The Pension Calculator Program for the Health and Retirement Study

Users Guide

August 1998

Program by Richard Curtin

Modified for Use with PC's by
Jody Lamkin and Bob Peticolas

Guide Adapted by
Tom Steinmeier

I. Introduction.

This guide contains instructions to calculate pension benefits for the Health and Retirement Study. In 1992, the HRS fielded an initial survey to respondents born in the 1931-1941 range and their spouses. The survey included sections on current employment and previous employment. For each job, the survey asked respondents for the name and address of the employer. It subsequently contacted the employer to obtain summary plan descriptions, which are the booklets describing benefits that employers distribute to their employees. In cases where the survey was unable to obtain the plan descriptions from the employer, it attempted to find the plan description in files maintained by the U.S. Department of Labor. These descriptions contain the formula for calculating benefits for workers at various ages and with various years of service with the firm, and they stipulate the requirements for collecting benefits under any particular formula. The descriptions have been coded into a standardized format, and the results have been used to create the pension calculation program.

The pension calculator program is actually two programs, one to calculate defined benefit plans and the other to calculate defined contribution and combination plans. In order to use either program, users will need three parts. First is the program itself. For the defined benefit plans the program file is CALC6DB.EXE, and for the defined contribution plans it is CALC6DC.EXE. Second is an input file describing the general economic environment in which the calculations are to be done. This is called the "parameters" file. The final part is an input file describing the individuals for which the pensions are to be calculated, and identifiers for the pensions that are applicable for each individual. This file is the "participant" file. Examples of the parameters and participants files have been distributed with the program and will be discussed in detail in the following sections. Note in particular that the program does not require direct access to the survey itself. All the information required for the calculations is contained in the input files.

The program can operate in three modes. In Mode 1, each line of input causes the program to calculate the pension available for a particular individual as of a particular date, which may be different from one individual to the next. This is useful if the object is, for instance, to calculate the value of respondents' pensions at a particular date or at a particular age. Mode 2 asks how different pensions would apply to a single individual. This would be useful for assessing how generous the various pensions are to individuals hired early or late, at high or low earnings, or with other characteristics. Mode 3 is similar to Mode 1, except that each input line causes the program to calculate the pension that would be available at alternative retirement dates. Researchers wanting to explore the retirement incentives of pensions will find this mode useful.

Three files are written by the program. The first lists the results in tabular format that is relatively easy for the researcher to look at and read. However, since the tabular format is less convenient to process further using standard statistical programs, a second file contains the same information in strictly numeric format. The last file is an error file, which flags calculations with

possible errors.

To use the program, the user places the program file in the same directory with the input files. From the DOS prompt, the user starts the program with a single command line. The program indicates its progress as it does the calculations. The general requirements for the program are an IBM/PC compatible computer with an Intel 486DX or compatible CPU and about 15 megabytes of available system memory. The output files, which are written to disk, can be substantial, especially if the program is run in the Mode 3.

The next section will describe the input files. Section III discusses how to run the program once the input files have been modified to suit the user's needs, and the following section describes the output files in a little more detail. The final section touches on some issues which may be of interest to advanced users.

II. Preparing the Input Files.

The first thing to note about the input files is that the calculation program is expecting simple ASCII files. It is possible, and usually even easier, to use word processing or spreadsheet programs to create the files. When these files are saved, however, the researcher must be sure to use the appropriate option in the word processing or spreadsheet program to save them in ASCII (sometimes called TEXT) format.

Before discussing these input files, it will be useful to reiterate the three modes that were discussed in the introduction. To keep things simple, we will number these modes in order:

Mode 1: For each line in the participant file, the program calculates a single pension for a specific participant as of a specific date.

Mode 2: A (possibly large) group of pensions is calculated for a single participant (or for a small group of participants).

Mode 3: For each line in the participant file, the program calculates a single pension for a specific participant, but for a range of alternative retirement dates.

The calculations are not really any different in the three modes. It is possible, though more laborious, to do the calculations of Mode 2 or Mode 3 using only Mode 1. However, there are differences in the variables included in the output of the three modes.

The Parameters File.

The first input file is the file describing the economic environment, as well as settings that the program will apply to all individuals in the sample. This file is identical for all three modes. Table 1 reproduces the sample file (PARAM) that is included with the program. This file may be renamed, but if so, the command line to run the program must be changed accordingly.

This file contains parameter values which should be appropriate for most users. If the user wants to change any of these values, it can be done with a simple text editor or word processor. Just replace the supplied value of the file with the new value you wish to use. It is even possible to create several versions of the file to be used with different runs of the program. Again, please remember that the file must be saved in simple ASCII TEXT format. No changes are required to this file if the values are acceptable to the user.

This file contains many values which are largely self explanatory, but several will not be so. It is probably best to go through these in order.

Line 1. Date of survey observations. This is the date on which the pension descriptions were

Table 1
Sample Parameter File

	1992.00	Date of survey observations
	1992.00	Date for present value calculations
	4.00	Inflation rate (annual percentage rate)
	2.30	Real interest rate (must be greater than 0)
	1.00	Real salary and wage growth rate
	1.00	Real SS wagebase growth rate
	16 70	Quit age range (min, max)
1925.0	2075.0	Quit date range (min, max)
42000.00	112000.0	Maximum benefits (age55, age65)
5.00	0.50	Profit sharing and forfeiture contrib percentages (1)
5.00	0.50	Profit sharing and forfeiture contrib percentages (2)
5.00	0.50	Profit sharing and forfeiture contrib percentages (3)
	50.00	Discretionary matching percentage (1)
	50.00	Discretionary matching percentage (2)
	50.00	Discretionary matching percentage (3)
5.00	5.00	Other unspecified employer contrib % / Unspecified mandatory employee contrib %
0.00	25.00	Voluntary contribution percentage bounds (min, max)
50.00	50.00	Unspecified matching percentage for mandatory and voluntary contributions
1.00	00.00	Annuity type / % of payment continued
0.00	5.00	New vesting (1=on,0=off) / Yrs for vesting
	0.00	Plan maximum retirement (1=on/0=off)
	0.00	Pension type (0=Retirement, 1=Disability, 2=Survivor)
	1.00	First year annuity / present value output flag, 0=FYA, 1=PV; default=0
	0.00	Wage histories flag, 1=no, 2=yes
	1.00	Adjust dollar amounts for inflation flag, 0=no, 1=yes
1.00	50.00	COLA to benefits for no COLA plans flag, 0=no, 1=yes; % of CPI
	1.00	Mortality tables: 1=1936 birth year, 0=all birth years
	1.00	Ignore benefits paid before date of present value calculations, 0=no, 1=yes

collected. It is probably best never to change this value from the original value.

Line 2. Date for present value calculations. Most values in the output tables are current dollar amounts, but some amounts are present values. This gives the date to which the present values are discounted. This parameter is relevant only for Modes 1 and 2. Mode 3, which computes present values for a range of alternative retirement dates, discounts the benefits to the relevant retirement date.

Line 3. Inflation rate. This is an annual percentage amount. Since many of the other rates in this file are real rates, a change in the inflation rate changes all the nominal rates simultaneously while preserving the real rates.

Line 4. Real interest rate. In combination with the inflation rate, this provides the nominal interest rate that is used in discounting.

Line 5. Real salary and wage growth. In combination with the inflation rate, this determines the level of overall nominal salary and wage growth. This does not force the wages of all respondents to grow at this rate; values in the participants file allow the wages or salary of an individual respondent to grow at either a greater or lesser amount. These value of this rate in the supplied file, and also for the previous two rates, are those used in the intermediate Social Security projections.

