual Participant Data with Given Quit Date
(RUNTYPE 1.2)

Order	Name	Definition
1	HHIDX	SCF Household ID
2	CODEID	Pension Plan ID
3	SEQ#	Pension Plan/Provider ID
4	SEX	Sex of participant
5	SPOUSEBD	Birth date of spouse, if married
6	BIRTHD	Birth date of participant
7	HIRED	Date participant began employment
8	QUITD	Date participant terminates employment
9	QWAGE	Annual wage or salary at quit date
10	QASY	Number of years covered at quit date
11	PENTYPE	Type of retirement benefit
12	PENAGE	Age benefit payments begin
13	%QW	Pension benefit as percent of final wages
14	%/YR	Proportion of final wages per service year
15	30YR	Proportion of final wages for 30 years
16	ANNUALA	Initial annual pension benefit amount
17	PRVALUE	Present value of entitlement
18	A65EQVL	Equivalent annual amount if start at 65

Table 6

Output Data File for Entitlement Estimates
Based on Simulated Participant Data
(RUNTYPE 4.2)

Order	Name	Definition
1	CODEID	Pension plan ID.
2	TYPE	Type of pension benefit
3	%QW	Benefit as percent of final wages
4	%/YR	Percent final wages per service year
5	ANNUALA	Initial annual benefit amount
6	PRVALUE	Present value of entitlement

Table 7

Output Data File for Entitlement Estimates
Based on Actual Participant Data Using All Potential Quit Dates
(RUNTYPE 3.2)

Order	Name	Definition
1	QUITD	Date employment terminates
2	QAGE	Age at quit date
3	QASY	Total service years at quit date
4	QWAGE	Wage or salary at quit date
5	SSB	Social Security benefit amount
6	PENBEN	Pension benefit amount
7	PENTYPE	Type of retirement benefit
8	VESTED	Deferred vested benefit
9	BENAGE	Age when benefits begin
10	VEST65	Vested benefit if begin at 65
11	DABLEB	Disability pension benefit
12	DTYPE	Type of disability pension
13	NRT	Normal retirement benefit
14	LRT	Late retirement benefit
15	ERT	Early retirement benefit
16	VDT	Deferred vested retirement
17	DRT	Disability retirement benefit
18	DST	Survivor retirement benefit
19	CNT	Benefit based on contributions

DATA BASE UTILITY PROGRAMS

The pension program produces one data record for each respondent or spouse for each pension plan in which they participate. Most analysts will want to aggregate these records to either the level of persons or households depending on the goals of the analysis. Analyses focused on the worker as the unit of analysis, for example, will want to aggregate the several potential records for each individual. When the household is the unit of analysis, the pension entitlements for both the respondent and spouse need to be aggregated. Both of the special conversion programs were designed to be used with RUNTYPE 1.2. When using simulated participant data (RUNTYPE 4.2), there is no need to covert or aggregate records before analysis.

Conversion to Person Records

The public use file that contains the program to create person records is SPFN:PP.PEREC. This Pascal program converts the pension program output to one line per individual, giving the pension entitlement amounts for each plan in which the individual participates. The output variables are in the same form as from the pension program, with the information repeated for up to four potential pensions. The first variable on each line is the ID variable (HHIDX). This variable is then followed by the same variables listed in Table 5 above—note that the ID variable is repeated at the start of each set. The program is invoked by the following command:

$RUN SPFN:PP.PEREC SCARDS=OUTDATA SPRINT=PERDATA

where the file OUTDATA is the output from the pension program, and PERDATA is the newly created data file.

Conversion to Household Records

The public use file that contains the program to create household records is SPFN:PP.HHREC. This program converts the pension program output to one line per household, giving the pension entitlement amounts for both the respondent and spouse. Up to eight potential pensions for each household are recorded in groups of four, with the first group applying to the respondent, and the second to the spouse. Each line of the new file starts with the household ID

(HHID, without the suffix), and is followed by up to eight sets of pension entitlement amounts. Whenever a respondent or spouse did not have a pension or was not covered by additional plans, the variables are padded with zeros. The order of the variables in each of the eight sets is the same as given in Table 5. The program is invoked by the following command:

$RUN SPFN:PP.HHREC SCARDS=OUTDATA SPRINT=HHDATA

where the file OUTDATA is the output from the pension program, and HHDATA is the newly created data file. Both types of data records can be generated from the same pension output file by simply using both record conversion programs on the same file.

