CREATING A BASEFILE FOR STINMOD

Kerrie Bremner, Gillian Beer, Rachel Lloyd and Simon Lambert

Technical Paper no. 27
June 2002
About NATSEM

The National Centre for Social and Economic Modelling was established on 1 January 1993, and supports its activities through research grants, commissioned research and longer term contracts for model maintenance and development with the federal departments of Family and Community Services, Health and Aged Care, and Education, Training and Youth Affairs.

NATSEM aims to be a key contributor to social and economic policy debate and analysis by developing models of the highest quality, undertaking independent and impartial research, and supplying valued consultancy services.

Policy changes often have to be made without sufficient information about either the current environment or the consequences of change. NATSEM specialises in analysing data and producing models so that decision makers have the best possible quantitative information on which to base their decisions.

NATSEM has an international reputation as a centre of excellence for analysing microdata and constructing microsimulation models. Such data and models commence with the records of real (but unidentifiable) Australians. Analysis typically begins by looking at either the characteristics or the impact of a policy change on an individual household, building up to the bigger picture by looking at many individual cases through the use of large datasets.

It must be emphasised that NATSEM does not have views on policy. All opinions are the authors’ own and are not necessarily shared by NATSEM.

Director: Ann Harding
CREATING A BASEFILE
FOR STINMOD

Kerrie Bremner, Gillian Beer,
Rachel Lloyd and Simon Lambert

Technical Paper no. 27
June 2002
Title: Creating a Basefile for STINMOD
Author(s): Kerrie Bremner, Gillian Beer, Rachel Lloyd and Simon Lambert
Series: Technical Paper no. 27
Key words: microsimulation; modelling; base population
Abstract

STINMOD is NATSEM’s static microsimulation model of income taxes and cash transfers. It is publicly available, runs on a personal computer and can be accessed via a user-friendly interface. STINMOD can be used to analyse the distributional impact of current tax-transfer policy or to estimate both the fiscal and distributional impacts of policy reform.

STINMOD applies the rules of income tax and government cash transfer programs to a database of income units representing the Australian population. This database, referred to here as the basefile, is a major component of STINMOD and the methodology underlying the creation of the basefile is the subject of this paper.

The paper provides an overview of the major changes to the STINMOD basefile that have taken place since earlier papers on the basefile methodology were written. It details the techniques used in each of the three main steps that are required to build a STINMOD basefile.
Author note

Kerrie Bremner and Rachel Lloyd are Senior Research Fellows at NATSEM. Gillian Beer is a Principal Research Fellow at NATSEM and Simon Lambert was a Principal Research Fellow at NATSEM when part of this paper was written.

Acknowledgments

The authors thank Marty Robinson for his work on the Outyears documentation, on which some of the information in this paper is based.

General caveat

NATSEM research findings are generally based on estimated characteristics of the population. Such estimates are usually derived from the application of microsimulation modelling techniques to microdata based on sample surveys.

These estimates may be different from the actual characteristics of the population because of sampling and nonsampling errors in the microdata and because of the assumptions underlying the modelling techniques.

The microdata do not contain any information that enables identification of the individuals or families to which they refer.
Contents

Abstract iii
Author note iv
Acknowledgments iv
General caveat iv
1 Introduction 1
2 Overview of changes 2
  2.1 Currency of basefile 2
  2.2 Calculation of weights 3
  2.3 Source of microdata 3
  2.4 Imputation of family links 5
  2.5 Imputation of workforce independence 6
  2.6 Basefile structure 6
  2.7 Parameters 7
  2.8 Accessibility 8
  2.9 Adjustment of incomes 8
3 Steps for creating a STINMOD basefile 9
  3.1 Calculation of weights 9
  3.2 Adjustment of incomes 17
  3.3 Imputations 20
4 Changing assumptions and generating a new basefile for STINMOD 24
5 Linking the STINMOD basefile to the original SIHC microdata 24
6 Future directions 26
References 28
1 Introduction

STINMOD is NATSEM’s static microsimulation model of income taxes and cash transfers. It is publicly available, runs on a personal computer and can be accessed via a user-friendly interface. STINMOD can be used to analyse the distributional impact of current tax–transfer policy or to estimate both the fiscal and distributional impacts of policy reform (see Lambert et al. 1994 for more information on STINMOD).

There are two major components to STINMOD. The first is the suite of entitlement modules. These modules simulate the policy rules of the major federal tax and transfer programs, including eligibility, entitlement and interaction. The rules are translated into computer code using SAS software.

The rules are then applied to a database of income units1 representing the Australian population. This database, from here on referred to as the basefile, is the second major component of STINMOD. The methodology underlying the creation of the basefile is the subject of this paper.

The first version of STINMOD was released in the middle of 1994. Around this time NATSEM published two papers that described the methodology used to create STINMOD’s basefile. One paper (Landt et al. 1994) describes in detail the approach used by NATSEM to calculate weights for individuals and income units in microdata sets destined for use in a microsimulation model. Such weights are used to create population estimates from a sample of the population.

The other paper (Percival 1994) describes the steps involved in creating the basefile for STINMOD-94A. It summarises the methodology used to calculate weights, describes the method for adjusting incomes to make them more contemporary, and outlines the imputations required to overcome omissions and limitations in the original microdata from which the STINMOD basefile was derived.

---

1 The Australian Bureau of Statistics defines an income unit as ‘one person or a group of related persons within a household, whose command over income is assumed to be shared. Income sharing is assumed to take place within married (registered or de facto) couples, and between parents and dependent children’ (ABS 2001).
Since that time there have been a number of significant changes in the way the STINMOD basefile is created. Although many aspects of the methodology remain unchanged, the current documentation no longer provides sufficient information for those who use STINMOD and want to understand how the basefile is created. While some technical notes have been produced, they have been aimed at experienced STINMOD users, particularly those who use the source code.

In response to this shortfall, this paper aims to summarise the current methodology underlying the STINMOD basefile. It should be read in conjunction with the papers by Landt et al. and Percival. The next section summarises the major changes that have occurred since the release of these two papers. In section 3, each of the three main steps involved in creating a STINMOD basefile is discussed in detail. The steps involved in creating a user-defined STINMOD basefile are outlined in section 4. Section 5 discusses how to link the STINMOD basefile back to the original microdata underlying the basefile. Finally, in section 6 there is some discussion of likely future directions in the STINMOD basefile methodology.

2 Overview of changes

2.1 Currency of basefile

Prior to the 00A version of STINMOD, the basefile was current at a single point in time. This point in time corresponded to the date of the most recently available Labour Force Survey (LFS) conducted by the Australian Bureau of Statistics (ABS). NATSEM obtains a detailed matrix file from the survey for May and November of each year and uses this in the reweighting process (described by Landt et al. 1994).

The 00A version was the first to have a series of basefiles, some of which are based on historical data and some of which are projected base datasets. That version, for example, has basefiles created for the December and June quarters of each financial year from 1998-99 to 2004-05.