Line 6. Real Social Security wage base growth rate. The Social Security wage base is the level of earnings that is subject to social security taxes in a given year. The wage base is currently indexed for the growth of average annual earnings for the population as a whole. This parameter specifies this growth rate. In general, this will be the same as the real salary and wage growth rate. However, if the sample is older and the growth rate of earnings has diminished, this parameter may plausibly exceed the real salary and wage growth rate by a small amount.

Line 7. Maximum and minimum quit age range. In Modes 1 and 2, the user specifies the date that each individual separates from the firm. That separation date in usually an important determinant of benefits. These parameters define the range of values which will be accepted; other values will be considered erroneous and will be reset by the program, as discussed later. In Mode 3, the maximum quit age places an upper limit on the range of retirement ages for which the pensions are computed.

Line 8. Minimum and maximum quit date range. These are similar to the previous parameters, except that the quit ages are here expressed in terms of calendar dates.

Line 9. Maximum benefits at age 55 and 65. These are the maximum annual benefits that the program will allow without truncation. These are meant to constrain the plans to fit within the IRS maximums for benefits from a tax-qualified plan. These only apply to defined benefit plans. Maximum benefits for ages between 55 and 65 are interpolated. These maximums apply to the

survey year; in other years they are adjusted for general wage and salary growth. Setting these maximums to a very large number will in effect remove this constraint.

Lines 10-12. Profit sharing and forfeiture contribution percentages. Defined contribution plans frequently have contributions that are related to some measure of profitability. Some plans also have provisions to redistribute forfeitures among remaining participants. A forfeiture in a defined contribution plan occurs when an employee who is not vested leaves the firm. Since there is no way to tell what the profits are at the firm, or how much money is available from forfeitures, the parameters on this line tell the program how to calculate these contributions. The first value is for profit sharing, and the second is for forfeitures. Both values are expressed as a percentage of salary. In Table 1, the parameters specify that 5% of salary is contributed in profit sharing plans, and 0.5% of salary is contributed if the plan calls for forfeitures to be redistributed.

There are three lines for the profit sharing and forfeiture contribution percentages. In the plans themselves, it is common for the plans to have different contributions for different periods. For instance, a plan might specify a certain percentage of profits before taxes before 1977 and a another percentage of profits after taxes after 1977. These would be coded as two separate variables, and the program would require values for both. The problem is that it is impossible to say exactly how these variables are defined without looking at the plans in detail. In general, one is should probably specify the same percentage contribution for all three lines of profits and forfeitures.

Lines 13-15. Discretionary matching contribution percentages. These are matching contributions, which means that the firm matches to some degree either mandatory or voluntary contributions which are made by employees. However, the match percentage is discretionary on the part of the firm, rather than being set as a given percent of the employees' contributions. The parameters in lines 13-15 specify what percentage of matching is to be assumed in the case of discretionary matching. For example, the parameters in Table 1 specify that in cases of discretionary matching, the program is to assume that the firm matches 50% of the employees contribution. As with the profit sharing and forfeiture contributions, the pension record allows up to three different variables to be defined, Again, there is little reason to assign different values for these three variables.

Line 16. Other unspecified employer contributions and unspecified mandatory employee contributions. The first value on this line refers to employer contributions that were neither profit sharing, forfeitures, nor matching. The second refers to mandatory contributions by an employee. In some cases the plan description makes it clear that these contributions are present but fails to specify how those contributions are determined. For instance, the plan may say that the firm will make a contribution based on years of service and salary but fail to say how the contribution is calculated. The first value on this line specifies the contribution, as a percentage of salary, that is to be assumed if there are unspecified employer contributions that are not profit sharing, forfeitures, or matching, and the second value specifies the contribution, again as a percentage of salary, if there are unspecified mandatory contributions required for the employee.

Line 17. Voluntary contribution percentage bounds. This line applies to plans with voluntary contributions. Most plans specify a minimum and a maximum amount of these contributions. In the participants file (to be discussed shortly), there is a value which allows the user to specify, for each worker, how much they are willing to contribute, subject to these minimum and maximum amounts. In some plans, however, the minimum and maximum amounts are not specified. This line allows the user to set the minimum and maximum amounts, expressed as a percentage of salary, for plans with voluntary contributions but unspecified minimums and maximums on those contributions.

Line 18. Unspecified matching percentage for mandatory and voluntary contributions. This is the final line which indicates what to do if there are unspecified contributions. The first value on the line refers to employer contributions which match mandatory contributions on the part of the employee. The second value refers to employer contributions which match voluntary contributions on the part of the employee. The difference between these values and the values on lines 13-15 is that these matches are not discretionary, just unspecified in the summary plan description. The values are given in percentage terms, so that a value of 50 means that the employer will match half of the corresponding employee contributions.

Line 19. Annuity type. This parameter specifies whether the annual benefits are calculated as single or joint annuities. A value of 1 specifies a single annuity, while a value of 2 indicates that the benefits are always expressed as a joint annuity. A value of 3 specifies a joint annuity if the worker is married (as indicated in the participant file), or a single annuity otherwise. The second parameter gives the percentage of the benefit for the survivor. For instance, a value of 50 means that if the spouse survives the worker, he or she will receive 50 percent of the original benefit. The life tables used to change single annuities to joint annuities are those used by the Social Security Administration in its intermediate projections.

Line 20. New vesting. Several plans in the sample had vesting periods of ten years or more, despite federal laws at the time which generally limited vesting periods to a maximum of five years. If this parameter is set equal to 1, the maximum vesting period is limited to the number of years specified in the second parameter on the line. If the parameter is left at 0, whatever vesting rules that were found in the summary plan description are used.

Line 21. Plan maximum retirement. Several plans had mandatory retirement provisions, despite federal laws to the contrary in 1989. The program will generally not calculate pensions for a quit date beyond the mandatory retirement age. However, if this switch is set to 0, the program will instead ignore any mandatory retirement provisions in the summary plan description and calculate the pension for any age up to the maximum age discussed earlier.

Line 22. Pension type. The amount of the pension may depend on the circumstances under which it is collected. This parameter tells the program which circumstance to apply. A value of 0 refers to a normal separation or retirement. A value of 1 applies if the worker collected the pension as a result of a disability, and a value of 2 refers to the benefits that would be available to

survivors in case the worker dies. In this case, the quit date in the participant file refers to the date the worker died. This parameter has no effect in Mode 3, since in Mode 3 the program will calculate all three values.

Line 23. First year annuity / present value output flag. In Mode 3, the program prints out a column to allow the value of the pensions to be compared at alternative retirement dates. If the value of this parameter is 1, the value of this column is a present value of the stream of benefits. If the value is of the parameter is 0, the output value is the annual value of an actuarially equivalent stream of benefits starting at age 65 (regardless of when the benefits actually start). Since the benefits start at a common date, they can be compared to tell whether the value of the pension is going up or down as alternative separation dates are considered.

Line 24. Wage histories flag. If this flag is set equal to 1, the participants file will contain a separate wage history for each participant. Otherwise, if the flag is set equal to 0, the earnings stream is assumed to follow a quadratic path, with parameters which are specified for each respondent in the participant file. Using wage histories allows earnings to vary year by year in any desired fashion. A partial wage history is allowable. The minimum wage history is one year. The program's normal wage estimation routines calculate values outside the wage history range. A wage history must be used for prior jobs to correctly associate the age at separation with the separation year.

Line 25. Adjust dollar amounts for inflation flag. Many plans contain dollar figures. For instance, a plan may specify that the monthly benefit is $25 for each year of service. If this flag is set equal to 0, the calculations assume that the $25 amount is the same regardless whether the participant quits in 1989 or 2020. If the flag is set equal to 1, then the $25 amount is indexed by the nominal salary and wage growth rates, as determined by lines 3 and 5 above. For instance, if the retirement date is 2010 and salaries are projected to have tripled between 1989 and 2010, the calculations would assume that firms would have raised the initial dollar benefit to $75 per year of service to adjust to the intervening wage and salary inflation.