OSIRIS OUTPUT DATA FILE CREATION

Dictionary Files

Several public files are made available on SPFN to facilitate analysis, and serve as an example of how to use the output from the pension program as input to widely available statistical software. With the use of the provided Osiris dictionary, the free format output from the pension program can be read by Osiris, and the desired statistical analysis performed. Even if users prefer other statistical packages, such as SPSS, SAS, or MIDAS, these software programs have commands that are specifically tailored to read Osiris datasets. Thus only in situations where Osiris datasets cannot be read by the user's software program, would the user prefer the free format version. Particularly helpful are the included variable names, field widths, and missing data codes that the Osiris dictionary provides.

Since the output data set differs depending on the RUNTYPE and record structure, three different Osiris dictionaries are provided. The dictionaries are in the files: SPFN:OSDICT1 for actual participant data aggregated to person records; SPFN:OSDICT2 for actual participant data aggregated to the household level; and SPFN:OSDICT4 for simulated participant data. Table 8 gives the dictionary records for data aggregated to person records. The dictionaries include the width of the variable, the number of decimal places, the type of data (character), and the first and second missing data codes. Table 9 shows the dictionary for the household records data base. Note that this dictionary includes 153 variables—eight sets of pension output variables, with each

set having 19 variables, plus the added ID. The dictionary records for the simulated data set are shown in Table 10.

Source Files for Modification

Since the provided dictionaries make no allowance for any pass through variables, the command files which created these dictionaries are also available to users on a read only basis. These files can be copied, and edited to reflect the addition of the user selected pass through variables. The file names are similar to the above with an added suffix: SPFN:OSDICT1.SOU, SPFN:OSDICT2.SOU, and SPFN:OSDICT4.SOU.

EXAMPLES OF COMMAND FILES

The complete set of appropriate pension program commands, record conversion command, and Osiris commands are available in public use files for person records, household records, and simulate data. The commands necessary to estimate pension entitlements when using actual participant data aggregated to person records are:

```
$RUN SPFN:PP4.0
1.2 SPFN:INDATA.SCF SPFN:INPARM -OUTDATA1 -OUTTAB1 -OUTERR1
$ENDFILE
$RUN SPFN:PP.PEREC SCARDS=-OUTDATA1 SPRINT=-OUTPEDA1
$RUN ISR:OSIRIS.IV SPRINT=-OUTOS1
&FREE DICTIN=SPFN:OSDICT1 DATAIN=-OUTPEDA1 DICTOUT=-DI1 DATAOUT=-
DA1
OSIRIS BUILD FOR PARTICIPANT DATA AGGREGATED TO PERSON RECORDS
VARS=ALL
&END
&STOP
```

The appropriate commands to use when estimates are based on actual participant data aggregated into household records are available in the public use file SPFN:RUNPP.HH. These commands are:

```
$RUN SPFN:PP4.0
1.2 SPFN:INDATA.SCF SPFN:INPARM -OUTDATA2 -OUTTAB2 -OUTERR2
$ENDFILE
$RUN SPFN:PP.HHREC SCARDS=-OUTDATA2 SPRINT=-OUTHHDA2
$RUN ISR:OSIRIS.IV SPRINT=-OUTOS2
&FREE DICTIN=SPFN:OSDICT2 DATAIN=-OUTHHDA2 DICTOUT=-DI2 DATAOUT=-
DA2
OSIRIS BUILD FOR ACTUAL PARTICIPANT DATA ON HOUSEHOLD BASIS
```

```
VARS=ALL
&END
&STOP
```