Previous STINMOD basefiles were based on known benchmark data. The move to projected basefiles required the development of a projection methodology, which is described later. Development of the projection
component was largely a response to requests from users in federal government departments who needed to provide ‘outyear’ costings as part of the budget process. As a result, STINMOD is now an outyears model, with users able to choose the financial year they wish to analyse.2

2.2 Calculation of weights

Since the mid-1990s a different approach has been taken to calculating weights. Prior to this the reweighting process involved using the LFS matrix file to calculate weights for people in the microdata. This was done by a somewhat sophisticated ratio approach, described in detail by Landt et al. While this approach gave good population estimates for the characteristics contained in the LFS matrix, STINMOD outcomes did not always closely match administrative numbers for some of its transfer programs. With the availability of the new CALMAR reweighting software in 1996, an additional step was added to the reweighting process. This involves reweighting to administrative benchmarks after the LFS matrix reweighting has been completed. This process is explained in more detail in section 3.1.

2.3 Source of microdata

The first release of STINMOD, 94A, was based on microdata from the 1990 ABS Survey of Income and Housing Costs and Amenities. The survey was a sample of the private dwelling population. These data were used for two subsequent releases of STINMOD as no more recent data were available. With the release of the microdata from the 1993-94 ABS Household Expenditure Survey (HES), the HES was used as the basis of the 96A version of STINMOD.

Around this time the ABS changed the method it used to conduct its income survey. It moved from collecting information on a periodic basis to a continuous basis. The survey was renamed the Survey of Income and Housing Costs (SIHC). From the beginning of 1994-95 the ABS ran the income survey on a monthly basis, sampling from the group of people who were exiting the Labour Force Survey. Microdata were publicly released on a financial year basis with a smaller time lag than

---

2 The development of the outyears capacity in STINMOD stemmed from an outyears version of the model created by the federal Treasury.
for the surveys that were conducted periodically. SIHC data were released in 1994-95, 1995-96, 1996-97 and 1997-98.

The main benefit of the new survey approach for STINMOD was that it provided more current microdata, which did not have to be ‘aged’ over as long a period as the earlier microdata. Such ageing always introduces an element of error into the microdata. This error is almost certain to increase as the length of time over which the data are aged increases.

However, the new approach had some disadvantages. Most importantly, the sample size was much smaller than that of the previous periodic surveys. This creates problems when undertaking disaggregated analysis with STINMOD as it often leads to unreliable estimates based on only a few observations from the sample.

The sample size problem was overcome by combining the data from the two most recent SIHCs. The only major implication of this was the need to calculate and use different uprating factors when updating incomes. This is discussed in more detail in section 3. SIHC microdata have been combined in all versions of STINMOD since 98A except 01A, which was based on the 1998-99 HES.

The absence of people living in non-private dwellings — the institutionalised population — was addressed in the 99B version by incorporating microdata to represent this component of the population. This population is of interest to STINMOD because of the high incidence of cash transfers to this group. Two obvious examples are aged pensioners in nursing homes and students in halls of residence.

The institutionalised population in the STINMOD basefile is derived from the 1 per cent sample from the 1991 Census of Population and Housing. To construct this synthetic population, a regression-based approach was used to allocate individuals in the 1 per cent sample to particular types of institution. Additional income information was imputed to enable STINMOD to calculate transfer payments and tax liabilities. Lim and Percival (1999) give a detailed description of the creation of the institutionalised dataset in STINMOD.

Validation of the constructed dataset to available benchmark data has given positive outcomes, the major limitations of the dataset being a lack of information in specific areas, notably about rental payments and family relationships.
Including the institutionalised sample also had implications for the calculation of weights and for the uprating of incomes. These are both discussed in section 3.

For each version of STINMOD released to date, table 1 shows the microdata used to construct the basefile and whether the institutionalised population is included.

Table 1  **Microdata used for STINMOD basefiles and the inclusion of the institutionalised population**

<table>
<thead>
<tr>
<th>Version of STINMOD</th>
<th>Survey(s)</th>
<th>Institutionalised population</th>
</tr>
</thead>
<tbody>
<tr>
<td>94A</td>
<td>1990 ABS Survey of Income and Housing Costs and Amenities</td>
<td>No</td>
</tr>
<tr>
<td>94B</td>
<td>1990 ABS Survey of Income and Housing Costs and Amenities</td>
<td>No</td>
</tr>
<tr>
<td>95</td>
<td>1990 ABS Survey of Income and Housing Costs and Amenities</td>
<td>No</td>
</tr>
<tr>
<td>96A</td>
<td>1993-94 ABS Household Expenditure Survey</td>
<td>No</td>
</tr>
<tr>
<td>98A</td>
<td>1994-95 and 1995-96 ABS Surveys of Income and Housing Costs</td>
<td>No</td>
</tr>
<tr>
<td>99A</td>
<td>1994-95 and 1995-96 ABS Surveys of Income and Housing Costs</td>
<td>Yes</td>
</tr>
<tr>
<td>99B</td>
<td>1995-96 and 1996-97 ABS Surveys of Income and Housing Costs</td>
<td>Yes</td>
</tr>
<tr>
<td>00A</td>
<td>1995-96 and 1996-97 ABS Surveys of Income and Housing Costs</td>
<td>Yes</td>
</tr>
<tr>
<td>00B</td>
<td>1996-97 and 1997-98 ABS Surveys of Income and Housing Costs</td>
<td>Yes</td>
</tr>
<tr>
<td>01A</td>
<td>1998-99 ABS Household Expenditure Survey</td>
<td>No</td>
</tr>
<tr>
<td>01B</td>
<td>1996-97 and 1997-98 ABS Surveys of Income and Housing Costs</td>
<td>Yes</td>
</tr>
</tbody>
</table>

### 2.4 Imputation of family links

The ABS microdata do not contain any information about parental income and family structure for single income units living away from home and still considered dependants for cash transfer purposes. This information is vital for STINMOD, as it is needed to correctly determine Youth Allowance (formerly AUSTUDY) and family assistance outcomes. As a result, since the first release of STINMOD, information necessary to determine outcomes for away-from-home single income units has been imputed onto their records.

Prior to the release of version 99A, imputed information about parental and family characteristics was added to the away-from-home single income units and used to determine their AUSTUDY outcomes. Importantly, the imputations were in one direction only — from parents to the away-from-home single income unit. The AUSTUDY outcomes for away-from-home single income units did not affect the cash transfers for the
parental income units. For example, the away-from-home single units did not count as dependants for cash transfer calculations for their parents.

This approach changed to a two-way link with the release of version 99A. Each away-from-home single income unit deemed to be dependent on a parent was directly linked to a particular parental income unit. The characteristics of the parental income unit have an impact on the cash transfer outcomes for the single income unit. In addition, the outcomes for the single income unit can affect the cash transfer outcomes for the parents and any other single income units living at home with the parents. The method used to link away-from-home single income units to parental income units is described in section 3.3.

2.5 Imputation of workforce independence

Some single income units are deemed not to be dependent on their parents if they satisfy a threshold of workforce participation that gives them cash transfer independence. The information needed to determine workforce independence is not collected in the surveys, so it must be imputed in STINMOD. Prior to the release of STINMOD-99A, this workforce independence allocation applied only to AUSTUDY recipients and it was imputed in a simple random fashion designed to meet administrative benchmarks.

The introduction of Youth Allowance greatly increased the population in scope for parental dependency. Also, the requirements for being considered to be independent on workforce grounds became less demanding. As a result, a new approach for imputing workforce independence was tried. This approach used the Australian Youth Survey to generate a probability of workforce independence for those in scope for Youth Allowance. More detail is provided in section 3.3.