Line 26. COLA to benefits for no COLA plans flag. Relatively few plans have formal mechanisms for adjusting benefits of existing retirees to accommodate increases in the cost of living. It is relatively common, however, for firms to adjust occasionally the pensions of its retirees on an ad hoc basis to at least partially offset the increases in inflation (as specified by line 3 in this file). If this flag is set equal to 0, the calculations assume that there are no cost of living adjustments (COLA's) once benefits are started. If the flag is set equal to 1, the calculations assume that once benefits are started, the firms increase them by a percentage of the inflation rate. The percentage is specified by the second parameter on the line. For instance, in Table 1, the calculations assume that once benefits start, they will be increased 2% per year. The 2% is 50% of the presumed 4% inflation rate. If the plan has a formal COLA, the plan COLA is used regardless of the setting of this flag. The difference between this switch and the previous one is that the previous switch pertains to the formula used to calculate initial benefits, while this one pertains to increases in benefits after they have started.

Line 27. Mortality tables flag. If this flag is set to 0, the program uses the mortality tables obtained from the Social Security Administration. These are the same tables the Social Security Administration uses to make its intermediate projections. The tables are by gender for single year birth cohorts. If the flag is set to 1, the tables for the 1936 birth cohort are used. Setting the flag to 1 causes the program to run moderately faster. For respondents and their spouses who were born within a decade or so of 1936, the effects on the pension values is slight. For spouses who were born several decades away from 1936, the differences in present values will be more noticeable.

Line 28. Ignore Benefits Paid Before Today's Date flag. This flag applies to calculations of the present value of the pension, and it affects only respondents whose benefits have started before the date in line 2 above. In some cases, what is desired is the total value of the pension, which is the discounted stream of benefits over the entire time the worker collects them. In other cases, what is of interest is only the current value as of a certain date, which includes the present value of current and future payments but ignores previous payments. If this flag is set equal to 0, the present value of the entire stream of benefits is calculated. If it is set equal to 1, then benefits before the date for the present value calculations (line 2 above) are disregarded. For example, to calculate the current value in 1994 of a pension begun in 1986, line 2 should be set to 1994 and line 29 should be set to 1. This flag is applicable to modes 1 and 2 only. Mode 3 always calculates the present value as of the separation date.

These parameters permit a great deal of flexibility in controlling how the program does the calculations. In order to make sure that the program does not try to use values for the parameters that would cause problems, it checks the values of these parameters to make sure that they are within reasonable bounds. These bounds are given in Appendix Table 1. The descriptions of the variables in that table correspond to the descriptions in Table 1 above. If the program does find parameters that are outside the bounds, it resets them to the values given by the third column of the appendix table and writes a message to the error file.

The Participant File for Modes 1 and 3.

The participant file has one line (or set of lines, if wage histories are included) for each participant for which a pension is being evaluated. It contains information about the participant which is crucial for the calculation of the pension, such as the date of birth, the date of hire, and the separation date. For Modes 1 and 3, it also contains identifiers for the plan that should be evaluated for this individual. In Mode 2, which calculates how different pensions would apply to a single participant or small group of participants, this information is not needed. Therefore, the participants file looks slightly different depending on which mode is being used.

Table 2 gives a sample participant file for Modes 1 and 3. The distribution package actually contains six participant files. There are separate participant files for the current job, covered in Section F of the survey, the last job, covered in Section G, and a previous job lasting at least 5 years, covered in Section H. For each section, there is a separate file for defined benefit

Table 2
Sample Participants File
for Modes 1 and 3

```
 1  127 378    0.00 2 1930.04 1979.12 1992.50 2080  27206 0 5 0 0  1546.69
 1 1067 658    0.00 2 1930.04 1979.12 1992.50 2080  27206 0 5 0 0  1546.69
 2  256 157 1933.21 2 1935.96 1971.04 1992.50 2080  40997 0 5 0 0  1996.09
 3   13   7 1933.29 2 1934.04 1974.71 1992.50 2080  15808 0 5 0 0  1563.55
 3  127 378 1934.04 1 1933.29 1959.50 1992.50 2340  25000 0 5 0 0  1563.55
 3 1067 658 1934.04 1 1933.29 1959.50 1992.50 2340  25000 0 5 0 0  1563.55
 4   13   7    0.00 2 1935.12 1976.62 1992.50 2080  14560 0 5 0 0  2426.90
 8 1143 551    0.00 2 1935.88 1965.62 1992.50 2080  20488 0 5 0 0  2212.61
12  240 838    0.00 2 1924.12 1968.71 1992.50 1820  28592 0 5 0 0  6718.13
19 1088 356 1930.54 2 1932.96 1971.88 1992.50 1040  15600 0 5 0 0  6313.77
24  101 575 1928.29 1 1928.54 1947.46 1992.50 2080  20000 0 5 0 0  7693.62
28  474  85 1936.29 1 1937.12 1970.62 1992.50 1820  40000 0 5 0 0  6734.19
38  411 586 1928.29 2 1930.29 1983.29 1992.50 1300  10400 0 5 0 0  7427.04
38 1197 593 1928.29 2 1930.29 1983.29 1992.50 1300  10400 0 5 0 0  7427.04
40  110 651 1934.29 1 1932.12 1983.46 1992.50 2040  45000 0 5 0 0  6841.46
```

plans and for defined contribution plans. The files are named as follows:

PARTIC-F.DB	Defined benefit plans for current (Section F) jobs.
PARTIC-F.DC	Defined Contribution plans for current (Section F) jobs.
PARTIC-G.DB	Defined benefit plans for last (Section G) jobs.
PARTIC-G.DC	Defined Contribution plans for last (Section G) jobs.
PARTIC-H.DB	Defined benefit plans for previous (Section H) jobs.
PARTIC-H.DC	Defined Contribution plans for previous (Section H) jobs.

The participant files containing the defined benefit plans must be run with the CALC6DB.EXE program, and the files containing a defined contribution (and combination) plans must be run with the CALC6DC.EXE program. Using the wrong program to evaluate a pension will cause that pension to be ignored in the calculations.

The name of the participant file is not important. The program will accept any participant file with a valid DOS name as long as the command line for the program is modified accordingly. The participant file in Table 2, and all of the output discussed in Part IV of this guide, does not come from any of the included participant files. This is to allow this users guide to be circulated without violating the rules for restricted data files, which govern the distribution of the program and its files.

In the participant file, there is one line (or set of lines, if wage histories are used) for each participant. For each participant, the values in the columns give information about the individual and the plan being evaluated. In the table, the columns are arranged vertically to make them more readable, but this is not necessary. A space between the values is sufficient to separate the values. Again, when producing this file with a word processor or spreadsheet, be sure to save it in a ASCII TEXT format.

It will probably be easiest to discuss the values in each column in turn.

Column 1. Person identifier. This does not affect the calculations in any way, but it is printed in the output to help identify the line in the output. The most logical variable to use is the HRS code id variable, but this is not necessary. Nor is it necessary that this number be unique, as is illustrated by the fact that the first two lines have the same value for this variable. This duplication also illustrates another point: the same worker may be covered by more than one plan on a particular job. If this is true, there will be a separate line in the participant file for each plan.

Column 2. Pension plan ID. This gives the ID number of the pension plan that is being evaluated. Plan ID's less than 3000 are for defined benefit plans and should be processed by CALC6DB.EXE, while plan ID's greater than 3000 are for defined contribution and combination plans and should be processed by CALC6DC.EXE. The combination of the person identifier and the pension plan ID as contained in the six included PARTIC files provides the only link between the HRS respondents and their pensions. There are no pension identifiers in the main HRS data

set. Note that although the six included PARTIC files are necessary to link the pensions to HRS respondents, it is possible to replace the actual HRS respondents with other respondents (as, for example, when pensions are imputed) or with completely hypothetical participants.

