NATSEM
UNIVERSITY OF CANBERRA

Longitudinal benchmarking and alignment of a dynamic microsimulation model

IMA Conference Paper

BY

Simon Kelly and Richard Percival

FOR

2nd General Conference of the
International Microsimulation Association,
Ottawa, Canada

8-10 JUNE 2009

Abstract

The National Centre for Social and Economic Modelling is developing a new dynamic microsimulation model of the Australian population known as APPSIM. A fundamental component in establishing the credibility of APPSIM will be the match between the simulation outcomes and 'real world' outcomes. Most dynamic microsimulation models in the past have primarily focused on matching cross-sectional outcomes over time. In APPSIM, the aim is to improve the outcome match by using two distinct types of alignment. The first will be the traditional cross-sectional alignment currently being used in other models. The second will be longitudinal alignment. This will ensure that the longitudinal distribution in the simulation also matches the real world. For example, the traditional cross-sectional alignment will ensure the simulated number of births in each period matches external benchmarks. The longitudinal alignment will ensure that the proportions of women having a certain number of children (0, 1, 2, 3, or 4+) during their lifetime match external benchmarks.

This paper discusses the data used for longitudinal benchmarking, the issues involved and methods to be used in APPSIM.

Author note

Simon Kelly is a Principal Research Consultant at the National Centre for Social and Economic Modelling (NATSEM) and Adjunct Associate Professor at the University of Canberra. Richard Percival is a Principal Research Fellow at NATSEM and Associate Professor at the University of Canberra.

Acknowledgements

The authors would like to gratefully acknowledge the funding provided by the Australian Research Council (under grant LP0562493), and by the 12 research partners to the grant : Treasury; Broadband, Communications and the Digital Economy; Education, Employment and Workplace Relations; Health and Ageing; Innovation, Industry, Science and Research; Finance and Deregulation; Families, Housing, Community Services and Indigenous Affairs; Immigration and Citizenship; Prime Minister and Cabinet; the Productivity Commission; Centrelink; and the Australian Bureau of Statistics. The authors would like to thank Ann Harding for helpful comments on earlier drafts.

Contact Details

Postal: NATSEM, University of Canberra, Canberra ACT 2601, Australia.
Email: simon.kelly@natsem.canberra.edu.au
Tel: +61 2 6201 2780
Fax: +61 2 6201 2751

1 INTRODUCTION

Over the last decade, dynamic microsimulation models have moved from being "academic exercises" to being an integral component of government social policy development in a number of countries. In these countries there is an increasing recognition that the capability of dynamic microsimulation to project both detailed and aggregate outcomes of policy change provides a level of understanding not otherwise available. Despite the strengths of dynamic microsimulation models, there is still resistance to using these models in some countries. The reason for this is that most countries have 'elaborate and complex welfare state programmes' (Harding and Gupta 2007) and the projection models to replicate these complex systems are expensive, extremely difficult to develop, difficult to validate, and difficult to align to real world benchmarks (Harding 2007) .

Within Australia, NATSEM commenced development of the Australian Population and Policy Simulation model (APPSIM) in 2005, with the first version due for completion at the end of 2009. The model is being developed with funding support from the Australian Research Council and 12 government agencies and is intended to provide an essential component of the modelling infrastructure for Australian policymakers. Credibility for APPSIM, a dynamic population microsimulation model,, is critical for its continued development and acceptance within government. One essential element that is required for APPSIM to be accepted as credible is its ability to track historic data and produce projections that are consistent with other related official and non-official projections (Bacon and Pennec 2007b). This requirement ensures a significant focus on validation and alignment.

Accordingly, this paper will take a detailed look at the cross-sectional and longitudinal validation and alignment in APPSIM. However, discussion of the individual validation elements within APPSIM is not enough. The simulation of a policy change is often the only way that the individual and joint effects of the all the processes in a model can be tested and problems, which otherwise would not be apparent, can be seen. To address this aspect of achieving credibility, the latter parts of this paper discuss the simulation of an economic downturn as a test of the overall validation and alignment of APPSIM.