2.6 Basefile structure

Until the release of version 99A, the STINMOD basefile consisted of a single SAS dataset with one record for each income unit in the microdata. Each microdata record contained information for the income unit as a whole (such as the state of residence), as well as information for
the reference person, spouse if there was one, and up to five full-time student dependants aged 15–24 years who were living at home.

As mentioned earlier, the introduction of Youth Allowance greatly increased the importance of dependency links in the microdata. As a result, since the release of 99A, STINMOD’s base data have consisted of a main basefile as detailed above and a second SAS dataset, which has facilitated the processing of dependency links when calculating Youth Allowance and family assistance outcomes. Initially, this second SAS dataset contained information for those in a family (which might consist of several income units) potentially in scope for Youth Allowance and who might be considered dependent on their parents. The dataset is known as the ‘families’ file.

This file was augmented when a new system of family assistance was introduced in July 2000. Information was added for older dependants who were in scope for Family Tax Benefit (and not Youth Allowance). Further information was added to allow Youth Allowance outcomes to be calculated for all income units in a family when the parental income unit was being processed. A pointer variable in the main basefile is used to indicate whether the families file needs to be read.

The other change to the basefile flowed from the shift to an outyears approach, which is discussed in section 2.1. There are now basefiles for each outyear. This means that separate main and families files are created for each outyear.

### 2.7 Parameters

In addition to the base datasets, early versions of STINMOD included a single set of SAS datasets that contained values for the parameters of the tax and transfer programs modelled in STINMOD. These datasets are referred to as parameter files.

As a result of the outyears functionality, one set of parameter files for each financial year covered is produced and provided with STINMOD. These files contain average financial year values for the parameters. These are generated from a spreadsheet (derived from one that was

---

4 In STINMOD the age of independence is a parameter that can go as high as 25.
originally developed by the Commonwealth Treasury) that implements the indexation provisions for these programs. Projections of the consumer price index (CPI) and male total average weekly earnings lie behind the indexation outcomes.

2.8 Accessibility

Until recently, most users of STINMOD used the basefiles constructed by NATSEM. In theory, users could generate their own basefiles using the NATSEM basefile creation programs and their own assumptions. However, this required advanced programming skills and a considerable investment in time.

The way in which the outyears functionality has been developed has made it easier for users to include their own assumptions about future labour force, economic and demographic outcomes. This is done by editing a spreadsheet containing those assumptions and then using a custom menu option to create new STINMOD basefiles that reflect those assumptions. The process is automated and much more user-friendly. Further details are provided in section 4.

2.9 Adjustment of incomes

The methodology used to adjust incomes from the period of the survey to the current or a future period is mostly the same as that described by Percival (1994). However, the way in which wage and salary income is uprated has changed.

In the first release of STINMOD, wage and salary income was adjusted by a factor reflecting movements in total average weekly earnings. However, with evidence of increasing dispersion in the distribution of wage and salary earnings, a more disaggregated approach was adopted. Now, 10 adjustment factors are used. These factors capture wage and salary growth rates by full-time and part-time labour force status and by quintile of income. This is discussed in more detail in section 3.2, as are the methods used to project movements in all other income components, as required by the outyears methodology.
3 Steps for creating a STINMOD basefile

There are three main steps required to generate a STINMOD basefile:

- calculate weights
- uprate private incomes
- impute family links and workforce independence, and make other miscellaneous imputations.

Each of these steps is now discussed in detail.

3.1 Calculation of weights

The current STINMOD basefile is derived from the 1996-97 and the 1997-98 SIHCs and a dataset that simulates Australia’s institutionalised population. Each record in the basefile represents an income unit and contains information about the income unit as a whole and information on each person in the income unit.

The first step in creating the STINMOD basefile is to assign each of the individuals in an income unit a weight. This weight represents the likelihood of finding persons with a similar set of characteristics in the Australian population. There are two main steps used to calculate the weights.

First, initial weights are calculated for each individual using a ratio approach. A LFS matrix file supplied by the ABS is used as the benchmark data in this process. The structure of this file is described in Landt et al. (1994). However, to create STINMOD basefiles for future points in time, the LFS matrix is firstly reweighted to represent the projected demographics at specified points in time.

The next step is to take the initial weights calculated from the LFS ratio reweighting process and adjust them to better match administrative program numbers. This is done using the CALMAR reweighting algorithm.

The CALMAR software (Deville and Särndal 1992) was developed by the Institut National de la Statistique et des Études Économiques (INSEE). In 1996 this software was provided to NATSEM by the ABS with the
permission of INSÉÉ. Given an initial set of weights and a set of benchmarks (in this case, actual and projected program numbers are used as the benchmark data), CALMAR adjusts the weights to more closely match those benchmarks. The new weights are the optimal solution to minimising a specified distance (between the old and new weights) function, given a specified set of constraints — in this case the administrative benchmarks. One of four different distance functions can be specified, with each having strengths and weaknesses in terms of the distribution of weight ratios. They are:

- linear method
- exponential method
- logit method
- linear truncated method.

The exponential method is used to adjust the weights in the STINMOD basefile. This produces unbounded weights but ensures that all weights are positive.

**Basefiles workbook**

The data required to calculate weights for a STINMOD basefile are contained in the Excel workbook ‘Outyears Basefiles.xls’. This workbook contains a number of worksheets, but only two of them relate to the calculation of the weights. They are:

- LFS Projections
- FaCS & DVA Program Numbers.

The LFS Projections worksheet contains both actual and projected data relating to demographics and labour force status for May and November of each year from 1995 to 2006. The information in this sheet is used to reweight the most recent LFS data available from the ABS to future

---

5 Providing the difference between the initial weights and the final weights is reasonably small, convergence will occur. A stopping threshold is set and there is convergence when the maximum absolute value of differences between a weight ratio calculated from successive iterations is less than the threshold. A maximum number of iterations is specified in order to achieve convergence and the algorithm stops if it has not converged after this number. Upon convergence, a new variable containing the adjusted weights is added to the STINMOD basefile.
points in time, based on the projections for each period. The May and November points in time are considered to be sufficiently close to the June and December quarters, for which STINMOD basefiles are created.

Projections of the number of people are made for May and November of each year for each of the following categories:

- age
- sex
- labour force status
- state of residence
- education status
- family status
- institution type
- dwelling type.

Various techniques are used to project the demographic and labour force characteristics of the population. The starting point is the ABS population projections by age and sex for May of each year. The projections for November are calculated by averaging the preceding and following May figures.

ABS projections of the labour force participation rate and unemployment rate are then applied to the population projections to give the number of people in the labour force, the number of people unemployed and the number of people not in the labour force for each projection period. The number of people in the labour force is then broken down according to employment status (for example, full-time wage and salary earner, full-time self-employed, and part-time wage and salary earner). This is done on the basis of average proportions for the preceding four quarters. The number of people in each state, study status category and family status category are also apportioned on the basis of average proportions for the preceding four quarters.