Column 3. Pension plan sequence number. This numbers the plans by the order in which they were collected, and is of no special significance outside the program. However, each pension plan ID has a unique sequence number, and the two numbers must match in the participant file. If pensions are evaluated for hypothetical participants, the hypothetical participants must still have valid pension plan ID's and corresponding sequence numbers, and the only place these can be obtained is from the six included PARTIC files.

Column 4. Birth date of the spouse. This is necessary for some calculations involving joint annuities. If there is no spouse, this variable is set equal to zero. Note that this variable can contain a fractional part. For instance, the value 1934.42 would indicate a birth date of approximately June 1, 1934. An integral value of this variable would be interpreted as January 1 of the given year.

Column 5. Gender of worker. This should be 1 for males, 2 for females. This affects the calculations because different life tables are used for the two sexes. These life tables, which are different for each birth cohort, are the life tables that the Social Security Administration uses for \their intermediate projections.

Column 6. Birth date of the worker. As in the case for the birth date of the spouse, the fractional part of this value indicates in what month and day in the year the individual was born.

Column 7. Date of hire at the firm. This variable is self explanatory. Again, fractional values indicate the part of the year that the worker was hired.

Column 8. Date the worker separates or retires from the firm. This variable is used in calculation Modes 1 and 2 only. In Mode 3, it is ignored, since in this mode the program calculates pension values for alternative separation dates. The month and year are reflected in the fractional part of the value.

Column 9. Number of hours that the participant works in a year. This is required because some plans specify that the years of service that a worker accumulates in a calendar year depends on the number of hours worked that year. For example, someone who worked 2500 hours in a given year may accumulate 1.25 years of service if the firm counts a year of service for every 2000 hours.

Column 10. Annual dollar amount of salary or wages. If the individual was working at the time of the survey, this value is the (nominal) amount of salary or wages in the survey year (as determined by line 1 of the parameters file). If the individual had left the job, the value is NOT the wage or salary at the separation date, but the wage or salary that the worker would have

received had he or she stayed on the job until the survey date. To project the wage between the separation date and the survey date, the analyst must take account of (i) the inflation rate, which is line 3 in the parameter file, (ii) the real salary and wage growth rate, which is in line 5 of the parameter file, and (iii) any quadratic growth specified in columns 13 and 14 below. A perhaps easier alternative, which does not require this projection, is to use the wage history to specify the wage at the separation date. If the wage histories flag is set in the parameters file, the value in column 10 is value is ignored.

Column 11. Differential wage growth. This allows the wages of different respondents to grow at different rates. For instance, if the inflation rate in the parameters file is set to 4% and the real wage and salary growth rate in that file is 1%, and if this variable is set to -2%, this individual's nominal wage growth rate would be 4% + 1% - 2% = 3%. Note that this value is a percentage, so that 2% would be written in the file as 2.0, not 0.02.

Column 12. Voluntary contribution rate. This value applies to defined contribution or combination plans which allow voluntary contributions. The figure is a percent of salary or wages. Thus, the value of 5 in Table 1 specifies that the worker who can choose the amount of contributions to the plan contributions an amount equal to 5% of wages. Note that this does not apply to any contributions that are mandatory; those contributions are figured separately.

Column 13. Linear wage growth coefficient (beta1).
Column 14. Quadratic wage growth coefficient (beta2).

These terms allow for real wages to be growing at a quadratic rate. They suppose that log earnings, after allowing for general and individual growth, are following a trajectory of:

$$\ln(\text{earnings}) = \text{beta0} + \text{beta1} * \text{age} + \text{beta2} * \text{age} * \text{age}$$

The growth of log earnings is the derivative of this expression, or

$$\text{growth of } \ln(\text{earnings}) = \text{beta1} + 2 * \text{beta2} * \text{age}.$$

Adding back in the growth of general and individual wages, the growth of log earnings is

ln (earnings at time t) - ln (earnings at time t-1)

=	infl rate / 100	[line 3 of parameter file]
+	real salary and wage growth rate / 100	[line 5 of parameter file]
+	differential wage growth rate / 100	[column 11 above]
+	beta1 + 2 * beta2 * age.	

If beta2 is negative, then as the worker ages, the wages grow at a slower and slower rate. By choosing appropriate values of beta1 and beta2, the user can accommodate any wage function which is quadratic in age, experience, and/or years of service. Note that unlike the inflation rate

and the previous two wage growth rates, which are expressed in percentage terms, the betas are fractions.

Column 15. Pass through variables. The remainder of the line is treated as a long character field. It may have numbers and/or words, and they may or may not be separated by spaces, commas, or other separators. The program does not use any information in this field in the calculations. The sole effect of these variables is that they are printed at the end of the output line in Mode 1. This character field can be used to pass through variables which will facilitate further analysis by standard statistical packages. The number of characters in this filed is limited to no more than 132. In the participant file supplied with the program, the pass through variables consists of weights. Many alternative uses are possible, however.

If the wage history switch is set in the parameter file, the previous information must be followed by a separate line or lines containing the wage history data. The wage history may be continued on successive lines if necessary. The format for this packet of information is:

Start year End year Start year wage .... End year wage -1

The first two entries are the year the wage history starts and ends. Note that the wage history may start after the hire date and end before the date of separation. If this is the case, the wage history will be used for the wages in those years, and the years before the beginning and after the end of the history will be constructed by using the previous parameters. The history can contain just one year, as when the final wage is given for a job that ended before the survey. If the wage history covers the entire period of employment, the previous parameters are in effect ignored. The wage history proper contains one entry for each year between the start year and the end year, inclusive, and may be spread out over more than one line. An example of a multiple line wage history is:

```
1980.00 1988.00 25440.00 26130.00 23650.00 28390.00 25990.00 29110.00 30120.00
29860.00 30960.00 -1
```

The -1 value (any negative number will serve the same purpose) serves to terminate the wage history, even if there are not sufficient number of entries in the wage history to cover all years between the start year and end year. Extra entries are ignored.

As was the case for the parameter file, the values in the participant file are checked for unreasonable values. Appendix Table 2 gives the minimum and maximum amounts for the values in this file, as well as the value to which the variable is set if the specified value lies outside these bounds. Again, if any variable is reset as a result of the bounds checking, a message will be written to the error file.

The Participant and Plan Files for Mode 2.

When the program operates in Mode 2, it requires a slightly different participant file, and

in addition it requires an extra file to specify the pension plans which are to be used. Recall that mode 3 calculates the values of all the pensions for a single or small group of individuals. Table 3 illustrates the participant and plan files required to use Mode 2.

The participant file is very similar to the file with Modes 1 and 3, with two exceptions. The first is that the first three columns of the previous file are omitted. The participant file in Mode 2 begins with the spouse birth date, and the order of the subsequent columns is the same as in the previous two modes. There can be no more than five lines in the participant file for Mode 2.

The plans to be evaluated are in a second file, which we will call the plan file. This file has three columns, as follows:

Column 1: Five-digit line identification number, starting with 9. These must be unique, but otherwise they have no consequence.

Column 2: Pension plan id. This is same variable which is in Column 2 of the participant file for Modes 1 and 3.

Column 3. Pension plan sequence number. This is the sequence number which corresponds to the pension plan id. See the previous discussion for Column 3 in the participant file for details.

Mode 2 evaluates all plans in the plan file for a set of up to five participants. This is useful for comparing how different plans treat the same individual. It is also useful for comparing how plans reward various characteristics. For example, the participant file could contain three hypothetical participants, one with 10 years of service, one with 20, and one with 30. The results could then give some insight on how the various plans reward differentially their longer service workers.