The commands needed to base estimates on simulated data and build an Osiris data base are included in the file SPFN:RUNPP.SIM. The required commands are as follows:

```
$RUN SPFN:PP4.0
4.2 INDATA.SIM(100) SPFN:INPARM -OUTDA4 -OUTTAB4 -OUTERR4
$ENDFILE
$RUN ISR:OSIRIS.IV SPRINT=-OUTOS4
&FREE DICTIN=SPFN:OSDICT4 DATAIN=-OUTDA4 DICTOUT=-DI4 DATAOUT=-DA4
OSIRIS FILE BUILD FOR SIMULATED DATA
VARS=ALL
&END
&STOP
```

PASCAL PROGRAM SOURCE FILES

PP4.0 Source Code

The Pascal source code that was used to create the program file PP4.0 is provided on a read-only basis in several files on SPFN. Users may copy this source code, and modify their own versions to meet any special analytic needs. Such an undertaking requires an extensive knowledge of the underlying data bases and procedures. The PP4.0 program is compiled in three phases, using the following commands:

```
$RUN *PASCALJB SCARDS=SPFN:CALCULATE 0=PP4.0(LAST+1)
1=*PASCALJBINCLUDE PAR=NOCHK,OPT
$RUN *PASCALJB SCARDS=SPFN:EVALUATE 0=PP4.0(LAST+1) PAR=NOCHK,OPT
$RUN *PASCALJB SCARDS=SPFN:OUTSIDE 0=PP4.0(LAST+1) PAR=NOCHK,OPT
```

The above programs need access to three input files: SPFN:CASES, SPFN:CASES2, and SPFN:PROCEDURES. The file SPFN:CASES contains a list of all valid case IDs. SPFN:CASES2 lists the case IDs for those plans which need to be integrated with another plan. Other than for the cases listed in this file, the total entitlement due from coverage under two or more pension plans is simply the sum of the individual plan entitlements. Only those plans whose benefits depended on the benefits received under another plan are included in the CASES2 file. The file SPFN:PROCEDURES contains all the benefit formulas and plan provisions in the form of Pascal procedures. Each procedure represents the plan provisions and benefit formulas of one

pension plan as a separate system of equations.

Procedures Source Code

The Pascal source code for the program that generated the input procedures file is also available to users for modification. As with the above program, the procedures program is compiled in three steps, using the following commands:

```
$RUN *PASCALJB SCARDS=SPFN:GENERATE   0=OBJ(LAST+1)
$RUN *PASCALJB SCARDS=SPFN:EMIT       0=OBJ(LAST+1)
$RUN *PASCALJB SCARDS=SPFN:MISCELLANY 0=OBJ(LAST+1)
```

After the above Pascal object file has been compiled, the PROCEDURES file is generated by the following commands:

```
$RUN OBJ SERCOM=-ERR1 SPRINT=FORMULAS SPUNCH=*DUMMY*
         0=PEN.DICT 1=PEN.DATA 2=COUNT 3=PROBLEMS
$RUN OBJ SERCOM=-ERR2 SPRINT=DATALIST
         0=PEN.DICT 1=PEN.DATA PAR=LIST
$RUN OBJ SERCOM=-ERR3 SPRINT=*DUMMY* SPUNCH=PROCEDURES
         0=PEN.DICT 1=PEN.DATA 2=*DUMMY* 3=*DUMMY*
```

The first command above causes the program to output a list of all formulas for each pension plan in a format suitable for printing. The second command instructs the program to output a list of all variables that the program accessed from the pension data set. The third command causes the program to output all benefit formulas in the form of Pascal procedures, one for each pension plan. Each procedure is named by its pension plan ID. The programs require two input files: PEN.DICT and PEN.DATA, the Osiris dictionary and data files for the pension survey. The output files COUNT and PROBLEMS give a summary of the number of formulas and cases that have been processed, and a listing of any potential problems encountered.

Utilities Source Code

The Pascal source code used to compile the programs that convert participant/plan records to either person or household records is available on SPFN on a read-only basis in the files:

```
SPFN:PP.PEREC.PAS
SPFN:PP.HHREC.PAS
```

Each of these programs is compiled separately, as follows:

```
$RUN *PASCALJB SCARDS=SPFN:PP.PERREC.PAS 0=PP.PEREC
$RUN *PASCALJB SCARDS=SPFN:PP.HHREC.PAS  0=PP.HHREC
```

ONLINE DOCUMENTATION

The contents of the document you are reading, as well as a shorter document which gives the name and a brief description of all public files, are available to users online. Simply edit or copy the files SPFN:PP4.DOC or SPFN:PP.FILES.