The projections in the LFS Projections worksheet are used to reweight the LFS matrix file (see ‘Reweighting the LFS matrix file’ below). Once the LFS matrix has been reweighted, initial weights for the STINMOD basefile are calculated using the ratio reweighting approach described below in ‘LFS matrix ratio reweighting’.
The FaCS & DVA Program Numbers worksheet contains the number of recipients of payments made by the Department of Family and Community Services (FaCS) and the Department of Veterans’ Affairs (DVA). The program numbers listed in this sheet are used as the targets in the CALMAR reweighting process. Actual numbers of recipients are taken from FaCS and DVA publications and FaCS administrative data. Projected recipient numbers are generally taken from departmental projections. The data in this sheet are described more fully below in ‘Administrative projections’.

Reweighting the LFS matrix file

The Labour Force Survey is considered to provide accurate, regular estimates of labour force participation by a range of socioeconomic indicators including state of residence, sex, age, occupation, industry of occupation, family status and education status. Thus, the LFS was selected as a suitable benchmark dataset against which to initially reweight the STINMOD basefile.

For the current version of STINMOD, 01B, the most recently available LFS matrix is for November 2000. However, the outyears facility requires STINMOD basefiles for the June and December quarters out to 2005-06. To derive initial weights for a STINMOD basefile for periods beyond November 2000, information in the LFS matrix file has to be projected to give benchmarks for May and November out to 2005-06.

A projected version of the LFS matrix is produced by using the CALMAR algorithm to reweight the most recent LFS matrix file according to the data specified in the LFS Projections worksheet. This gives a projected LFS matrix file for May and November of each year to 2005-06.

LFS matrix ratio reweighting

As discussed in section 2.3, the sample of individuals living in private dwellings for the current STINMOD basefile is taken from the 1996-97

---

6 Projections of the number of recipients of FaCS payments are provided by the Department of Family and Community Services. Projections of the number of recipients of DVA payments are made using a linear trend function.
and the 1997-98 SIHCs. To this is added a synthetic institutionalised population derived from information in the Census of Population and Housing.

Once the projected LFS matrix file has been produced, a ratio reweighting approach is used to assign initial weights to the two SIHC samples. The institutionalised population is excluded from this process due to the limited range of variables in the institutionalised population dataset. The institutionalised have their weights adjusted using CALMAR (see ‘Using CALMAR to adjust the weights on the STINMOD basefile’ below).

The ratio reweighting process involves matching records from the SIHCs and the LFS private dwelling population. A detailed description of the process can be found in Landt et al. (1994). However, a summary is given here.

Initially, records from the two surveys are matched by state of residence, sex, age, labour force status, family status and education status. Each record in the combined SIHC data that matches the LFS by all of these categories is assigned the LFS population figure. The weight for the SIHC record is then calculated by dividing the LFS population figure by the number of SIHC records it has matched. For example, if there are 10 SIHC records matching a LFS record that has a population weight of 200, then each SIHC record is given an initial weight of 20.

Not all LFS data records find a match in the SIHC, and vice versa. Those records in the SIHC that are not matched with the LFS at the highest level (by state of residence, sex, age, labour force status, family status and education status) are passed through to the next stage of matching. This stage involves dropping one of the matching variables in the LFS and repeating the process described above. Education status is the first variable to be dropped, followed by family status and then labour force status.

It could also be that there are records in the LFS that are not matched in the SIHC. To account for this, factors are calculated to inflate the values

---

7 The LFS includes individuals from both private and non-private dwellings. However, since the ratio reweighting technique is used only for the SIHCs and the SIHCs do not contain individuals from non-private dwellings, the LFS matrix file is subsetted to exclude individuals from non-private dwellings.
of the new weights calculated in the process above. The factors are calculated for each state by sex and age groupings.

After the ratio reweighting process has been completed, initial weights have been calculated for each person in the STINMOD basefile that lives in a private dwelling. Weights for individuals living in institutions are derived using a different technique.

For the institutionalised population, the initial weights are those that were calculated as part of the creation of the synthetic data. These initial weights were calculated using CALMAR with age groups, labour force status, student status and gender as the benchmark data. Lim and Percival (1999) discuss in detail the process for creating the weights for the institutionalised population.

Having calculated initial weights for the private dwelling population and using the stored weights for the institutionalised population, the next step is to refine these weights using CALMAR, so as to better match administrative projections for various FaCS and DVA payments.

**Administrative projections**

Actual numbers and projections for selected government programs are recorded in the FaCS & DVA Program Numbers worksheet of the Outyears Basefiles.xls workbook. These numbers are used as the targets in the CALMAR reweighting process to refine the initial weights calculated from the LFS matrix ratio reweighting process described above.

The FaCS & DVA Program Numbers worksheet contains information about the actual or projected number of recipients of each FaCS and DVA payment by age group and, in some cases, marital status. Data on the actual number of recipients of FaCS payments is taken from FaCS publications or administrative data. Projections of the total number of recipients of the payments are provided by FaCS. Once the projected total is known, the breakdown by age group and marital status is calculated based on the distributions shown in the most recent actual data.

The projection figures provided for FaCS payments are for no particular point within the financial year, so they are assumed to be estimates for
the December quarter, which is the mid-point of the year. Projected figures for the June quarters are calculated as the average of the preceding and following December figures.

For DVA payments, projections are made slightly differently. The number of recipients of each DVA payment is predicted for the December quarter using a linear trend function. For the June quarter the number is taken as the average of the preceding and following December figures.

Not all of the payments that are listed in the FaCS & DVA Program Numbers worksheet are used as reweighting targets in CALMAR and, of those payments that are used as targets, the breakdown of the categories is not always as it appears in the spreadsheet. For instance, recipient numbers categorised by marital status are not used as targets for CALMAR (due to concerns about CALMAR converging).

For most of the payments, recipient numbers categorised by age group are used as the CALMAR targets. However, the age groups listed in the worksheet are sometimes collapsed. This is necessary to ensure the CALMAR reweighting process will converge, since CALMAR will not converge if the numbers in the target groupings are too small.

Although Carer Payment and Wife Pension numbers are recorded separately in the FaCS & DVA Program Numbers worksheet, they are amalgamated into a single category for the CALMAR reweighting process. This is because the SIHC does not distinguish separately between these two payments.

Unlike the FaCS payments, DVA recipient numbers are not categorised by age. Only the total numbers of recipients are used in the CALMAR reweighting process.

In summary, the payments used as reweighting targets in CALMAR and the age groupings used in the reweighting process are as follows:

- Age Pension (60–64, 65–69, 70–74, 75+)
- Disability Support Pension (<30, 30–39, 40–49, 50–59, 60–64, 65+)
- Carer Payment and Wife Pension (<30, 30–39, 40–49, 50–59, 60–64, 65+)
- Widow B Pension (45+)
• Newstart Allowance, including Mature Age Allowance and Youth Allowance for the unemployed (<21 (Youth Allowance), 21–24, 25–34, 35–44, 45–54, 55–59, 60–64 (Mature Age Allowance))
• Sickness Allowance (21–24, 25–34, 35–44, 45–54, 55–64)
• Special Benefit (<35, 35–59, 60+)
• Parenting Payment Unpartnered (<30, 30–39, 40–49, 50+)
• Department of Veterans’ Affairs (Service and DVA Disability Pension, Service Pension, DVA Disability Pension, DVA War Widows Pension).

