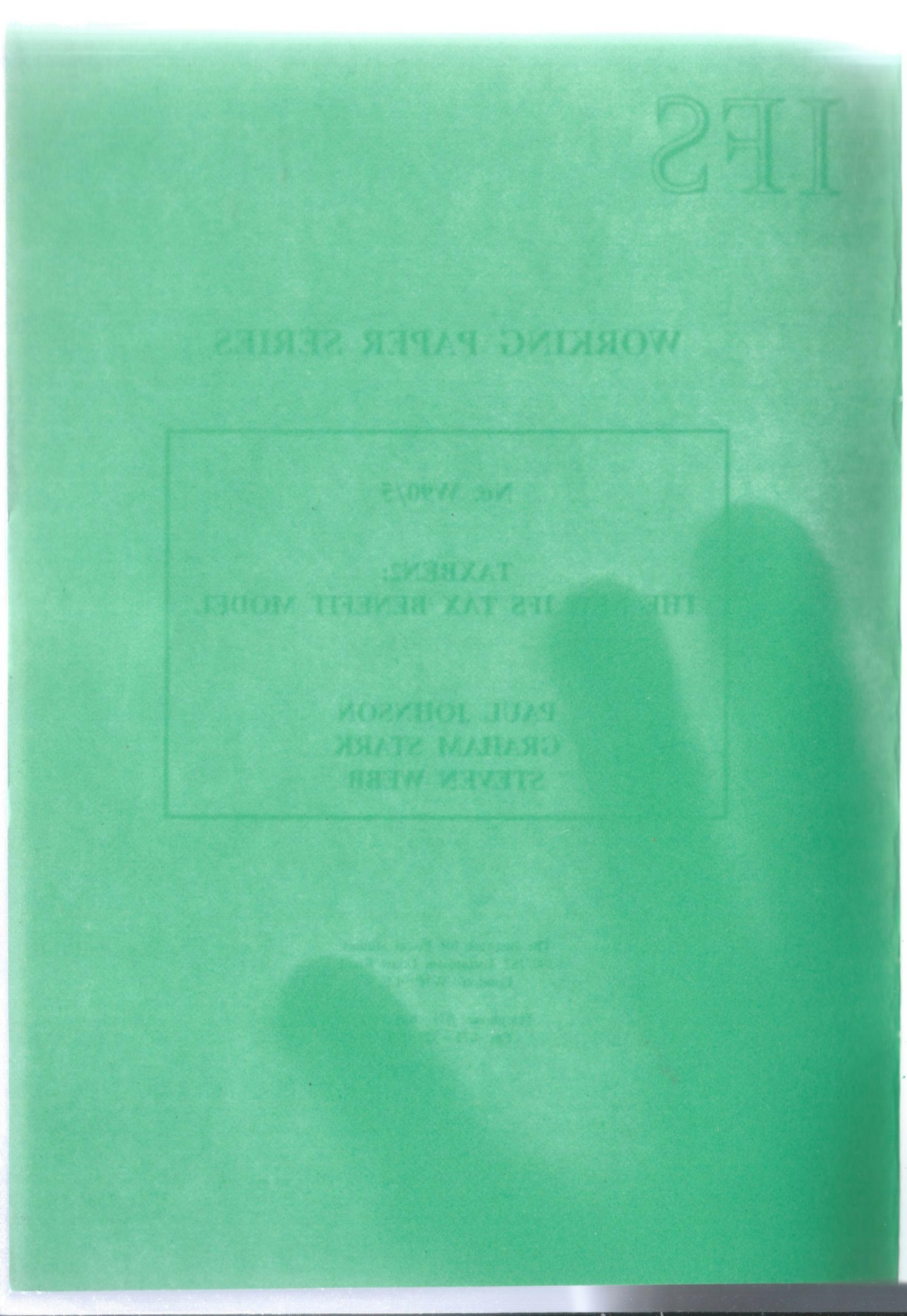


IFS Working Paper W90/5

TAXBEN2: THE NEW IFS TAX BENEFIT MODEL

Paul Johnson
Graham Stark
Steven Webb

May 1990

The financial assistance of the Economic and Social Research Council is gratefully acknowledged.