Table 3
Sample Files for Mode 2

Sample Participant File

```
   0.00 2 1930.04 1979.12 1992.50 2080  27206 0 5 0 0  1546.69
1934.04 1 1933.29 1959.50 1992.50 2340  25000 0 5 0 0  1563.55
```

Sample Plan File

```
90000  127 378
90001 1067 658
90002  256 157
90003  267 799
90004   13   7
90005 1062 456
90006 1143 551
90007  240 838
90008  537 758
90009  380  83
```

III. Running the Program.

The program is run from the DOS command prompt. The command line is slightly different for the three modes. The first mode, Mode 1, calculates the value of the pension benefits for each participant based on the separation date specified in the participant file. For this mode, the command line to evaluate defined benefit plans is as follows:

```
CALC6DB 1 PARTIC PARAM OUTDATA TABLES ERRORS
```

The command line for CALC6DC is exactly analogous, except that the participant file in this case must contain only defined contribution and combination plans. The first parameter following the program name is the mode number, in this case 1. The next two files are the participant and parameter files, as discussed in the last section. These files do not have to be called PARTIC and PARAM, and in particular users will probably want to use one of the participant files included with the program. If the files have different names, however, those names must appear in the command line in the appropriate order. The next three files are output files, which will be discussed in the next section. TABLES is a file in tabular format which is relatively readable using a viewer. OUTDATA is the same information in numeric format, which will be easier to import into subsequent statistical packages. ERRORS is a listing of errors and potential errors that the program found.

Mode 2 allows the program to calculate the value of many pensions for a limited set of individuals. The command line for this mode is:

```
CALC6DB 2 PARTIC PARAM PLAN OUTDATA TABLES ERRORS
```

Note that this command line contains one additional parameter. PLAN is the file containing the pension plan id's and sequence numbers for the plans to be evaluated. In this case, PARTIC contains the input values of the participants for whom these plans are to be evaluated.

Mode 3 allows the program to calculate the value of the pension at alternative separation dates for participants in the participant file. The format of the command line for Mode 3 is:

```
CALC6DB 3 PARTIC PARAM OUTDATA TABLES ERRORS
```

This is exactly the same format as for Mode 1, except that the mode number after the program name is 3. All of the files have the same functions as with Mode 1.

During the execution, the program will indicate its progress. Since the program may take an extended amount of time for large participant or plan files, this is an assurance that the program is still running. The program is fairly robust against errors in the input files, but there are a couple of things that may cause a premature halt. The first of these is if there are not

enough parameters on the command line. In this case, the program will print out the message:

```
Missing parameter from command line
```

The second error is if the program cannot find one of the input files, or if one of the files has an invalid name (for instance, if it has two periods). This will cause the program to print out the not very enlightening message:

```
Runtime error 002
```

If this message appears, check (i) that the input files exist, and (ii) that all files have proper DOS names (up to 8 characters plus an extension of up to 3 characters).

IV. Interpreting the output.

The output varies depending on the mode in which the program is run. Perhaps the easiest way to examine the output is to look at the TABLES file for each of the three modes. We will say a few words about the OUTDATA file, and only a couple about the ERROR file. Each of these files can be given any legitimate DOS name on the command line, and as a result, the files discussed in this section will have the name you have given it.

The first part of the TABLES file is the same for all three modes. It is simply a listing of the parameters in the parameter file. This is admittedly slightly repetitious, but it alleviates the necessity of keeping track of what parameters file went with which output file.

Output of Mode 1.

Table 4 gives an example of the tabular output using Mode 1. An asterisk at the beginning of the row means that the program has written a message to the error file for this plan. The first eight columns simply repeat what is in the input file. The wage is the nominal wage as of the quit date. The age is calculated from the difference between the quit date and the birth date, and the asy (which stands for actual service years) is the difference between the quit date and the hire date.

The next column is the type of benefit. The symbols in the table are as follows:

NR	Defined Benefit Normal Retirement Benefits
LR	Defined Benefit Late Retirement Benefits
ER	Defined Benefit Early Retirement Benefits
VD	Defined Benefit Deferred Vested Benefits
DB	Defined Benefit Disability Benefits
SB	Defined Benefit Survivor Benefits
CN	Defined Contribution Benefits
NC	Combination Normal Retirement Benefits
LC	Combination Late Retirement Benefits
EC	Combination Early Retirement Benefits
VC	Combination Deferred Vested Benefits
DC	Combination Disability Benefits
SC	Combination Survivor Benefits

In general, the plans are of three types: (i) defined benefit, which calculate benefits based on a formula usually involving pay and/or years of service, (ii) defined contribution, which calculate benefits based on the amount accumulated in an account, and (iii) combination plans, whose benefits have both defined benefit and defined contribution elements.

For defined benefit plans, the three benefits are (i) normal retirement benefits, which are

Table 4
Sample Tabular Output of a Case from Run Mode 1

	HHIDX	Code	Seq#	Sex	SpouseBD	BirthD	HireD	Date	Wage	Age	Asy	Tp	Age	%QW	%/Yr	30y	AnualAm	Pr.Value	A65eqvl
*	1	127	378	2	0.00	1930.04	1979.12	1992.50	27206	59.46	10.38	ER	59.46	12.90	1.24	37.28	3509	50933	5453
	1	1067	658	2	0.00	1930.04	1979.12	1992.50	27206	59.46	10.38	CN	59.46	6.21	0.60	17.94	1689	20130	2625
	2	256	157	2	1933.21	1935.96	1971.04	1992.50	40997	53.54	18.46	VD	55.00	11.25	0.61	18.28	4611	64975	9932
	3	13	7	2	1933.29	1934.04	1974.71	1992.50	15808	55.46	14.79	ER	55.46	6.00	0.41	12.16	948	14764	1984
*	3	127	378	1	1934.04	1933.29	1959.50	1992.50	25000	56.21	30.00	ER	56.21	37.20	1.24	37.20	9299	128402	19140
	3	1067	658	1	1934.04	1933.29	1959.50	1992.50	25000	56.21	30.00	CN	56.21	23.47	0.78	23.47	5869	67191	12079
	4	13	7	2	0.00	1935.12	1976.62	1992.50	14560	54.38	12.88	VD	55.00	4.91	0.38	11.45	715	10670	1541
	8	1143	551	2	0.00	1935.88	1965.62	1992.50	20488	53.62	23.88	CN	53.62	11.49	0.48	14.44	2355	29862	5585
	12	240	838	2	0.00	1924.12	1968.71	1992.50	28592	65.38	20.79	NR	65.38	35.84	1.72	51.71	10246	131005	10008
	19	1118	356	2	1930.54	1932.96	1971.88	1992.50	15600	56.54	17.62	CN	56.54	9.64	0.55	16.41	1503	18322	2874
	24	101	575	1	1928.29	1928.54	1947.46	1992.50	20000	60.96	42.04	VD	62.00	25.10	0.60	17.91	5019	55112	6471
	28	474	85	1	1936.29	1937.12	1970.62	1992.50	40000	52.38	18.88	ER	52.38	17.92	0.95	28.47	7166	87021	19772
	38	411	586	2	1928.29	1930.29	1983.29	1992.50	10400	59.21	6.21	VD	65.00	4.01	0.65	19.35	417	3513	417
	38	1197	593	2	1928.29	1930.29	1983.29	1992.50	10400	59.21	6.21	CN	59.21	2.11	0.34	10.17	219	2610	346
*	40	110	651	1	1934.29	1932.12	1983.46	1992.50	45000	57.38	6.04	VD	65.00	2.78	0.46	13.81	1252	7732	1252
	MEAN			1.67	1932.06	1932.21	1971.60	1992.50	24417	57.29	17.90		58.40	14.05	0.73	22.00	3655	46149	6632

unreduced benefits collectable immediately on retirement, (ii) early retirement benefits, which are reduced but collectable immediately, and (iii) vested deferred benefits, which are not payable immediately but only after a period of time. Late retirement benefits are normal retirement benefits that are augmented beyond the regular formula because the worker has stayed with the firm beyond the normal retirement age. Disability benefits are payable if the worker has suffered a disability, and survivor benefits are payable to the survivor if a worker dies. These last two benefits are triggered by line 22 in the parameter file.