While some aspects of alignment in APPSIM have already been discussed in other working papers and technical notes (see Bacon and Pennec 2007a, 2007b; Pennec and Bacon 2007; Keegan 2007; and Keegan and Thurecht 2008), this paper provides considerably more detail on the validation processes employed. An overview of validating dynamic microsimulation models is given in Section 2 and a description of APPSIM is provided in Section 3. The APPSIM validation tools and techniques, both cross-sectional and longitudinal, are discussed in detail in Section 4. The focus then changes to overall validation and alignment of the model by applying a policy scenario to APPSIM. To enable the example to be discussed, Section 5 describes how superannuation and social security are modelled. Section 6 provides a timely example of the usefulness of a credible dynamic microsimulation model by analysing the impact of the global financial crisis on superannuation from an individual and government perspective.

2 VALIDATING DYNAMIC MODELS

One of the most successful dynamic models in the world, in both the thoroughness of its validation and its continued use by policy makers, is DYNACAN, the Human Resources Development Canada dynamic microsimulation model. DYNACAN has overcome the many hurdles that face the builders of dynamic models and developed a model that aligns with known outcomes and is an integral part of the policy development process for the Canada Pension Plan (Morrison, 2008). These hurdles in the Australian context are discussed in Kelly and King (2001).

Quality is the essential ingredient needed for a model to be accepted by government. The policy makers need a 'high quality' product. High quality is used here in the sense of its classic definition – a high level of *fitness for use* - that is, it is free from manufacturing defects and conforms to the design specifications (Montgomery 1991). Morrison suggests that this level of quality can be achieved in dynamic microsimulation models through validation (2008). This validation or 'quality control' ensures that both the internal workings of the model are correct (there are no manufacturing defects) and that the output is reasonable (the design specifications are met). Morrison (2008) provides significant detail of the methods used by DYNACAN to ensure its outputs are valid. He groups validation within DYNACAN into five types:

1. Data, coefficient and parameter validation;
2. Algorithmic validation;
3. Module-specific validation;
4. Multi-module validation; and
5. Policy impact validation.

Clearly, unless each of these areas produces a credible output, the overall model will not be of a high quality and will not be accepted by policy decision makers. The normal metric used by policy makers to decide whether a model is credible or not is the comparison of the simulated outcomes with known external benchmarks. Being able to demonstrate that the model successfully matches a large number of benchmarks greatly increases the chances of the model being accepted and used by policymakers.

APPSIM is focusing on the five validation types suggested by Morrison to ensure a high quality model is delivered to Australian policy makers. In addition, within the module-specific and multi-module validation, the checks have been expanded to cover both cross-sectional and longitudinal checks. This will ensure the outcomes are accurate both from a point-in-time and from a lifetime perspective.

3 APPSIM

3.1 OVERVIEW

The Australian Population and Policy Simulation Model (APPSIM) simulates all of the major events that happen to Australians during their lifetime, on the basis of the probability of such events happening to real people in Australia. The simulated events include death, overseas migration, marriage, divorce, childbirth, ageing, education, labour force participation, retirement, etc. Through these events, people earn income, receive social security, pay taxes and accumulate assets.

APPSIM has a similar structure to most other dynamic microsimulation models – an initial starting population, a simulation cycle and an output. Within the simulation cycle are sets of functions or tables of probabilities for calculating the chances of events occurring.

To calculate the probability of an event occurring, the simulation uses transition probabilities (calculated from equations or tables), based on the person's characteristics, history and the simulated time. As the simulation clock steps through time, the chance of a person transitioning from one state to another is considered (for example changing from the state of 'employed full-time' to the state of 'unemployed'). In the case of a transition between labour force states, the circumstances that influence a transition are the person's age and sex; their labour force status in the previous two years; their education qualifications; whether they are married; the age of the youngest child in the family; whether they are old enough to access their retirement savings; whether they are eligible for a government age pension; and their health at that time.