Contents

1 INTRODUCTION

A tax and benefit model is a computer program that calculates the effects of possible changes in the fiscal system on a sample of families. It allows inferences to be made about both the aggregate effects of tax and benefit changes, and the effect on specific groups. Tax and benefit modelling is playing an increasingly important role in policy analysis in the UK, both inside and outside the government. Correspondingly, tax and benefit models are rapidly becoming more sophisticated.

There has been a tax and benefit model at IFS for a number of years, and it has formed an essential part of much of our analysis of reforms to the tax and benefit system. The last major revision of the model occurred five years ago and since then there have been major advances in both the computer hardware and software available to us. Furthermore, through many years of microcomputer-based modelling a good deal of new expertise has been acquired in, for example, the handling of large data sets and the structuring of programs. These factors combined to convince us that the writing of a completely new model would be a worthwhile exercise, affording us the opportunity to improve both our use of the data and the structure, speed and efficiency of the model.

This paper begins with a brief discussion of the purpose of tax benefit models. We then go on to outline in some detail the nature of the data set used, and the way in which we make efficient use of it. Finally, in the section that forms the bulk of this paper we concentrate on the way in which benefit entitlements and tax liabilities are assessed for every household in the survey data used, and how these calculations are then used to examine the effects of a reform on the population as a whole.

1.1 The Purpose of Tax Benefit Models

During the 1980s, there have been extensive changes to the UK tax and benefit system. The top rate of income tax has fallen from 83% in 1979 (98% for investment incomes) to 40% today. The rate of VAT has nearly doubled. The structure of means-tested benefits was completely revised in 1988. A new system of independent taxation of personal income was introduced in 1990. There were also countless proposals which never came to pass. The Government had made two previous attempts at personal tax reform, for instance, and all the opposition parties would have made radical alterations to the tax and benefit system.

Clearly, we need to know not only the aggregate costs of such changes, but also the impact on individual families. Frequently this is done by showing the effect on a "typical" family. British newspapers and TV are addicted to using examples like this. But it can be highly misleading. The typical family with a working husband, non-working wife and two children, for instance, in fact represents only about 1 family in 20 in Britain today, and single-earner couples with two children themselves differ widely in income, expenditure patterns, type of tenure and so forth. It is usually possible to prove almost anything with a well chosen "typical" family. Nowadays the tax and benefit system is so complicated and family circumstances are so diverse that it is impossible not to be surprised by some of the consequences of all but the simplest possible change.

This is where the strength of tax and benefit models lies. If the effects of a reform can be calculated not on a few example households but on a large, representative sample of families, something informative about both the aggregate and the specific effects of possible reforms can be said, and any anomalies identified. For this we need a data source containing enough information on incomes and family

circumstances to perform the calculations that the benefit authorities and tax inspectors perform, and which is representative of the population as a whole. For the UK, such a data source exists: the Family Expenditure Survey (FES).

Much of the work done with the new model has been based on the 1984 FES, but the model has recently been adapted to use the 1987 FES. Clearly the more up-to-date the data on which the model is based the more confident we can be about the results. However, comparing output using data from these two years shows that the results we derived using 1984 data are very similar to those derived from 1987 data.

The new model has already been used in an extensive analysis of the effects of tax and benefit changes over the past ten years (Johnson and Stark, 1989) making use of its ability to run comparisons between several systems at once. This follows numerous studies on a wide variety of subjects for which the previous model was used.

1.2 Structure

The model is divided into two completely separate programs. The first is written in Pascal and the second in Modula – 2, a Pascal extension with more powerful features. By contrast, the previous version was written in FORTRAN – 77, a language mainly intended for numerical analysis. Compared with FORTRAN, Pascal and Modula – 2 offer a richer variety of data types, less scope for error and more easily readable code.

The first program is used to create the data file used in calculations by the main part of the model. It takes as input a complete "raw" FES data set and produces as output a compact file containing most of the information on personal income and characteristics contained in the raw data, plus some extra derived variables added for use in the main program. Users of the model need never see or know of this process since the data set is created only once for any year of FES data, and then used by the model.

The second program, the model proper, makes use of the created data set to do all the required calculations on the tax benefit system, the results of which it can then write out in tabular form.