The next column is the age at which the benefit can begin. Most benefits begin immediately upon retirement, but deferred vested benefits begin in the future. The next three columns relate the annual annuity amount (in the third to the last column) to wages. The first column, labeled as %QW, gives the annuity amount as a percentage of the last wage, and the %/YR column further divides this by years of service. The value of the pension, divided by the wage and by years of service, is sometimes referred to as the "generosity" of the plan. The third column multiplies this "generosity" by 30. This is meant to give a (very rough) indication as to the percent of final pay that the plan would replace if the worker had stayed with the firm for 30 years.

The last three columns bear some explanation. Let us start with the column labeled Pr.Value. This is the present value of the benefits, discounted to the date specified in the second line of the parameter file. The present value calculations do take account for the likelihood of surviving to collect the benefits. The survival probabilities are taken from the social security intermediate projections, and are specific to gender and birth cohort. Recall that if the last line of the parameter file is set equal to 1, the present value calculations ignore benefits paid before the date specified in line 2 of the parameter file.

The previous column, labeled as AnualAm, is the nominal amount of an annuity which has the same present value as the next column. If line 26 in the parameter file is set to 1, this annuity is presumed to be adjusted upward at some percentage of inflation, as given by the second parameter of line 26 of the parameter file. In this case, the value given in the AnualAm column is the nominal value of the first year annuity. Note that this may or may not be the same as the first year benefit. Many plans have benefits which change discontinuously as some date, typically 62 or 65. Some of these plans offer a supplement which continues until the worker is eligible for social security benefits, and others subtract off from their benefits a fraction of social security benefits after the worker begins collecting them. The AnualAm value, by contrast, assumes that the benefit will be constant (or will be adjusted only for inflation) after the benefit begins. This means that if the benefits are in fact from an offset plan, where the benefits are reduced when social security begins, the value in the AnualAm column will underestimate the pension before the offset begins and will overestimate it thereafter. It bears emphasizing that the present value of the pension is calculated correctly, and the only caveat is that the AnualAm value is not necessarily the same as the pension benefit in any given year.

The final column, labeled as A65eqvl, is the nominal amount of an annuity which has the

same present value as the previous column, but which starts at age 65. The same caveats about the adjustment for inflation, as discussed in the last paragraph, apply. Note that because these are nominal (and therefore undiscounted) amounts as of age 65 (which will occur in different years for different individuals), they do not necessarily give the same ranking of the pensions as do the present value amounts in the previous column, which are discounted back to a common year.

The OUTDATA file presents essentially the same information, but without the headers and entirely in numeric format. There is one line for each plan in the participant file; the summary information from the last line of Table 4 is omitted. The line corresponding to the first line in Table 4 is as follows:

```
  1  127  378 2    0.00 1930.04 1979.12 1992.50    27206.00  59.46  10.38 03
     59.46  12.90   1.24  37.28    3509.00    50933.36    5453.38   1546.69
```

In the actual output, this would all be on a single line. There are several things to note about this line. First, the dollar amounts are more precise, being given to two decimal places rather than being rounded to the nearest dollar, as appears in the table. Second, the pension type column (which had an entry of ER in the first line of the tabular format) has been converted to a numeric quantity. The codes for the benefit types is as follows:

1	Defined benefit NR benefits	11	Combination NR benefits
2	Defined benefit LR benefits	12	Combination LR benefits
3	Defined benefit ER benefits	13	Combination ER benefits
4	Defined benefit VD benefits	14	Combination VD benefits
5	Defined benefit DB benefits	15	Combination DB benefits
6	Defined benefit SB benefits	16	Combination SB benefits
7	DC benefits	19	Combination plan, DC benefits only
9	No benefits		
-1	Error in calculations		

For the meaning of NR, LR, ER, etc., please see the discussion of types of benefits earlier. Finally, note that the last item in the OUTDATA line is the pass through variable.

The ERRORS file frequently generates a substantial amount of output, but most of the output can probably be disregarded. The program contains many checks regarding suspicious calculations, for instance, particular sub-formulas which yield a negative value. Usually these are irrelevant, such as when the plan provides for an alternate formula which gives positive results. In general, there are two instances of entries in the ERRORS file where the one might want to pay particular attention. The first of these are lines that begin with DATA. This is the message which indicates that some parameter or value in the input file was outside the bounds given in the appendix tables and was reset to some other value. The calculations are probably all right, but they may not correspond to the calculations that the user had intended. The second place where caution is advised are errors that are accompanied by negative entries in the tabular

or data files. This usually indicates a genuine error in the calculations, as opposed to a case where there was an unusual value in an intermediate calculation.

Output of Mode 2.

An example of the output of Mode 2 is given in Table 5. This table has two parts. The first is a listing of the participants, up to five of them, from the participant file. The first 7 variables in this part of the table are many of the same participant variables discussed in Table 4, although in a slightly different order. The variable labeled as "IndVr" is the differential wage and growth rate, which is column 11 from the participant file. The next variable, labeled "VolCt", is the voluntary contribution rate, from column 12 of the participant file. The last two columns are age and service years, as calculated from the birth date, hire date, and quit date.

The second part of the table contains the computed results. Down the left side of the table are the pension plan identification numbers, with each line corresponding to one line in the plan file. Across the table are up to five groups of five variables each. Each group, which is delimited by the dashes across the first line of the heading, pertains to one of the participants listed in the first part of the table. Each group, in turn, lists five calculated values of the pension as applied to the indicated participant. These five values are the same as the variables on the right side of Table 4: (i) the type of benefit, indicated by a single letter rather than a two letter code, (ii) the ratio of the pension to the final wage, expressed as a percent, (iii) this ratio divided by the years of service, which is a measure of generosity, (iv) the annual amount of the annuity which has the same present value of the pension, and (v) the present value of the pension. These variables are defined in exactly the same way as in Mode 1, with only minor difference. To avoid lines longer than 132 characters, the benefit type is now represented by a single character, which means that the benefits of defined benefit and combination plans are not separately identified in this output. Also, for the same reason, the pension to wage ratio is now rounded to the nearest percent. One note of caution: the amount in the fourth column is not necessarily the same as the annual pension benefit if those benefits are not constant. See the discussion of this variable in the section for the output of Mode 1.

The OUTDATA file has one line for every line in the plan file. The first line of that file that corresponds to the tabular results in Table 6 is as follows:

```
 127 90000  127  378 3  12.90   1.24    3509.00    50933.36 90000  127  378 3
   37.20   1.24    9299.23   128401.73
```

As before, in the actual file this is a single line. The line begins with the pension plan identification number. Before each group of five values, corresponding to one of the participants, appear the identification number in the first column of the plan file. the pension plan identification number, and the pension sequence number. As before, all percentages and dollar amounts are expressed to two decimal places in the data file. The numeric codes for the plan

Table 5
Sample Tabular Output of a Case from Run Mode 2

#	Sex	SpouseBD	BirthD	HireD	QuitD	Hour	Survey Wage	IndVr	VolCt	Age	Asy
1	2	0.00	1930.04	1979.12	1992.50	2080	27206	0.00	5.00	59	10
2	1	1934.04	1933.29	1959.50	1992.50	2340	25000	0.00	5.00	56	30

	1					2				
Code	T	%W	%/Yr	AnualA	PrValue	T	%W	%/Yr	AnualA	PrValue
127	E	13	1.24	3509	50933	E	37	1.24	9299	128402
1067	C	6	0.60	1689	20130	C	23	0.78	5869	67191
256	E	8	0.79	2222	32256	E	20	0.66	4931	68091
267	V	3	0.33	932	7249	V	22	0.73	5481	28409
13	E	4	0.39	1091	15836	E	9	0.30	2232	30822
1062	C	16	1.56	4413	52590	C	57	1.90	14278	163470
1143	C	5	0.45	1264	15064	C	17	0.56	4211	48210
240	V	18	1.72	4868	41680	E	25	0.84	6281	86727
537	V	21	2.00	5636	78917	N	66	2.21	16542	281820
380	V	6	0.59	1657	18870	V	19	0.64	4789	37456
MEAN		10	0.97	2728	33353		30	0.99	7391	94060

types are similar to those in the data outputs for the other two Modes, except that numeric codes 11-17 (which are for combination plans) have been mapped into the corresponding 1-7 values, reflecting that in Mode 2, defined benefit plans are not distinguished from combination plans.