After calculation of a transition probability (in the range 0.0-1.0), this 'chance' is compared with a random number. Based on the result of this comparison, the transition may be flagged to occur. For example, if person A's chance of transitioning to unemployment in year *t* is 0.015 and a random number of 0.345 is drawn, then person A will not be flagged to transition to unemployment in year *t*. A feature of the model that will be discussed later is that it has the ability to adjust its outcomes to align with external reference data.

The APPSIM model is written in C#, with the simulation reading in the starting population from a Microsoft Access® database and a series of Microsoft Excel® spreadsheets that contain the parameters (that can be readily changed by users) and benchmarking data.

4 ALIGNMENT AND VALIDATION IN APPSIM

A number of tools and features have been developed for APPSIM to ensure that the output is valid from both a cross-sectional and longitudinal perspective. Each of the DYNACAN identified validation priority areas - data, coefficient and parameter validation; algorithmic validation; module-specific validation; multi-module validation; and policy impact validation – are checked. The tools and features also ensure that APPSIM has the ability to track historic data and produce projections that are consistent with official projections.

One of the most common means of ensuring that a model operates within acceptable boundaries is to use explicit alignment procedures. 'Alignment' here means that the model's results are constrained to match external benchmarks.

Before outlining the particular methods used in APPSIM, it is worth acknowledging that there are opposing views on alignment.

The main argument against alignment is that the aligned outcome is not consistent with the original summed micro-results and influences produced within the model. The acceptance that alignment is required is in some cases an acknowledgement that the process being simulated within the dynamic model has been inadequately or wrongly specified. The result is that the outcomes have been compromised and the integrity of the model has been reduced. While not denying the compromise, the argument in defence of alignment acknowledges that it is impossible to completely specify a process and alignment is used to 'repair' the consequences of insufficient estimation data. Within Australia, for example, some six waves of the HILDA panel data are used to estimate many of the transition probabilities within APPSIM. HILDA contains a relatively small sample, of around 7500 households each year, so it is inevitable that the equations produced from it (particularly for relatively rare events such as childbirth) will contain some 'noise'. In addition, by chance, HILDA captures a period of strong economic growth within Australia and it is not certain that the behaviour of individuals estimated from it will continue to be the same during the current economic turmoil. So while in theory, alignment should not be necessary, in reality, the estimation data is never perfect and alignment will always be required (Kelly and King 2001).

In the following paragraphs the tools themselves and how they are used to validate and align the model are discussed.

4.1 VALIDATION TOOLS

The model has a range of tools integrated into its design to assist with the validation process. These include:

- Data structure – a 'strongly typed' database is used and most attributes have a preset range of options;
- User-selected alignment;
- Individual Data Output – an output of every characteristic of a specific individual or a range of individuals and up to ten user-defined other factors at any time;
- Cohort Tracking – an annual output of the characteristics of a user-defined birth cohort;
- Everyone Output – every characteristic of every individual at a range of points in time; and
- Summary Statistics – over 600 summary statistics collected annually.

Data structure

The data structure that is employed within APPSIM is designed to minimise errors and quickly identify errors that do initially get through. The database that underpins APPSIM and contains all of the details on people, families and households is a 'strongly typed' database. This type of database uses names for each column of the database and will only accept the correct type of values. For example, internally, the column of the population database that contains a certain field is referred to as People.Columns[*x*] where *x* is the column number. However, using strongly typed fields, column 44 which contains the flag of whether a person is retired or not uses the label RETIRED to refer to this column and the range of accepted values is limited to Boolean values of 0 or 1. This allows the programmer to refer to the retired flag as People.Retired rather than the very vague People.Columns[44]. The strongly typed field will only accept Boolean values and will reject all others.