1.3 The Family Expenditure Survey

As mentioned above, the IFS model uses data drawn from the FES.[1] The initial intention of the survey was to gather information on household expenditure which could be used to calculate the weights to be used in the construction of the retail price index. This extremely detailed information on the expenditure of the members of each household forms a large part of the data. Also included in the survey are household characteristics such as tenure type, region, etc. and individual characteristics such as work status, occupation, age, education, sex and marital status. Most valuable for our purposes, the FES contains detailed information on incomes. Thus it provides much of the information required to calculate tax liabilities and benefit entitlements, as well as providing information on actual tax payments and benefit receipts. There are, however, some shortcomings in the data.

1 See Kemsley, Redpath and Holmes (1980) for a detailed description of the sampling frame and other survey procedures.

Perhaps most important is the fact that the FES is a voluntary survey with a response rate of around 70%. (See Kemsley, Redpath and Holmes (1980) and Atkinson and Micklewright (1983) for further analysis.) Furthermore, some of the answers given seem to be unrepresentative of the population as a whole – for example expenditure on alcohol is seriously under-reported – and some information may not be available to the respondent – for example the split in mortgage payments between the interest and capital components. A final problem is that some of the information we would like for our calculations, for example a record of National Insurance contributions made in earlier years, simply is not included in the survey.

In the text that follows we will be referring to FES codes. These are simply a series of code numbers and letters which are assigned to each value recorded in the FES. For example the recorded value of the individual's last net wage is coded as personal income 303. Making use of these codes allows us unambiguously to identify the recorded values to which we are referring. Exact definitions of the codes are published annually along with the FES itself.

2 CREATING THE DATA SUBSET

As discussed above, our first task is to convert a raw FES data set into a form suitable for use with our model. The "raw" data set we use is from the IFS library of micro-computer-readable FES data sets. This library is held on laser disk and spans the years 1968 to 1988 inclusive.

There are four important units used in creating the data set:

1) The "definitions" unit: This simply contains record definitions of everything in the original FES data set, allowing any information in the FES to be used.

2) The "read" unit: This reads in all the information from the original data set, which allows the creation of the household records which contain a complete description of each FES household. It also allocates people to tax units and fixes some of the problems caused by coding errors in the FES.

3) The "create" unit: This is where the derived variables used in the model are created from the original FES data.

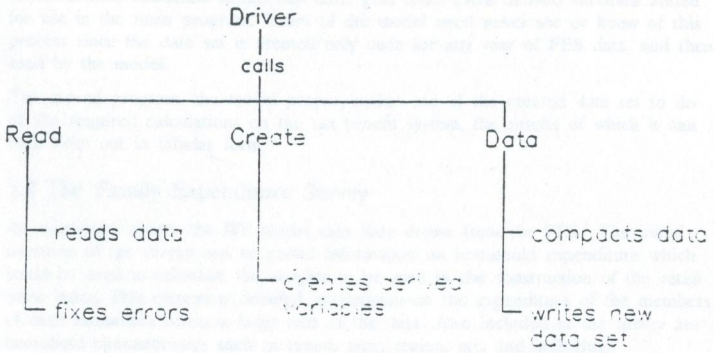
4) The "data" unit: This unit compacts the data and creates the file used by the model.

5) The "driver" unit: This calls the other units and runs the program.

A simplified idea of the structure is shown in figure 1.

Figure 1

Structure of the Data Creation Program



2.1 The "Definitions" Unit

This unit contains definitions of every variable in the FES. The definitions are arranged in a record structure, in such a way that one record, the household record, contains all the information recorded in the FES for each household. The household record is itself made up of tax unit records and certain variables, such as tenure type, which are relevant only at the household level. The tax unit records are themselves made up of adult records and tax unit level variables. The adult

records, in turn, are made up of a number of sub-records such as a main employment record and an education record. These sub-records contain the actual elements found in the FES such as last wage, income tax paid, level of education reached, etc.

This use of records is extremely helpful when dealing with a data set as large and complex as the FES and in performing the complicated manipulations required to simulate the tax and benefit system. They allow us to group similar things together. For instance variables relevant only at the tax unit level can be kept separate from those relevant at either the adult or household level; information on income from employment can be stored in a different record from that on investment incomes. In such a fashion variables are more easily recognised and simple mistakes avoided.

Perhaps the easiest way to understand this structure is to illustrate with an example. Take a simple element like usual gross income from main employment. This is an element of the main employment record, which is itself an element of the adult record. The adult record is an element of the tax unit record which is an element of the household record.

Thus the last net wage for the head of the first tax unit would be written in full like this:

HH.TU[1].Ad[Head].MainEmp.LastNetWage

(although Pascal allows this to be abbreviated).