Using CALMAR to adjust the weights on the STINMOD basefile

After initial weights have been calculated using the LFS ratio reweighting process, the CALMAR reweighting algorithm (discussed earlier in this section) is used to adjust the weights of the combined SIHC and institutionalised datasets so as to better align aggregates with program benchmarks. CALMAR is also used to adjust the weights of people according to their dwelling type and whether they live in a private dwelling or an institution.

The first step in the CALMAR reweighting process involves creating a person-level dataset from the combined SIHC and institutionalised datasets. For each payment type used as a reweighting target in CALMAR, each individual in the SIHC dataset is placed in a category on the basis of their age and whether they receive that payment. For example, as outlined above in ‘Administrative projections’, one of the payments used in the CALMAR reweighting process is Age Pension. People shown in the SIHC data as receiving this payment are placed in one of four age groupings coinciding with the target age groups specified in CALMAR. People shown in the SIHC data as not receiving any Age Pension are placed in a residual category (not based on age). A similar process is used to allocate individuals according to their dwelling type.

The next step is to read in the program recipient numbers and dwelling type marginals (that is, the reweighting target numbers). The program recipient marginals are taken from the FaCS & DVA Program Numbers worksheet, and the dwelling type marginals are taken from the reweighted LFS matrix (based on the data in the LFS Projections worksheet). The CALMAR reweighting program is then run to estimate
new weights for the persons dataset using the program and dwelling type marginals as targets.

At this point there is a set of data with weights that reflect the Australian population based on labour force and administrative payment projections.

3.2 Adjustment of incomes

The income data collected in the SIHC need to be updated to the current time from either 1996-97 or 1997-98, depending on the year of the survey in which the individual participated. However, only private incomes are uprated as part of the STINMOD basefile creation process. The value of transfer payments recorded in the surveys is not uprated, since STINMOD calculates these amounts.

Private incomes in the STINMOD basefile are uprated using income uprating factors calculated in the Outyears Basefiles.xls workbook. Different factors need to be applied to the two surveys to take into account the different years in which they were conducted. The incomes in the institutionalised dataset are uprated using the factors used to uprate the 1996-97 SIHC.8 The two relevant worksheets in the Outyears Basefiles.xls workbook for uprating incomes are:

- Income Factors Data
- Income Uprating Factors.

The Income Factors Data worksheet contains historical data and projections for the various sources of private income such as wage and salary, business and dividend income and some expenditure items such as private and public rent and child support payments.

The Income Uprating Factors worksheet uses the actual and projected income series from the Income Factors Data sheet to calculate the adjustment factors to be applied to the private income variables in the

---

8 Despite the fact that the institutionalised population is derived from the 1991 Census of Population and Housing (as discussed in section 2.3), at the time this synthetic population was constructed, incomes were uprated to 1996 levels. Hence, in the STINMOD basefile creation process the incomes of the institutionalised population are taken from 1996 to the required point in time.
1996-97 and the 1997-98 SIHCs and the institutionalised dataset. Both current and period (previous year’s annual income) private incomes need to be uprated, so factors are calculated for both. The general rules for calculating adjustment factors are as follows:

- **current income variables**
  \[
  \text{Adjustment factor} = \frac{\text{Projected income level for future point in time}}{\text{Income level for survey point in time}}
  \]

- **period income variables**
  \[
  \text{Adjustment factor} = \frac{\text{Projected financial year average income for future point in time}}{\text{Financial year average income for previous year to survey}}
  \]

**Uprating wage and salary income**

Using the ABS Employee Earnings and Trade Union Membership data (ABS Cat. No. 6310.0), wage and salary earnings by quintile are calculated for both full-time and part-time workers. These figures are then projected forward using the average weekly earnings projections located in the Outyears Parameters.xls workbook. Uprating factors are calculated from these figures according to the formulas given above and stored in the Income Uprating Factors worksheet.

In the STINMOD basefile, individuals are classified according to whether they are wage and salary earners and then whether they are working full-time or part-time. Full-time wage and salary earners are then ranked by quintiles according to their income from wages and salaries and the relevant income uprating factor is then applied to that income. The same process is repeated for part-time wage and salary earners.

**Uprating self-employment income**

Adjustment factors for income from self-employment are derived from ABS National Accounts data and unpublished LFS data. Two separate factors are calculated — one for self-employment income relating to farming industries and one for all other (non-farm) self-employment income.

An estimate of the average non-farm self-employment income is calculated by dividing the total national non-farm income for
Creating a Basefile for STINMOD

unincorporated enterprises (from the National Accounts) by the number of self-employed people not in the farming industry recorded in the LFS. An estimate of the average farm income of self-employed people is derived in a similar way. However, since average farm income fluctuates greatly it has been ‘smoothed’ using a log function.

Projections of the total national farm and non-farm income for unincorporated enterprises are made using a linear trend function. Projections of the number of self-employed people are taken from the LFS Projections worksheet and apportioned between the farm and non-farm industry on the basis of the average split in the known data.

**Uprating child support payments**

Unpublished child support payments data on the number of people making child support payments and the total amount paid have been collected from the Child Support Agency. These figures are used to derive the average payment amount. The income adjustment factor is calculated on the basis of the change in average payments over time. Where actual data are not available, projections are made on the basis of a fitted linear regression equation.

The factors calculated from the child support payments data are used to uprate child support payments received and paid. Although the number of families paying and receiving child support has been increasing over time, this is not accounted for specifically in the current STINMOD basefile creation process. However, it is an area for possible development (see section 6).

In the process of uprating child support payments for the current STINMOD basefile, only those who are shown in the survey to be either receiving or paying child support have the amounts adjusted. No new child support recipients or payers are generated.

**Uprating income from other sources**

Incomes from dwelling rents, interest, dividends, royalties and other investments are uprated using factors derived from unpublished ABS National Accounts data. The National Accounts data provide the
aggregate level of income for each of these categories, enabling factors to be derived for each.

Projections for each series — dwelling rents, interest, dividends, royalties and other investments — are done in two different ways in the Income Uprating Factors worksheet. Users can choose the one they prefer. The first method uses the CPI to project future levels of each of the income categories. The second method smooths and projects the income series using a logarithmic trend function. In the current version of STINMOD and in the basefile creation code, the trend function is used to project future values of these incomes.

All other sources of private income, including superannuation income, workers compensation payments, income from persons not in the household and other regular income, are uprated using the CPI.

Uprating housing costs

Unpublished ABS data from the CPI housing series are used to uprate housing costs. Separate factors are calculated for each state of residence for the following housing cost categories:

- private rent
- public rent
- mortgages
- rates.

The CPI housing series has undergone a number of changes in recent years and unfortunately data are no longer collected separately for publicly and privately owned dwellings. Thus, rents for publicly and privately owned dwellings are projected using the headline CPI.9

3.3 Imputations

The SIHC dataset does not contain all the information needed to be able to accurately apply the rules of the government cash transfer and income tax systems modelled in STINMOD. Therefore, a number of imputations need to be performed when constructing the basefile.

9 The headline CPI used is for all groups, weighted average of eight capital cities.
Imputing workforce independence

The payment requiring the most imputation is Youth Allowance. Youth Allowance is paid to full-time students and non-full-time students. Recipients of Youth Allowance can be either dependent or independent. One way of being considered independent is on the basis of previous workforce experience. To be workforce independent, an individual must satisfy one of the following criteria <www.centrelink.gov.au/internet/internet.nsf/payments/qual_ind_yal.htm>.