AVAILABILITY TO NON-MTS USERS

All the pension program files available on SPFN can be transmitted via telephone lines to other computer sites, or copied to magnetic tape and mailed to other computer facilities. Requests for magnetic tapes should be directed to the Survey Research Center, Pension Study Program Archive. The MTS computer system operates on an IBM 3090–400 mainframe. The version of Pascal used is a product of Plug Compatible Software, Inc. This version is very similar to the IBM version, Pascal/VS, but has a capacity for larger programs. When the program is compiled to include all pension plan formulas, it requires at least 3.5 megabytes of main memory to execute. The size of the object program can be controlled by limiting the number of pension plans included in the compiled version, and using multiple versions of the program. The minimum size of the compiled program—including only one pension plan—would be approximately 250K.

Table 8

OSIRIS Dictionary for Actual Participant Data
Formatted as Person Records

Var#	Name	Width	Ndec	Type	MD1	MD2
V1	HHIDX	5	0	C	0	
V101	R1:HHIDX	5	0	C	0	
V102	R1:CODEID	4	0	C	0	−1
V103	R1:SEQ#	4	0	C	0	−1
V104	R1:SEX	1	0	C	0	
V105	R1:SPOUSEBD	7	2	C	−1	
V106	R1:BIRTHD	7	2	C	0	−1
V107	R1:HIRED	7	2	C	0	−1
V108	R1:QUITD	7	2	C	0	−1
V109	R1:QWAGE	11	2	C	0	−1
V110	R1:QAGE	6	2	C	0	−1
V111	R1:QASY	6	2	C	0	−1
V112	R1:PTYPE	2	0	C	0	−1
V113	R1:PAGE	6	2	C	0	−1
V114	R1:%QW	6	2	C	0	−1
V115	R1:%/YR	6	2	C	0	−1
V116	R1:30YR	6	2	C	0	−1
V117	R1:PENSION	10	2	C	0	−1
V118	R1:PRVALUE	11	2	C	0	−1
V119	R1:A65EQIV	10	2	C	0	−1
V201	R2:HHIDX	5	0	C	0	
V202	R2:CODEID	4	0	C	0	−1
V203	R2:SEQ#	4	0	C	0	−1
V204	R2:SEX	1	0	C	0	
V205	R2:SPOUSEBD	7	2	C	−1	
V206	R2:BIRTHD	7	2	C	0	−1
V207	R2:HIRED	7	2	C	0	−1
V208	R2:QUITD	7	2	C	0	−1
V209	R2:QWAGE	11	2	C	0	−1
V210	R2:QAGE	6	2	C	0	−1
V211	R2:QASY	6	2	C	0	−1
V212	R2:PTYPE	2	0	C	0	−1
V213	R2:PAGE	6	2	C	0	−1
V214	R2:%QW	6	2	C	0	−1
V215	R2:%/YR	6	2	C	0	−1
V216	R2:30YR	6	2	C	0	−1
V217	R2:PENSION	10	2	C	0	−1
V218	R2:PRVALUE	11	2	C	0	−1
V219	R2:A65EQIV	10	2	C	0	−1
V301	R3:HHIDX	5	0	C	0	
V302	R3:CODEID	4	0	C	0	−1
V303	R3:SEQ#	4	0	C	0	−1
V304	R3:SEX	1	0	C	0	
V305	R3:SPOUSEBD	7	2	C	−1	

Table 8 (continued)