2.2 The "Read" Unit

As explained above the main function of the "read" unit is simply to read in all the values from the original FES data set for the creation of the household records, defined in the "definition" unit, for every household. First, however, we should examine the two other main procedures found in this unit.

The information in the FES is based on diary records kept by each household. All these diary records are then keyed into a computer to provide the data sets we use. Clearly, in an operation of this scale there is scope for error. Although the errors we found were relatively few, we decided that where it was relatively clear what the error was, we would attempt to correct it rather than simply rejecting the household. Most of these errors concern personal characteristics codes ("A" codes). For instance, some single women are coded as being wife of head of an income unit (A009 = 2), even though no husband exists. These women have been made head of their own tax unit. Also, more than one person is sometimes allocated to the headship of a particular income unit. Where this occurs the second person is put in their own income unit.

The second important procedure in this unit allocates people to tax units. In the FES, households are divided up into income units, but in a slightly different way from that which we would want for the purposes of tax benefit modelling, in that all young dependants under 25 are allocated to their parents' income unit. For the purposes of determining tax and benefit liability and entitlement we allocate all those over 16 not still in approved full-time non-advanced further education (i.e. at school, Sixth Form College, FE College, etc.) to their own income unit. All other people are allocated to the income unit recorded by code A008 in the FES. The young dependants are allocated their own units, starting with the next available unit once all other people have been allocated to theirs.

Another complication is that foster children in the FES are not allocated to any income unit during coding. If they are 16 or under we allocate them as children to the first tax unit in apparent receipt of foster allowances or, if there are no such receipts, to the first tax unit. If over 16, we treat them as we do other under-25 dependants.

2.3 The "Create" Unit

As explained earlier this unit creates a number of derived variables for inclusion in the model's data set. These derived variables are required where the original data does not contain explicit information that is needed for the purposes of the model, but where the requisite information can be deduced from the information that is directly available.

Firstly, a number of income variables are defined which are used extensively in the model. Thus, for instance, last gross main income is made up as follows (the meanings of the variables are hopefully self-evident):

Inc [LastGrossMain] = MainEmp.LastNetWage + MainEmp.BonusesNet
+ MainDed.TaxDeduction + MainDed.NIContrib
+ MainDed.Superann + MainEarnedDeductions

where the prefix MainEmp indicates that the variable is an element of the main employment record, and MainDed indicates it is a member of our main deductions record. (The FES codes identifying each of these income variables and those used below are shown below.)

A current gross main income is also derived. If employment status (FES code A202) is greater than or equal to 6 (i.e. if the person is retired, unoccupied or unemployed), then this is set equal to zero. (Remember, these are all main *employment* incomes.) If the employment status is less than 4 (i.e. the person is self-employed or a part-time or full-time employee), then current gross main income is set equal to last gross main income.

The two remaining employment status values, 4 and 5, indicate an employee temporarily away from work. If such people are receiving pay while away then current gross main income is again set equal to last gross income, otherwise it is set equal to zero.

(FES variables used above:

Variable	Code Number (all personal incomes)
LastNetWage	303
BonusesNet	330
TaxDeduction	305
NIContrib	306
Superann	318
MainEarnedDeductions	422 + 424 + 425 + 426 + 427 + 428 + 429)

The following procedures concern the level of entitlement to National Insurance benefits. The FES does not provide information on an individual's lifetime record of National Insurance contributions which is in principle necessary to calculate entitlement to a range of National Insurance benefits. In the past, the approach adopted by the IFS model, and many other tax benefit models, has been to use general rules of thumb to fill the gaps in incomplete data. These have included simple uprating of actual receipts data, or the random allocation of benefit positions to known recipients subject to certain known overall distributions.

Such approaches, however, impair the accuracy of the model both in reflecting the current situation and in the modelling of a reform. Simplifying assumptions not based directly on the data may introduce biases because they will typically not be sufficiently tailored to reflect individual circumstances.

It should, however, be possible to infer from the actual data the relevant information needed to determine entitlement to various National Insurance benefits. Such a course is taken in this unit to determine entitlement to the following: attendance allowance, invalidity benefit, retirement pensions, widow's benefit and mobility allowance. The different possible entitlements may reflect differences in National Insurance contribution records, as with retirement persons, or in supposed needs, as with attendance allowance. We will go through some of the procedures here to show the basis on which we are working.