- The individual has been out of school for at least 18 months and has earned at least 75 per cent of the Commonwealth Training Wage in the 18 months prior to lodging a claim.
- The individual has supported himself or herself by working full time (at least 30 hours a week) for at least 18 months during the preceding two years (or for a period or periods of 12 months if considered disadvantaged).
- The individual has worked part-time (at least 15 hours a week) for at least two years since last leaving school.

The information required to determine whether an individual is workforce independent is not available from the SIHC and so workforce independence status has to be imputed for any young people in scope for Youth Allowance. To do this, a logistic regression equation was estimated from the Australian Youth Survey (a longitudinal survey conducted between 1989 and 1994), which contains data about the workforce participation patterns of young people. This equation is used to generate a probability of workforce independence for young people in scope for Youth Allowance. Based on these probabilities, workforce independence is randomly imputed to a proportion of the population in scope.

Linking away-from-home dependants to a parental record

For recipients of Youth Allowance who are considered to be dependent on their parents, the amount of Youth Allowance they are entitled to depends on the income of their parents, the number of other dependants in the family and their own income. As described in section 2.6, a families dataset is used to collate this information. This dataset contains family-level information needed to perform the parental income test for
dependants eligible for Youth Allowance (for more information see sections 2.4 and 2.6).

The families dataset contains information on the combined parental income, the number of dependants aged less than 16 years in the family and the ages and incomes of up to five full-time students and up to five non-student dependants eligible for Youth Allowance. One student and one non-student may be living away from home. The families dataset also contains information about children of the family reference person who are not in scope for Youth Allowance, but who are single, have no children and live at home. This information about children of the family reference person who are not in scope for Youth Allowance is required because Family Tax Benefit Part A is now available for all non-Youth Allowance dependants aged 16–20 years and not studying or aged 16–24 years and studying full-time and not workforce independent.

It is possible to identify from the SIHC all dependants in a family who are living at home. However, the structure of the Youth Allowance parental income test means that it is also necessary to account for dependants living away from home. Entitlement to Youth Allowance for away-from-home dependants is based on parental income and the number of other dependants in the family. Similarly, Youth Allowance entitlement for at-home dependants is based on parental income and the number of dependants in the family, including those living away from home. Unfortunately, there is no way of identifying from the SIHC the parents of a dependant who is living away from home. Indeed, it is likely that the parents are not even included in the survey. Thus, away-from-home dependants need to be linked to suitable parental income units (who are representative of their parents) by imputation.

Dependent students and other people eligible for Youth Allowance who are living away from the parental home but who are likely to be subject to the parental income test for Youth Allowance are identified first. These are single person income units, never married, aged 16–24 years, not workforce independent and in receipt of unemployment or sickness benefits, or are full-time students and not the sons or daughters of family heads.
In STINMOD these away-from-home dependants are linked to a parental income unit drawn from the families file. In other words, they are linked to a family that has other young people either in scope for Youth Allowance or living at home, single, with no children but not in scope for Youth Allowance. The parental income units are classified according to their income group and whether they are headed by a couple or single parent.

The probabilities of young people having parents in each of the income and couple–single parents groups are determined from the Australian Youth Survey. The probability distributions are determined separately for full-time students and non-full-time students. These probabilities are then used to randomly link each away-from-home dependant with a family in the families file. The away-from-home dependant is added to the families file as student 5 or unemployed 5. Each family can have a maximum of one dependent full-time student living away from home and eligible for Youth Allowance and one dependent non-full-time student living away from home and eligible for Youth Allowance.

**Other imputations**

Other imputations performed as part of the STINMOD basefile creation process include the following.

- A disability level for DVA pension recipients is imputed onto the basefile. This is because the level of DVA entitlements depends on the level of disability suffered by the claimant.

- Each person in the STINMOD basefile is assigned a random number for private health insurance take-up. This is used to assess whether a person or family is subject to the Medicare levy surcharge in the STINMOD entitlement modules.

- Recipients of Carer Payment and Wife Pension are imputed. Because the receipt of these benefits is recorded as a single variable in the SIHC, imputation is required to determine whether an individual receives either Carer Payment or Wife Pension.
4 Changing assumptions and generating a new basefile for STINMOD

As discussed, the generation of STINMOD basefiles relies on various assumptions. If a user wants to change any of these assumptions, it is usually very easy to do through the Outyears Basefiles.xls workbook. Assumptions about current or future economic conditions, labour market information, growth in private incomes, changes in average weekly earnings and the CPI, the number of FaCS and DVA customers, and the demographic breakdown of the population can be adjusted by the user.

Because many of the processes are automated, a new STINMOD basefile reflecting the different assumptions can be created with the click of a button. To create a new STINMOD basefile, the first step is to select Reweight LFS Matrix from the STINMOD Outyears menu. This needs to be done only if a change is made in the LFS Projections worksheet, but it does not hurt to do it. After the LFS matrix has been reweighted, the second step is to select Create Basefiles from the STINMOD Outyears menu. Users can recreate STINMOD basefiles for all June and December quarters from the December quarter of 1998 to the June quarter of 2006, or they can choose to create new basefiles for particular points in time.

5 Linking the STINMOD basefile to the original SIHC microdata

The variables in the STINMOD basefile are a subset of the variables in the SIHC dataset. In general, only those variables that are needed to create a STINMOD basefile or to run the STINMOD basefile through the entitlement modules are retained. There are many other SIHC variables that are not stored in the STINMOD basefile for reasons of storage space and simplicity.

Sometimes, a user may need information from a SIHC variable that is not stored on the basefile. Additional variables from the SIHC dataset can be merged into the STINMOD basefile using person identification

NATSEM Technical Paper no. 27
numbers, thus allowing the user to access greater information about the person or their family.

Each household, family, income unit and person in the SIHC dataset has a unique identifier attached to it. The identifiers are calculated to allow users to link people in the same income unit or family. They are calculated using the following equations:

\[
\text{Household ID} = \text{random identification number} \\
\text{Family ID} = (\text{household ID} \times 10) + \text{family number} \\
\text{Income unit ID} = (\text{household ID} \times 100) + (\text{family ID} \times 10) + \text{income unit number} \\
\text{Person ID} = (\text{household ID} \times 1000) + (\text{family ID} \times 100) + (\text{income unit ID} \times 10) + \text{person number}
\]

Each observation in the STINMOD basefile is an income unit and each income unit has a unique identification number — an income unit ID — assigned to it. Similarly, each person aged 15 years or more within the income unit has a unique identifier — a person ID — assigned to them. These identifiers are calculated in a similar way to the SIHC identifiers. However, the income unit and person identification numbers in the STINMOD basefile are not the same as those stored in the SIHC microdata for the following reasons.

- Because the STINMOD basefile comprises data from two SIHCs and the institutionalised population from the census, it is necessary to ensure that identification numbers coming from different sources are not duplicated.
- It is necessary to be able identify the survey in which each income unit was interviewed.
- When recipients of away-from-home dependent Youth Allowance are linked to a parental income unit (see section 3.3), their original income unit number from the SIHC is replaced with the income unit number of their imputed parents.