Var#	Name	Width	Ndec	Type	MD1	MD2
V306	R3:BIRTHD	7	2	C	0	-1
V307	R3:HIRED	7	2	C	0	-1
V308	R3:QUITD	7	2	C	0	-1
V309	R3:QWAGE	11	2	C	0	-1
V310	R3:QAGE	6	2	C	0	-1
V311	R3:QASY	6	2	C	0	-1
V312	R3:PTYPE	2	0	C	0	-1
V313	R3:PAGE	6	2	C	0	-1
V314	R3:%QW	6	2	C	0	-1
V315	R3:%/YR	6	2	C	0	-1
V316	R3:30YR	6	2	C	0	-1
V317	R3:PENSION	10	2	C	0	-1
V318	R3:PRVALUE	11	2	C	0	-1
V319	R3:A65EQIV	10	2	C	0	-1
V401	R4:HHIDX	5	0	C	0	
V402	R4:CODEID	4	0	C	0	-1
V403	R4:SEQ#	4	0	C	0	-1
V404	R4:SEX	1	0	C	0	
V405	R4:SPOUSEBD	7	2	C	-1	
V406	R4:BIRTHD	7	2	C	0	-1
V407	R4:HIRED	7	2	C	0	-1
V408	R4:QUITD	7	2	C	0	-1
V409	R4:QWAGE	11	2	C	0	-1
V410	R4:QAGE	6	2	C	0	-1
V411	R4:QASY	6	2	C	0	-1
V412	R4:PTYPE	2	0	C	0	-1
V413	R4:PAGE	6	2	C	0	-1
V414	R4:%QW	6	2	C	0	-1
V415	R4:%/YR	6	2	C	0	-1
V416	R4:30YR	6	2	C	0	-1
V417	R4:PENSION	10	2	C	0	-1
V418	R4:PRVALUE	11	2	C	0	-1
V419	R4:A65EQIV	10	2	C	0	-1

Table 9

OSIRIS Dictionary for Actual Participant Data
Formatted as Household Records

Var#	Name	Width	Ndec	Type	MD1	MD2
V1	HHID	4	0	C	0	
V101	R1:HHIDX	5	0	C	0	
V102	R1:CODEID	4	0	C	0	-1
V103	R1:SEQ#	4	0	C	0	-1
V104	R1:SEX	1	0	C	0	
V105	R1:SPOUSEBD	7	2	C	-1	
V106	R1:BIRTHD	7	2	C	0	-1
V107	R1:HIRED	7	2	C	0	-1
V108	R1:QUITD	7	2	C	0	-1
V109	R1:QWAGE	11	2	C	0	-1
V110	R1:QAGE	6	2	C	0	-1
V111	R1:QASY	6	2	C	0	-1
V112	R1:PTYPE	2	0	C	0	-1
V113	R1:PAGE	6	2	C	0	-1
V114	R1:%QW	6	2	C	0	-1
V115	R1:%/YR	6	2	C	0	-1
V116	R1:30YR	6	2	C	0	-1
V117	R1:PENSION	10	2	C	0	-1
V118	R1:PRVALUE	11	2	C	0	-1
V119	R1:A65EQIV	10	2	C	0	-1
V201	R2:HHIDX	5	0	C	0	
V202	R2:CODEID	4	0	C	0	-1
V203	R2:SEQ#	4	0	C	0	-1
V204	R2:SEX	1	0	C	0	
V205	R2:SPOUSEBD	7	2	C	-1	
V206	R2:BIRTHD	7	2	C	0	-1
V207	R2:HIRED	7	2	C	0	-1
V208	R2:QUITD	7	2	C	0	-1
V209	R2:QWAGE	11	2	C	0	-1
V210	R2:QAGE	6	2	C	0	-1
V211	R2:QASY	6	2	C	0	-1
V212	R2:PTYPE	2	0	C	0	-1
V213	R2:PAGE	6	2	C	0	-1
V214	R2:%QW	6	2	C	0	-1
V215	R2:%/YR	6	2	C	0	-1
V216	R2:30YR	6	2	C	0	-1
V217	R2:PENSION	10	2	C	0	-1
V218	R2:PRVALUE	11	2	C	0	-1
V219	R2:A65EQIV	10	2	C	0	-1
V301	R3:HHIDX	5	0	C	0	
V302	R3:CODEID	4	0	C	0	-1
V303	R3:SEQ#	4	0	C	0	-1
V304	R3:SEX	1	0	C	0	
V305	R3:SPOUSEBD	7	2	C	-1	