2.3.1 Entitlement to attendance allowance

Attendance allowance is a non-means-tested benefit available to those who are severely disabled and who require attention or supervision from another person because of their disability. It is payable at two rates – a higher rate for those who require attention both day and night, and a lower rate for all other cases. To determine entitlement we simply compare receipt with the rates payable at the time. Thus if receipt equals the higher rate (either before or after the annual uprating which takes place during the FES year) then we set entitlement equal to higher rate entitlement. Conversely, if receipt equals the lower rate then entitlement is set equal to lower rate entitlement. This entitlement information is recorded and included in the file created by the "data" unit.

The importance of doing this in such a way is the increased flexibility it allows us. For instance, if the lower and higher rates were to be differentially uprated it would be a simple matter for us to assign the correct rates to the correct people. This is something that could not be done if the only information concerned whether any attendance allowance was being received.

2.3.2 Entitlement to invalidity benefit

Invalidity benefit (IVB) is a contributory benefit payable to any adult who is incapacitated from work for a period of more than 28 weeks and who has made the necessary National Insurance contributions. The modelling of this benefit is complicated by the following factors:

1) It comprises two elements: an invalidity pension similar in structure to the retirement pension, and an invalidity allowance the value of which varies according to the age of the recipient when his incapacity began. There are three rates of invalidity allowance, the highest rate going to those whose incapacity began before they had reached the age of 40.

2) Additions are payable in respect of dependants, subject to certain earnings rules.

3) Invalidity pension may be paid at a reduced rate if insufficient National Insurance contributions have been made. This reduced rate applies also to any additions for a dependent spouse, but not to additions for dependent children.

4) Invalidity pension may be supplemented by an earnings-related component.

In the program we try to determine entitlement from receipts data. We take the following steps:

1) We calculate earned income for husband and wife, needed for use in determining the level of spouse additions.

2) If someone in the tax unit is receiving some IVB, we determine whether it is husband or wife.

3) Knowing how much, if anything, is being earned by the spouse, we determine the size of the dependent spouse's addition, which depends on the relationship between the earnings and an upper and lower earnings limit.

4) Modelled invalidity pension becomes the basic invalidity pension plus modelled dependent spouse's addition plus child addition times the number of children.

5) We now come on to determine the entitlement to invalidity allowance. If the recipient is younger than the lowest relevant age (40 in 1984) then he clearly receives the highest allowance. If, however, he is older than this and is recorded as having been on IVB for 52 weeks (this being the number of weeks of receipt during the previous year) it is not immediately clear to which level of allowance he is entitled. For instance, if he is over the highest age (60 in 1984) he may be entitled to no invalidity allowance or to the low, middle or high level depending upon his age when he first become incapacitated.

6) Those who could not be matched for certain in (5) have their receipts compared with their modelled invalidity pension plus each of the possible allowances to which they could be entitled. Where matching occurs, entitlement is recorded as being for the level of allowance which brings about this matching.

Where matching has still not occurred a valid allowance is ascribed randomly. Then for those whose invalidity benefit receipt is higher than their modelled entitlement we assume entitlement to SERPS equal to the difference between receipt and modelled benefit. Where benefit is lower than modelled entitlement we assume an incomplete contributions record and the contributions ratio is calculated as basic invalidity benefit receipt less modelled child additions less modelled invalidity allowance, divided by the basic modelled invalidity pension (including spouse's additions but excluding child additions).

2.3.3 Entitlement to retirement pensions

We shall not describe this procedure in as much detail as the last since the basic principles are fairly clear. We start off, as usual, by calculating what we think entitlement should be on the basis of the level of the basic pension and dependants' and age additions. If receipt matches this modelled entitlement then the entitlement is recorded. If not then receipt is compared with a number of possible levels of supplementary benefit receipt, since receipt of pensions and SB is often confused in the FES. If actual receipt matches a possible level then basic SB entitlement is given. Otherwise if receipt is greater than basic modelled entitlement then we assume entitlement to SERPS, the entitlement being equal to the difference between the benefit received and the benefit modelled. Finally if receipt is less than the modelled pension we assume this is the result of an incomplete contributions

record. Because only the basic allowance can be scaled down, any child and age additions are taken away from the actual receipt and a contributions rate is calculated as this adjusted receipt divided by the basic modelled pension.