Output of Mode 3.

Table 6 illustrates an example of the tabular output generated for Mode 3. There are several differences between this output and that for Mode 1 and Mode 2. One of the most evident is in the volume of output. All of the output listed in the Table 6 is generated from a single line in the participant file. For each participant, the program generates two parts to the output. The first part, along the top of the table, is simply a listing of the various components of the input line from the participant file. Note that the value of the separation date, which is column 8 of the participant file, is not used in this mode and is not listed in the first part of the output. Nevertheless, some value for the separation date must be present in the participant file.

The second part of the output has one year for each potential separation date from the hire date through the maximum quit age or date from the parameter file. Note that if the plan has a mandatory retirement age, and if this age is allowed to stand by setting the value in line 21 of the parameter file to one, the calculations will end at the maximum retirement age.

The first four columns are the date, age, number of service years, and the wage as of that date, as discussed previously. The fifth column is the annual social security benefit which would be payable at that age. Note that since social security is not available until age 62, this column has a value of zero until that age. The next two columns give the annual pension benefit (in nominal terms) and the type of benefit. Like the social security benefit, this column has entries only if the worker is immediately eligible for the benefit; hence, it has a value of zero in years before early retirement. This variable has a label which is slightly misleading, since the value in this column is actually the value of an annuity whose present value is the same as the actual stream of benefits. That is, in cases where there is a social security supplement or offset, or some other reason that the benefit is not a constant stream, the program first calculates a present value of the stream of benefits. It then calculates the value of an annuity amount which has the same value as the present value of the actual stream of benefits, and reports it in this column. Refer to the discussion of the AnualAm column from Mode 1. The same caveats as were made in that discussion apply with regard to increases to accommodate inflation, as specified by line 26 in the parameter file.

The next two columns refer to the amount of the vested benefit and the date that it may first be collected. If the worker is already eligible for benefits at retirement, the amount in the first of these columns should be the same as the amount in the previous column, and the age should be the current age on that line. Otherwise, the amount will be the amount the individual is eligible for as a deferred vested participant, and the age (in the future) when the benefits can start to be collected. If the worker is not yet vested, this value will be zero.

Table 6
Sample Tabular Output of a Case from Run Mode 3

HHID	CodeId	Seq#	Sex	SpouseBirthDate	BirthDate	HireDate	Hours	IndVr	VolCt	Beta1	Beta2
1	127	378	2	0.00	1930.04	1979.12	2080	0.00	5.00	0.000000	0.00000000

FF	Date	Age	ASY	Wage	SSB	PensnB	Tp	Vested	BA	VestPV	DableB	Tp	NRT	LRT	ERT	VDT	DRT	DST	CNT
	1979	49	0	16435	0	0	--	0	60	0	0	--	0	0	0	0	0	0	0
*	1980	50	1	17285	0	0	--	0	60	0	0	--	0	0	0	0	0	0	0
	1981	51	2	18178	0	0	--	0	60	0	0	--	0	0	0	0	0	0	0
	1982	52	3	19118	0	0	--	0	60	0	0	--	0	0	0	0	0	0	0
	1983	53	4	20106	0	0	--	0	60	0	0	--	0	0	0	0	0	0	0
	1984	54	5	21146	0	0	--	0	60	0	0	--	0	0	0	0	0	0	0
	1985	55	6	22239	0	0	VD	2029	60	21157	0	--	0	0	0	1220	0	1066	0
	1986	56	7	23388	0	0	VD	2372	60	26455	0	--	0	0	0	1579	0	1246	0
	1987	57	8	24597	0	0	VD	2714	60	32406	0	--	0	0	0	2001	0	1426	0
	1988	58	9	25869	0	0	VD	3057	60	39083	0	--	0	0	0	2495	0	3211	0
	1989	59	10	27206	0	0	VD	3400	60	46570	0	--	0	0	0	3071	0	3571	0
	1990	60	11	28612	0	3743	ER	3743	60	54959	3743	ER	0	0	3743	3743	0	3931	0
	1991	61	12	30091	0	4188	ER	4188	61	60281	4188	ER	0	0	4188	4188	0	4291	0
	1992	62	13	31647	9276	4815	ER	4815	62	67879	4815	ER	0	0	4815	4815	0	4815	0
	1993	63	14	33283	10392	5483	ER	5483	63	75641	5483	ER	0	0	5483	5483	0	5483	0
	1994	64	15	35003	11808	6169	ER	6169	64	83234	6169	ER	0	0	6169	6169	0	6169	0
	1995	65	16	36813	13332	6918	NR	6918	65	91238	6918	NR	6918	6918	6918	6918	0	6918	0
	1996	66	17	38715	14652	7573	NR	7573	66	97557	7573	NR	7573	7573	7573	7573	0	7573	0
	1997	67	18	40717	16092	8251	NR	8251	67	103776	8251	NR	8251	8251	8251	8251	0	8251	0
	1998	68	19	42821	17604	8969	NR	8969	68	110048	8969	NR	8969	8969	8969	8969	0	8969	0
	1999	69	20	45035	19260	9709	NR	9709	69	116107	9709	NR	9709	9709	9709	9709	0	9709	0
	2000	70	21	47363	21036	10103	NR	10103	70	117644	10103	NR	10103	10103	10103	10103	0	10103	0

The next line may take on one of two forms, depending on the value of the switch set in line 23 of the parameter file. In the illustration, that switch has be set to 1, which has caused the present value of the pension to be printed out. Note that this present value is the value discounted back to the age of separation, and not discounted to the date in line 2 of the parameter file. Because they are discounted to different dates, unless the quit date is the same as the date in line 2 of the parameter file, the present values in Mode 3 will differ from the present values calculated in either Mode 1 or Mode 2. If the switch on line 23 of the parameter file is set to 0, then this column will print out the value of the annuity if it were started at age 65. This value corresponds to the A65eqvl column in Mode 1.

The next two columns print out the benefit amount and type of benefit that a disabled worker would be eligible for if he were disabled at the given age. Since the plan illustrated in Table 6 has no specific provisions for disability, this worker would receive the same benefits as a regular retired worker. Please be cautioned that the benefits reported in this column are only the benefits payable through the pension plan. The actual disability benefits available to the worker may well include amounts from other plans specifically for disability, and hence the pension disability benefits may be a misleading indicator of the income available to a disabled worker.

The next 7 columns are the benefits that are available under various formulas. These seven columns are, in order, normal retirement benefits, late retirement benefits, early retirement benefits, deferred vested benefits, disability benefits, survivor benefits, and defined contribution benefits. Zeros in these columns indicate either that the firm does not offer that type of benefit, or that the worker is not eligible for that type of benefit based on the age and number of years of service.