Thus, the income unit and person identification numbers in the STINMOD basefile are modified versions of the SIHC identification numbers. The modifications entail multiplying the original SIHC income unit or person identifier by 10 and then adding either ‘6’ or ‘7’ to indicate which SIHC the record is from. If the record is from the 1996-97 survey, a
‘6’ is added. If it is from the 1997-98 survey, a ‘7’ is added. Records from the institutionalised dataset have a unique identifier ending in ‘0’ to indicate their origin.

Because recipients of away-from-home dependent Youth Allowance have their SIHC income unit identifier replaced by the income unit identifier of their imputed parents, the only way of linking the STINMOD basefile with the SIHCs is to use person identification numbers and linking at the person level. The original SIHC identifiers can be derived from the STINMOD person identifiers. This is done by separating individuals according to the survey they originated from (using the ‘6’ or ‘7’ on the end of the person identification number) and then truncating their person identification number (that is, removing the ‘6’ or ‘7’) to revert it to the original SIHC number. Records can be matched with the relevant SIHC.

A couple of important things need to be kept in mind when merging variables from the SIHC dataset on to the STINMOD basefile. Firstly, any variables relating to income that are merged from the SIHC will not have been uprated to the period of the STINMOD basefile. If required, the user would need to uprate these variables separately.

Secondly, due to the already limited number of variables in the institutionalised population, it is not possible to merge extra variables for this group. Thus, variables that are merged on to the STINMOD basefile from the SIHC dataset will not have values for the institutionalised people within the STINMOD basefile.

NATSEM has some pre-written SAS code that takes the STINMOD basefile from an income unit to a person level and merges specified variables from the SIHC dataset on to the STINMOD basefile. This code is available to STINMOD source code users.

6 Future directions

While much work has been undertaken in recent years to improve the STINMOD basefile, there is still scope for improvement. Some known concerns and possibilities for enhancing the base population follow.
• Some techniques to improve the match of STINMOD customer numbers to known administrative customer numbers are currently being investigated. One approach being considered involves running a preliminary income-uprated basefile through the STINMOD entitlement modules to flag payment recipients. This information would then be used in the CALMAR reweighting stage of the basefile creation process. The reason behind this is that people who are recorded as receiving a payment in the SIHC may no longer qualify for that payment after incomes have been uprated. By using only people who are known to be receiving payments after incomes have been uprated in the CALMAR stage it is hoped that STINMOD program numbers will align better to administrative data.

• Currently the numbers of Wife Pension and Carer Payment recipients are combined for the purposes of reweighting with CALMAR because the two payments were not distinguished in the original surveys. It may be possible in the future to better impute which of the payments a person is receiving and separate the two for reweighting. This may lead to closer matches to administrative data for these payments.

• Some sources of income, such as child support, have become more widespread since the SIHCs were undertaken. This problem will grow as the dates of the surveys lag further behind the point in time at which the basefile is reweighted. More sophisticated and accurate uprating techniques may need to be developed to address this issue.

• A change in the format for collecting National Accounts data by the ABS in 1999 has meant that information used to update some sources of income (namely, self-employed incomes, dwelling rental income, interest, dividends and royalties) is no longer available in the format required. The uprating process for these income sources therefore needs to be reviewed.

• NATSEM envisions that it may be possible in the future to reweight basefiles to a regional level to support more detailed analysis. NATSEM is currently working on a regional version of STINMOD.

• Currently the STINMOD basefile does not contain information on assets. This prevents the assets test for payments from being applied. As better information on asset holdings of people becomes available, it may be possible to impute assets onto the basefile. This has been attempted in the past, but due to limited available information it was deemed to be not accurate enough for the purposes of STINMOD.
References


NATSEM publications

Copies of NATSEM publications and information about NATSEM may be obtained from:

Publications Officer
NATSEM
University of Canberra ACT 2601
Australia
Ph: + 61 2 6201 2750 Fax: + 61 2 6201 2751
Email: natsem@natsem.canberra.edu.au

See also NATSEM’s website <www.natsem.canberra.edu.au>.

Periodic publications

NATSEM News keeps the general community up to date with the developments and activities at NATSEM, including product and publication releases, staffing and major events such as conferences. This newsletter is produced twice a year.

NATSEM’s Annual Review gives the reader an historical perspective of the centre and its achievements for the year.

Discussion Paper series

<table>
<thead>
<tr>
<th>No.</th>
<th>Authors</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0</td>
<td>Harding, A</td>
<td>Lifetime Repayment Patterns for HECS and AUSTUDY Loans, July 1993</td>
</tr>
<tr>
<td>2.0</td>
<td>Mitchell, D and</td>
<td>Changes in Poverty among Families during the 1980s: Poverty Gap Versus Poverty Head-Count Approaches, October 1993</td>
</tr>
<tr>
<td></td>
<td>Harding, A</td>
<td></td>
</tr>
<tr>
<td>3.0</td>
<td>Landt, J, Harding, A, Percival, R and Sadkowsky, K</td>
<td>Reweighting a Base Population for a Microsimulation Model, January 1994</td>
</tr>
<tr>
<td>No.</td>
<td>Authors</td>
<td>Title</td>
</tr>
<tr>
<td>-----</td>
<td>--------------------------------</td>
<td>-------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>5</td>
<td>Landt, J, Percival, R, Schofield, D and Wilson, D</td>
<td><em>Income Inequality in Australia: The Impact of Non-Cash Subsidies for Health and Housing</em>, March 1995</td>
</tr>
</tbody>
</table>
| 6   | Polette, J                     | *Distribution of Effective Marginal Tax Rates Across the Australian Labour Force*, August 1995  
(contributed to article in *Australian Economic Review*, 3rd quarter, pp. 100–6, 1995) |
| 9   | Paul, S and Percival, R        | *Distribution of Non-Cash Education Subsidies in Australia in 1994*, September 1995       |
| 10  | Schofield, D, Polette, J and Hardin, A | *Australia’s Child Care Subsidies: A Distributional Analysis*, January 1996 |
| 11  | Schofield, D                   | *The Impact of Employment and Hours of Work on Health Status and Health Service Use*, March 1996 |
| 13  | Schofield, D and Polette, J    | *How Effective Are Child Care Subsidies in Reducing a Barrier to Work?*, May 1996  
<p>| 14  | Schofield, D                   | <em>Who Uses Sunscreen?: A Comparison of the Use of Sunscreen with the Use of Prescribed Pharmaceuticals</em>, May 1996 |</p>
<table>
<thead>
<tr>
<th>No.</th>
<th>Authors</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>Lambert, S, Beer, G and Smith, J</td>
<td>Taxing the Individual or the Couple: A Distributional Analysis, October 1996</td>
</tr>
<tr>
<td>16</td>
<td>Landt, J and Bray, J</td>
<td>Alternative Approaches to Measuring Rental Housing Affordability in Australia, April 1997</td>
</tr>
<tr>
<td>17</td>
<td>Schofield, D</td>
<td>The Distribution and Determinants of Private Health Insurance in Australia, 1990, May 1997</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(published as ‘Ancillary and specialist health services: equity of access and the benefit of public services’, <em>Australian Journal of Social Issues</em>, vol. 34, no. 1, pp. 79–96, February 1999)</td>
</tr>
<tr>
<td>25</td>
<td>Percival, R and Fischer, S</td>
<td>Simplicity Versus Targeting: A Legal Aid Example, December 1997</td>
</tr>
<tr>
<td>No.</td>
<td>Authors</td>
<td>Title</td>
</tr>
<tr>
<td>-----</td>
<td>---------</td>
<td>-------</td>
</tr>
<tr>
<td>30</td>
<td>Harding, A</td>
<td><em>Tomorrow’s Consumers: A New Approach to Forecasting Their Characteristics and Spending Patterns</em>, June 1998</td>
</tr>
<tr>
<td>40</td>
<td>Harding, A and Robinson, M</td>
<td><em>Forecasting the Characteristics of Consumers in 2010</em>, March 1999</td>
</tr>
<tr>
<td>41</td>
<td>King, A and McDonald, P</td>
<td><em>Private Transfers Across Australian Generations</em>, March 1999</td>
</tr>
<tr>
<td>No.</td>
<td>Authors</td>
<td>Title</td>
</tr>
<tr>
<td>-----</td>
<td>---------</td>
<td>-------</td>
</tr>
<tr>
<td>43</td>
<td>King, A, Bækgaard, H and Harding, A</td>
<td>Australian Retirement Incomes, August 1999</td>
</tr>
<tr>
<td>45</td>
<td>Walker, A</td>
<td>Distributional Impact of Higher Patient Contributions to Australia’s Pharmaceutical Benefits Scheme, September 1999</td>
</tr>
<tr>
<td>46</td>
<td>Wilson, J</td>
<td>An Analysis of Private Health Insurance Membership in Australia, 1995, November 1999</td>
</tr>
<tr>
<td>49</td>
<td>Szukalska, A and Robinson, M</td>
<td>Distributional Analysis of Youth Allowance, July 2000</td>
</tr>
<tr>
<td>51</td>
<td>Lloyd, A, Harding, A and Hellwig, O</td>
<td>Regional Divide? A Study of Incomes in Regional Australia, September 2000</td>
</tr>
<tr>
<td>53</td>
<td>Lloyd, R and Hellwig, O</td>
<td>Barriers to the Take-Up of New Technology, November 2000</td>
</tr>
<tr>
<td>54</td>
<td>Reiman, C</td>
<td>The Gender Wage Gap in Australia: Accounting for Linked Employer–Employee Data from the 1995 Australian Workplace Industrial Relations Survey, March 2001</td>
</tr>
<tr>
<td>56</td>
<td>Harding, A and Greenwell, H</td>
<td>Trends in Income and Expenditure Inequality in the 1980s and 1990s, December 2001</td>
</tr>
</tbody>
</table>
### DYNAMOD Technical Paper series