2.3.4 Entitlement to widow's benefit

Somewhat more complex is the calculation of entitlement to widow's benefit. The process is similar to that for attendance allowance in that a direct comparison is made between the receipt data and the various widow's benefit rates of the day. The first step is to calculate modelled entitlement under various assumptions – that they are receiving widow's allowance pre- or post-November (which is when the benefit rates changed), or that they are receiving widow's pension or widowed mother's allowance pre- or post-November. These modelled entitlements are then compared with actual receipts. If receipt is zero then there is no entitlement. If receipt exactly equals widow's allowance plus any dependants' additions plus age addition, then entitlement is for widow's payment plus the relevant widow's benefit. Equally if receipt matches widow's pension or widowed mother's allowance and any dependants' additions, then it is to this that entitlement is set. Otherwise if receipt is greater than modelled widow's pension we assume entitlement to SERPS and if it is less we assume an incomplete contributions record, much as with retirement pensions.

2.3.5 Other benefits

We determine entitlement to unemployment benefit by looking at whether it is currently being received, entitlement to child benefit by looking for the existence of children, and entitlement to statutory sick pay and sickness benefit by looking at the length of time away from work and the reasons for this absence. We do not attempt anything more sophisticated with unemployment benefits and sickness benefits partly because the data for these short-term benefits is relatively unreliable.

2.3.6 National Insurance contributions

The main point of having a National Insurance contributions procedure here is to determine whether employees are contracted in or out of the State Earnings-Related Pension Scheme, SERPS, and to determine which of the self-employed are liable to class 2 contributions. (It is possible to contract out of SERPS, in which case National Insurance contributions are levied at a rate 2% less than the normal rate, between the floor and the ceiling.)

For employees the first step is to work out what each person's NI would be for each job, assuming first that they are contracted in to SERPS and then that they are contracted out. For women a calculation is also performed to determine what their payment would be if they were paying contributions at the special reduced rate. Each of these is compared with the recorded payment for main and subsidiary employments and the calculated payment nearest to the recorded payment is taken to indicate whether the person in question is contracted in or out of SERPS and whether they are paying contributions at the reduced rate for women. If somebody's earnings are under the floor then he is assumed not to be contracted out.

As far as the self-employed are concerned – that is, those with any recorded self-employment income – all are made potentially liable to class 2 NI contributions, with the exception of those with no recorded payment whose income is over the class 2 floor and who are classed as married women, old age pensioners or widows.

2.3.7 Other

Also created in this unit are usual and current gross incomes for tax units and for households. Also at the household level, rates are calculated as the product of rate poundage and rateable value and similar calculations are performed for sewerage, water and environmental rates.

2.4 The "Data" Unit

As mentioned earlier this is the unit which writes out a subset of the information which has been read in or created, into the file which is to be used by the model. However, it does not merely write it out in a format, and therefore size, similar to the data set originally read in, but compresses it into a file of just 600K.

We employ a number of techniques to do this, including run-coding (so as to eliminate zeros), condensing money values into 2 bytes instead of the usual 4 or 8, and compressing several categorical variables (such as region, tenure, etc.) into single binary integers.

3 THE MAIN MODEL

As mentioned earlier the main part of the model where all the calculations are done is written in Modula-2. In this section we will first give a very brief outline of the structure of the model as a whole before going on to take a much more detailed look at each unit in turn.

The main model is comprised of three parts: a "set-up" routine which sets up the record structure and reads in the data and parameters; a calculation routine which calculates the tax and benefit liabilities and entitlements for every household in the FES; and an output routine which can be used to write out a wide variety of tables showing effects by household or tax unit, by income range or family type. The driver routine, which is very short, calls all these modules and runs the model.