Table 9 (continued)

Var#	Name	Width	Ndec	Type	MD1	MD2
V306	R3:BIRTHD	7	2	C	0	−1
V307	R3:HIRED	7	2	C	0	−1
V308	R3:QUITD	7	2	C	0	−1
V309	R3:QWAGE	11	2	C	0	−1
V310	R3:QAGE	6	2	C	0	−1
V311	R3:QASY	6	2	C	0	−1
V312	R3:PTYPE	2	0	C	0	−1
V313	R3:PAGE	6	2	C	0	−1
V314	R3:%QW	6	2	C	0	−1
V315	R3:%/YR	6	2	C	0	−1
V316	R3:30YR	6	2	C	0	−1
V317	R3:PENSION	10	2	C	0	−1
V318	R3:PRVALUE	11	2	C	0	−1
V319	R3:A65EQIV	10	2	C	0	−1
V401	R4:HHIDX	5	0	C	0	
V402	R4:CODEID	4	0	C	0	−1
V403	R4:SEQ#	4	0	C	0	−1
V404	R4:SEX	1	0	C	0	
V405	R4:SPOUSEBD	7	2	C	−1	
V406	R4:BIRTHD	7	2	C	0	−1
V407	R4:HIRED	7	2	C	0	−1
V408	R4:QUITD	7	2	C	0	−1
V409	R4:QWAGE	11	2	C	0	−1
V410	R4:QAGE	6	2	C	0	−1
V411	R4:QASY	6	2	C	0	−1
V412	R4:PTYPE	2	0	C	0	−1
V413	R4:PAGE	6	2	C	0	−1
V414	R4:%QW	6	2	C	0	−1
V415	R4:%/YR	6	2	C	0	−1
V416	R4:30YR	6	2	C	0	−1
V417	R4:PENSION	10	2	C	0	−1
V418	R4:PRVALUE	11	2	C	0	−1
V419	R4:A65EQIV	10	2	C	0	−1
V501	S1:HHIDX	5	0	C	0	
V502	S1:CODEID	4	0	C	0	−1
V503	S1:SEQ#	4	0	C	0	−1
V504	S1:SEX	1	0	C	0	
V505	S1:SPOUSEBD	7	2	C	−1	
V506	S1:BIRTHD	7	2	C	0	−1
V507	S1:HIRED	7	2	C	0	−1
V508	S1:QUITD	7	2	C	0	−1
V509	S1:QWAGE	11	2	C	0	−1
V510	S1:QAGE	6	2	C	0	−1
V511	S1:QASY	6	2	C	0	−1
V512	S1:PTYPE	2	0	C	0	−1
V513	S1:PAGE	6	2	C	0	−1
V514	S1:%QW	6	2	C	0	−1
V515	S1:%/YR	6	2	C	0	−1

Table 9 (continued)