The OUTDATA file for Mode 3 contains one line for each line in the second section of the tabular output, excluding the heading lines. The line in the data file corresponding to the line for 1989 in Table 6 is:

```
1  127  378  1989.04  59   9.92    27206.00        0.00       0.00 04
 3399.78 60.00    46570       0.00 09       0.00       0.00       0.00
 3071.06       0.00    3571.20       0.00
```

Again, in the file this information all appears on a single line. It is very much the same as the data in the tabular output, with three exceptions. First, three variables have been appended to the front of the line, representing the household identification number, the pension plan identifier, and the pension plan sequence number. Secondly, all the dollar amounts are given to two decimal places, as are the date, number of service years, and one of the age columns. Finally, the columns representing the types of benefits have numeric codes in the data file. These codes are the same codes as were discussed in the Mode 1 results. Note that unlike the Mode 1 results, the pass through variables do not appear in the Mode 3 data file.

V. Further Details of the Pension Calculations.

The overall purpose of the above sections is to give the user enough information to run the pension calculation program and alter important elements of the calculations. In the parameter input file, the user is free to alter the inflation, interest rate, wage growth, and other assumptions which affect the calculations for pensions in general. In the participant input file, the user can change the characteristics of the participants and the way the participants are matched to the pensions. In the command line, the user is able to specify three modes of calculation, each useful for a class of problems. By varying the input files and choosing the appropriate mode, most users will be able to adapt this program to their needs.

However, the actual plans and the program for calculating them are still a black box. For the user who wants to know more about the plans, these are available in a separate restricted data file which is about 33 megabytes long. There are 1717 pensions in this file, each with a single record of 19249 bytes. The records are in ASCII format and contain a large number of variables. There are two additional files which have split the plans into defined benefit plans and defined contribution and combination plans.

To understand the variables in the pension records, it is helpful to recall how the pensions were collected. When the HRS was fielded in 1992, respondents were asked for the name and address of the firms which offered the most important pension. Subsequently, the survey staff contacted the employers and obtained a document called the summary plan description of the pension. Next, they used the summary plan description to fill in a "questionnaire" about the pension. This questionnaire asked, among other things, how service years and pay variables were defined, how these were combined to calculate the pension, what requirements (including vesting) were required to collect different kinds of benefits, and many others. These questionnaires were then entered into computer to generate the records in the file.

Three pieces of documentation are helpful to interpret the records. First, a dictionary file lists the variables, and for each variable it indicates the position of the variable in the record and the length of the variable. Secondly, to interpret the variable, a codebook is required. The codebook lists the variables and gives the codes for each variable. For instance, a final average pay variable may be coded as 04, which means the "average during the final # mo/yrs." Other variables give the number and tell whether this value represents the number of months or the number of years. Finally, as is frequently the case with any survey, a copy of the questionnaire is helpful.

These records contain a great deal of information about the pensions, and there are undoubtedly some questions that the actual pension records will be useful in answering. However, the records are fairly involved, and one must be cautious in examining these records. They contain well over a thousand variables, and small changes in a single variable can sometimes have large effects on the values of the pensions that are calculated. Even relatively simple questions can be tricky. For instance, an early retirement benefit might be calculated with

different formulas, depending on age and years of service, and each with a different reduction factor. To find the early retirement reduction, it is necessary to find out with formula applies before evaluating the reduction factor.

Although the pension records contain a wealth of information, they are difficult to use with standard statistical packages. The pension calculation program spares most users of the burden of working with these records, which can be time consuming. The calculation program compiles these records and uses them as the base for its calculations. The compiled records are in fact embedded in the CALC6DB.OVR and CALC6DC.OVR files, which means that the program does not need access to the file containing the pension records during its execution. It combines this information with the input from the parameter and participant files, and calculates the values discussed in the previous section.

The calculation program itself is even more involved. It is produced in two steps. First, for each pension record, a program translates that record into a subroutine which spells out in computer code exactly how to calculate the value of the pension, given the values in the parameter and participant files. Next, the defined benefit subroutines are combined into a single program, along with other subroutines which perform specialized functions. The defined contribution and combination routines are likewise gathered into a second program. The results are the CALC6DB and CALC6DC programs. It is possible, though not for the faint of heart (or computer skills) to modify the way the calculation program works, or to modify the pension records and incorporate the modified records into these programs. The program (even exclusive of the subroutines for the pensions) contains many thousands of lines of PASCAL code, and ascertaining how it works is a major undertaking. As a result, most users will not find that it is worthwhile to try to make changes in the program.

For users who want to obtain further information about the pension records, questionnaires, or codebooks, or who need to make modifications to the program, please contact the Health and Retirement Study at the Institute for Social Research, University of Michigan.

Appendix Table 1
Valid Ranges of Parameter File Values

	Valid Range		Recode
	Minimum	Maximum	Value
Date of survey observations	1980.00	2000.00	1992.00
Date for present value calculations	1980.00	2000.00	1992.00
Inflation rate	- 25.00	25.00	0.00
Real interest rate	0.01	20.00	1.00
Real salary and wage growth rate	- 20.00	20.00	0.00
Real Social Security wage base growth rate	- 20.00	20.00	0.00
Quit age, minimum	16.00	80.00	16.00
Quit age, maximum	16.00	80.00	80.00
Quit date, minimum	1925.00	2075.00	1925.00
Quit date, maximum	1925.00	2075.00	2075.00
Maximum benefit, age 55	0.00	500000.00	500000.00
Maximum benefit, age 65	0.00	500000.00	500000.00
Profit sharing contribution percentage (1)	0.00	30.00	5.00
Forfeiture contribution percentage (1)	0.00	10.00	0.50
Profit sharing contribution percentage (2)	0.00	30.00	5.00
Forfeiture contribution percentage (2)	0.00	10.00	0.50
Profit sharing contribution percentage (3)	0.00	30.00	5.00
Forfeiture contribution percentage (3)	0.00	10.00	0.50
Discretionary matching percentage (1)	0.00	200.00	50.00
Discretionary matching percentage (2)	0.00	200.00	50.00
Discretionary matching percentage (3)	0.00	200.00	50.00
Other unspecified employer contrib %	0.00	200.00	0.00
Unspecified mandatory employee contrib %	0.00	200.00	0.00
Voluntary contribution percentage, minimum	0.00	30.00	0.00
Voluntary contribution percentage, maximum	0.00	30.00	0.00
Unspecified matching % for mandatory contrib	0.00	200.00	100.00
Unspecified matching % for voluntary contrib	0.00	200.00	100.00
Annuity type	1.00	3.00	1.00
Percent of payment continued	0.00	100.00	50.00
New vesting switch	0.00	1.00	0.00
Years for vesting	0.00	25.00	5.00
Mandatory retirement flag	0.00	1.00	0.00
Pension type	0.00	2.00	0.00
First year annuity flag	0.00	1.00	0.00
Wage history flag	0.00	1.00	0.00
Adjust dollar amounts for inflation flag	0.00	1.00	0.00
COLA (for non-COLA plans) flag	0.00	1.00	0.00
Percent of COLA	0.00	100.00	0.00
Mortality table flag	0.00	1.00	0.00
Include past benefits flag	0.00	1.00	0.00

Appendix Table 2
Valid ranges for the Participant File Values

	Valid Range		Recode
	Minimum	Maximum	Value
Birth date of spouse	0.00	2000.00	0.00
Gender	1.00	2.00	1.00
Birth date	1880.00	1970.00	1950.00
Hire date	Birth date + 16	Birth date + 80	Birth date + 20
Date of separation	Hire date	Birth date + 80	Birth date + 65
Work hours per year	1.00	5200.00	2080.00
Annual earnings	1.00	8000000.00	25000.00
Differential wage growth rate	-20.00	20.00	0.00
Voluntary contribution percent	0.00	30.00	0.00
Linear wage growth (Beta1)	0.00	1.00	0.00
Quadratic wage growth (Beta2)	-1.00	1.00	0.00