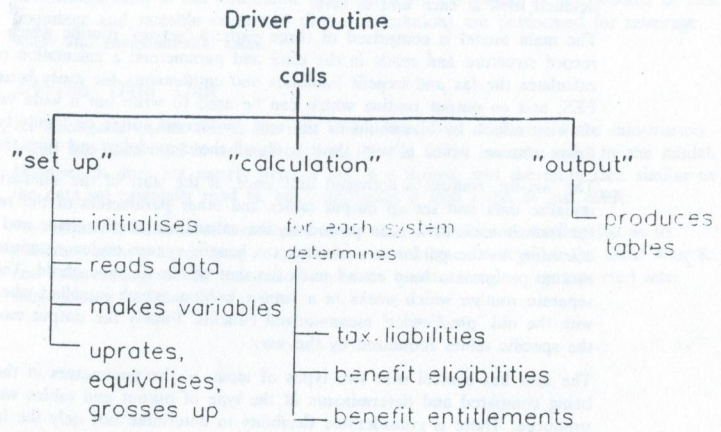
The "set-up" routine is activated only once, at the start of the simulation, to initialise data and set up output tables and other parameters of the run. The data for each household is then passed to the calculation routine once and is processed according to the parameters of each tax benefit system under examination. The routine performs a loop round each tax unit within the household. There is a separate routine which works in a similar fashion, which is called when working with the old "pre-Fowler" means-tested benefits. Finally the output module prints the specific tables requested by the user.

The user has control over two types of input – the parameters of the systems being compared and determinants of the type of output and tables which will be produced. There is considerable flexibility to determine not only the levels of benefits and rates of tax in the systems being compared but also, for instance, whether the benefit structure is as it was before or after the 1988 "Fowler Reforms" to the social security system, the structure of husband's and wife's taxation and the structure of National Insurance contributions. The model can be made to print out detailed information on one, or a number, of specific households, or give summary tables for the population as a whole. These summary tables can take several different forms as determined by the user, showing the distribution of gains and losses from a reform according to a number of different criteria such as family type, income decile and net or gross income.

A much simplified outline of the structure of the model is shown in figure 2.

Figure 2

Structure of the Model



3.1 The "Set-Up" Units

There are two of these, but one is simply a definitions file, very similar to the definitions unit in the data creation half of the model. In a similar fashion it sets up and defines the record structure for the FES information to be used in the rest of the model. But as well as a record structure for the FES data, it defines a record structure for the parameters of the tax and benefit systems that are read in and also for the output tables to be used.

The second is the unit which reads in the created data set, as well as containing procedures to make the household types we use in the output, to make the grossing-up factors for each tax unit type, to update incomes etc. as desired, to equivalise incomes according to household size, to read in the parameters and some others.

The reading routine is, in a sense, the reverse of the "data" routine which wrote out the data set initially. One procedure reads an index file which shows where each household is in the main data set. This makes it possible to seek the byte where the household record begins. It is this use of an index file which allows us the option of going straight to specific households and writing out information on them individually. Another reads the series of 1s and 0s indicating the existence of non-zero values for the income variables. When it comes across a 1 it goes to the relevant position in the main data set to read the income, otherwise the income is set at zero. Finally there is a procedure designed to read the compacted data set, reading the relevant bits from each 16-bit integer.

3.1.1 Making family type

This procedure is used to define the types of family used in the output tables to show the effects of reforms on various family types. The precise definitions used are set out below.

Single-parent family: Any tax unit with just one adult and one or more children. Their definition as single-parent families takes precedence over their employment position in determining family type.

Pensioner: Those whose employment status is recorded as retired or who are over the state pension age with employment status recorded as unoccupied. A married couple is classified as couple pensioners if, and only if, the head (i.e. the husband) is classified as a pensioner, irrespective of the employment position of his wife.

Unemployed: Included among the unemployed are those seeking work and the sick, including those sick who are not seeking work. We do not count as unemployed those who are healthy and not seeking work.

Employed: These are divided into single and married, the married being divided into single-earner and two-earner couples with and without children.

Others: Comprised mainly of students and those not in work and not looking for work.

When we want to classify by household type as opposed to family type, the situation is complicated by the existence of multiple tax unit households (i.e. households which contain more than one family or tax unit). We simply refer to these households as multiple tax unit households, rather than the alternative of assigning them to the same group as the first tax unit in the household. This prevents many confusions that may otherwise occur in the results.

3.1.2 Grossing up

The FES is a broadly representative sample of the British population. However, being a survey to which response is voluntary it is not perfect and some groups are less adequately represented than others. For instance married couples with children are over-represented, the very old significantly under-represented.

There may be circumstances in which it is desirable to "gross up" the sample that we have so that it is more representative of the population as a whole and so that we can see the total effects of changes in terms of their cost. It is not, however, always necessary to do this as determining the effects of a change on the actual sample that we have is interesting in its own right. The model allows the user to determine whether or not grossing up occurs.