<table>
<thead>
<tr>
<th>No.</th>
<th>Authors</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Antcliff, S</td>
<td>An Introduction to DYNAMOD: A Dynamic Microsimulation Model, September 1993</td>
</tr>
</tbody>
</table>

*Discontinued series. Topic is now covered by the broader Technical Paper series.*

### Dynamic Modelling Working Paper series

<table>
<thead>
<tr>
<th>No.</th>
<th>Authors</th>
<th>Title</th>
</tr>
</thead>
</table>

*Discontinued series. Topic is now covered by other series, including the broader Technical Paper series.*

### STINMOD Technical Paper series

<table>
<thead>
<tr>
<th>No.</th>
<th>Authors</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>Percival, R</td>
<td>Building STINMOD's Base Population, November 1994</td>
</tr>
<tr>
<td>3</td>
<td>Schofield, D and Paul, S</td>
<td>Modelling Social Security and Veterans' Payments, December 1994</td>
</tr>
<tr>
<td>4</td>
<td>Lambert, S</td>
<td>Modelling Income Tax and the Medicare Levy, December 1994</td>
</tr>
<tr>
<td>5</td>
<td>Percival, R</td>
<td>Modelling AUSTUDY, December 1994</td>
</tr>
<tr>
<td>6</td>
<td>Landt, J</td>
<td>Modelling Housing Costs and Benefits, December 1994</td>
</tr>
<tr>
<td>7</td>
<td>Schofield, D</td>
<td>Designing a User Interface for a Microsimulation Model, March 1995</td>
</tr>
<tr>
<td>8</td>
<td>Percival, R and Schofield, D</td>
<td>Modelling Australian Public Health Expenditure, May 1995</td>
</tr>
<tr>
<td>9</td>
<td>Paul, S</td>
<td>Modelling Government Education Outlays, September 1995</td>
</tr>
<tr>
<td>10</td>
<td>Schofield, D, Polette, J and Hardin, A</td>
<td>Modelling Child Care Services and Subsidies, January 1996</td>
</tr>
<tr>
<td>11</td>
<td>Schofield, D and Polette, J</td>
<td>A Comparison of Data Merging Methodologies for Extending a Microsimulation Model, October 1996</td>
</tr>
</tbody>
</table>

*Series was renamed the Technical Paper series in 1997.*
<table>
<thead>
<tr>
<th>No.</th>
<th>Authors</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>13</td>
<td>Galler, HP</td>
<td>Discrete-Time and Continuous-Time Approaches to Dynamic Microsimulation Reconsidered, August 1997</td>
</tr>
<tr>
<td>17</td>
<td>Lim, PP and Percival, R</td>
<td>Simulating Australia’s Institutionalised Population, May 1999</td>
</tr>
<tr>
<td>18</td>
<td>Szukalska, A, Percival, R and Walker, A</td>
<td>Modelling Child Care Utilisation and Benefits, November 1999</td>
</tr>
<tr>
<td>19</td>
<td>King, A, Bækgaard, H and Robinson, M</td>
<td>DYNAMOD-2: An Overview, December 1999</td>
</tr>
<tr>
<td>20</td>
<td>King, A, Bækgaard, H and Robinson, M</td>
<td>The Base Data for DYNAMOD-2, December 1999</td>
</tr>
<tr>
<td>21</td>
<td>Abello, A, King, A and Robinson, M</td>
<td>Demographic Projections with DYNAMOD-2, May 2002</td>
</tr>
<tr>
<td>22</td>
<td>King, A, Walker, A and Bækgaard, H</td>
<td>Modelling Overseas Migration in DYNAMOD-2, June 2002</td>
</tr>
<tr>
<td>23</td>
<td>Robinson, M and Bækgaard, H</td>
<td>Modelling Students In DYNAMOD-2, June 2002</td>
</tr>
<tr>
<td>24</td>
<td>Bækgaard, H</td>
<td>Modelling the Dynamics of the Distribution of Earned Income, June 2002</td>
</tr>
<tr>
<td>26</td>
<td>Percival, R</td>
<td>Comparing Simulated and Survey Incomes, June 2002</td>
</tr>
</tbody>
</table>
### Policy Paper series

<table>
<thead>
<tr>
<th>No.</th>
<th>Authors</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Harding, A and Polette, J</td>
<td><em>The Distributional Impact of a Guns Levy</em>, May 1996</td>
</tr>
<tr>
<td>2</td>
<td>Harding, A</td>
<td><em>Lifetime Impact of HECS Reform Options</em>, May 1996</td>
</tr>
<tr>
<td>3</td>
<td>Beer, G</td>
<td><em>An Examination of the Impact of the Family Tax Initiative</em>, September 1996</td>
</tr>
</tbody>
</table>