Var#	Name	Width	Ndec	Type	MD1	MD2
V516	S1:30YR	6	2	C	0	-1
V517	S1:PENSION	10	2	C	0	-1
V518	S1:PRVALUE	11	2	C	0	-1
V519	S1:A65EQIV	10	2	C	0	-1
V601	S2:HHIDX	5	0	C	0	
V602	S2:CODEID	4	0	C	0	-1
V603	S2:SEQ#	4	0	C	0	-1
V604	S2:SEX	1	0	C	0	
V605	S2:SPOUSEBD	7	2	C	-1	
V606	S2:BIRTHD	7	2	C	0	-1
V607	S2:HIRED	7	2	C	0	-1
V608	S2:QUITD	7	2	C	0	-1
V609	S2:QWAGE	11	2	C	0	-1
V610	S2:QAGE	6	2	C	0	-1
V611	S2:QASY	6	2	C	0	-1
V612	S2:PTYPE	2	0	C	0	-1
V613	S2:PAGE	6	2	C	0	-1
V614	S2:%QW	6	2	C	0	-1
V615	S2:%/YR	6	2	C	0	-1
V616	S2:30YR	6	2	C	0	-1
V617	S2:PENSION	10	2	C	0	-1
V618	S2:PRVALUE	11	2	C	0	-1
V619	S2:A65EQIV	10	2	C	0	-1
V701	S3:HHIDX	5	0	C	0	
V702	S3:CODEID	4	0	C	0	-1
V703	S3:SEQ#	4	0	C	0	-1
V704	S3:SEX	1	0	C	0	
V705	S3:SPOUSEBD	7	2	C	-1	
V706	S3:BIRTHD	7	2	C	0	-1
V707	S3:HIRED	7	2	C	0	-1
V708	S3:QUITD	7	2	C	0	-1
V709	S3:QWAGE	11	2	C	0	-1
V710	S3:QAGE	6	2	C	0	-1
V711	S3:QASY	6	2	C	0	-1
V712	S3:PTYPE	2	0	C	0	-1
V713	S3:PAGE	6	2	C	0	-1
V714	S3:%QW	6	2	C	0	-1
V715	S3:%/YR	6	2	C	0	-1
V716	S3:30YR	6	2	C	0	-1
V717	S3:PENSION	10	2	C	0	-1
V718	S3:PRVALUE	11	2	C	0	-1
V719	S3:A65EQIV	10	2	C	0	-1
V801	S4:HHIDX	5	0	C	0	
V802	S4:CODEID	4	0	C	0	-1
V803	S4:SEQ#	4	0	C	0	-1
V804	S4:SEX	1	0	C	0	
V805	S4:SPOUSEBD	7	2	C	-1	
V806	S4:BIRTHD	7	2	C	0	-1

Table 9 (continued)

Var#	Name	Width	Ndec	Type	MD1	MD2
V807	S4:HIRED	7	2	C	0	−1
V808	S4:QUITD	7	2	C	0	−1
V809	S4:QWAGE	11	2	C	0	−1
V810	S4:QAGE	6	2	C	0	−1
V811	S4:QASY	6	2	C	0	−1
V812	S4:PTYPE	2	0	C	0	−1
V813	S4:PAGE	6	2	C	0	−1
V814	S4:%QW	6	2	C	0	−1
V815	S4:%/YR	6	2	C	0	−1
V816	S4:30YR	6	2	C	0	−1
V817	S4:PENSION	10	2	C	0	−1
V818	S4:PRVALUE	11	2	C	0	−1
V819	S4:A65EQIV	10	2	C	0	−1

Table 10

OSIRIS Dictionary for Simulated Data
Formatted as Pension Plan Records

Var#	Name	Width	Ndec	Type	MD1	MD2
V1	CODEID	4	0	C	0	-1
V101	1ST PTYPE	1	0	C	0	
V102	1ST %QW	6	2	C	0	-1
V103	1ST %/YR	6	2	C	0	-1
V104	1ST PEN AMT	10	2	C	0	-1
V105	1ST P.VALUE	11	2	C	0	-1
V201	2ND PTYPE	1	0	C	0	
V202	2ND %QW	6	2	C	0	-1
V203	2ND %/YR	6	2	C	0	-1
V204	2ND PEN AMT	10	2	C	0	-1
V205	2ND P.VALUE	11	2	C	0	-1
V301	3RD PTYPE	1	0	C	0	
V302	3RD %QW	6	2	C	0	-1
V303	3RD %/YR	6	2	C	0	-1
V304	3RD PEN AMT	10	2	C	0	-1
V305	3RD P.VALUE	11	2	C	0	-1
V401	4TH PTYPE	1	0	C	0	
V402	4TH %QW	6	2	C	0	-1
V403	4TH %/YR	6	2	C	0	-1
V404	4TH PEN AMT	10	2	C	0	-1
V405	4TH P.VALUE	11	2	C	0	-1
V501	5TH PTYPE	1	0	C	0	
V502	5TH %QW	6	2	C	0	-1
V503	5TH %/YR	6	2	C	0	-1
V504	5TH PEN AMT	10	2	C	0	-1
V505	5TH P.VALUE	11	2	C	0	-1