The problem in deciding how, precisely, grossing up should be carried out is that there may be sampling errors in more than one dimension. That is, not only is it possible that certain family types are under-represented, but there may also be uneven representation of income groups or among those in different types of employment, and grossing up to correct for one of these need not necessarily correct for any of the others. One ambitious attempt to tackle this problem is proposed by Atkinson, Gomulka and Sutherland (1988). Weights for each household were chosen so as to guarantee that there would be the correct numbers of each family type, income group and tenure type when the sample was grossed up. From the set of all possible weights that would achieve this, they chose those which minimised a measure of the dispersion between individual weights. The problem was thus one of constrained optimisation. The difficulty with this is that the weights that

are produced vary very widely (from 800 to 21,000), and although the aggregate performance may be better, this is at the cost of likely distortions at the disaggregated level. And although this method will, for example, give the right number of single parents, the right number of council tenants and the right number of people earning below £100 p.w., it does not go anywhere near giving the right number of single parents who are council tenants earning less than £100 p.w.. Thus, we decided to use the simpler method of just grossing up for family type, firstly because this is common practice in tax benefit models and secondly because it is easily understood.

We determined the grossing up numbers quite simply by comparing the numbers of each family type in the 1984 FES with the numbers in the population as a whole, and dividing the latter by the former. The grossing-up factors we use are the following:

Family type				Grossing-up factor
Marital status	Age of head	Sex	No. of children	
Married	< 65	–	0	2834.3
Married	< 65	–	1	2796.0
Married	< 65	–	2	2691.1
Married	< 65	–	3	2685.6
Married	< 65	–	4+	2578.9
Married	65 – 69	–	–	3195.0
Married	69 – 74	–	–	3405.4
Married	75 – 99	–	–	3625.0
Single	–	–	1+	2907.0
Single	< 65	Male	0	3991.8
Single	65 – 69	Male	0	2786.7
Single	70 – 74	Male	0	2706.7
Single	75 – 99	Male	0	2971.6
Single	< 60	Female	0	3607.3
Single	60 – 64	Female	0	2205.7
Single	65 – 69	Female	0	3413.8
Single	70 – 74	Female	0	2703.6
Single	75 – 99	Female	0	4144.2

Thus, in the grossing-up routine, each tax unit is grossed up by (i.e. multiplied by) the relevant number above, according to its family type. This means effectively that each tax unit represents the number of such tax units in the population as a whole given by its grossing-up factor. Beyond this, however, we do make some attempt in a separate procedure to adjust for the fact that those on very high incomes are notoriously under-represented in the FES by multiplying the grossed-up number of all those in receipt of income of more than £20,000 p.a. by 1.40.

As far as the income data itself is concerned, this is generally accurate. However, both self-employment income and investment incomes are understated in the FES as compared with other published figures. In grossing up we therefore multiply self-employment incomes by 1.16 and investment incomes by 1.20. These numbers come from Atkinson (1983).

3.1.3 Equivalising incomes

When comparing the incomes of tax units we may wish to take into account the differing numbers of people within those tax units. For instance, a single person earning £10,000 p.a. is clearly better off than a married couple with a combined income of £10,000 p.a. However, he is not necessarily twice as well off since it doesn't typically take twice as much money to feed, house, etc. a couple as a single person. In estimating how well off someone's income makes them it is also necessary to take into account the number of children they have to support, whether or not they are householders, and the number of other people living in the household.

We are able to use a range of equivalence scales such as the McClements scales (pre-housing costs) developed at the DHSS in the mid-1970s, based on expenditure data from the 1971 and 1972 FESs. Although these may be somewhat out of date now, they are still in regular use at the DSS. These scales are shown below as an illustrative example. (It is important to note once again that this is a procedure which is called only if the user desires it.)

1) *If head of household is married*

Number of additional adults in household	Equivalence value
0	1
1	1.42
2	1.78
>2 (where number of adults = n)	1.78 + 0.36 * (n – 2)

2) *If head of household is single*

0	0.61
1	1.07
2	1.49
>2	1.49 + 0.36 * (n – 2)

3) *Total number of children in household*

Age of child	Value for each child
< 2	0.09
2 – 4	0.18
5 – 7	0.21
8 – 10	0.23
11 – 12	0.25
13 – 15	0.27
> 15	0.36