Despite the use of strongly typed fields, previous experience with the development of dynamic microsimulation models has shown that undefined characteristics being attributed to an individual are a major source of errors. For example, the gender attribute which should have only two states (1 = male and 2 = female) can be found to be not assigned, given a value of 99 (if unknown) or, worse, it could be accidently incremented (either changing the gender from male to female or creating a new gender "3"). To minimise the possibility of assigning non-existent states to a field, APPSIM always refers to a value through an enumerator list. This ensures only valid states are used. For example, to assign a baby's gender, the 'sex' enumerator list is used. This list has only two values (1 = male and 2 = female) and the coding would be baby.Gender = sex.Male or baby.Gender = sex.Female. By only referring to variable states by the enumerated value, coding mistakes and data inconsistencies are greatly reduced and almost eliminated. A final example to emphasise this point is a change in labour force status from unemployed (3) to employed full-time (1). Traditionally this would be coded as

```
IF person.LFST = 3 THEN person.LFST = 1.
```

If the value 3 (unemployed) was accidently entered as 4 (not in the labour force), it would be very difficult to identify the error. However using enumeration, the code becomes

```
IF person.LFST = lf.unemployed THEN person.LFST = lf.employedFullTime
```

and any errors in the coding (as above, if person.LFST = lf.NILF …) are minimised and easily observed and corrected.

A second major advantage of enumeration is the readability of the code. This readability makes errors easier to identify and broadens the range of people that can work on the model. Rather than coding being checked by one person and logic being checked by another, the processes can be combined and the chances of errors are reduced.

User-selected alignment

During the initialisation of the simulation, the user has the ability to turn alignment on or off. By selecting alignment ON, the outcomes of the module align with external benchmarks and with alignment set to OFF, the outcomes of the module will be purely an outcome of the transition probabilities or equations within the module. While it would be theoretically very nice for the outcomes generated by the regression equations or probability tables to accurately replicate reality, limitations in the data used to model transitions and changing behaviour make this impossible. For APPSIM, most equations are based on the HILDA survey which is only a reasonably small panel survey and only seven waves of data are available (discussed later in the paper).

Individual Output

The Individual Output Tool (IOT) provides an output of every field or characteristic of a specific individual along with a range of user-defined other fields at a specific point in the simulation. A group of people can be output by use of a simple loop mechanism and a person can be tracked over time by repeating the procedure each simulated year. The IOT has two major purposes – the first is the validation of coefficients, parameters and algorithms (DYNACAN steps one and two); and the second is longitudinal validation (DYNACAN steps three, four and five).

The IOT enables each individual in scope for a transition to be output along with the algorithm coefficients and parameters generated within the simulation at that point in time. This output provides a valuable source dataset for validation. In the case of logistic regression equation algorithms every simulated coefficient, parameter and outcome is able to be checked and compared with theoretical or externally calculated values. Similarly, in the case of applying a distribution of outcomes based on a probability distribution, the simulated outcomes for individuals and the overall groups can be compared with theoretical outcomes for those in scope. This actual versus theoretical outcome at an individual or group level provides a very quick and thorough validation that the model is functioning within its specifications.

The IOT ability to track a person or group over time also provides a tool to observe output over time. This capability can be used to ensure that the outputs from each module and from the overall model overall are valid. As a dynamic microsimulation model enables the probability of an event to vary throughout a simulation, IOT can be used to ensure that the simulated individual and group outcomes from an algorithm vary in accordance with the changing probabilities. The IOT can also be used to ensure that the correct individuals are being selected for transition. For instance, while we want a certain proportion of women to have a child in every year, we do not want the same women to be chosen each year. In other words we can use the IOT to observe the lifetime childbirths of a person and ensure these lifetime results are valid as well as the cross-sectional results.

The output provided by IOT is available with the enumerated values discussed above. The plain language output allows simple computerised checking or a visual inspection to be undertaken to quickly identify errors in the data. For example, if the gender column is being checked, the only valid values are male and female. A value like '99' will be easily detected using visual inspection of the data or the production of a frequency table.

Cohort tracking

The group output feature of the IOT is used in conjunction with user-input birth cohort parameters to provide a detailed cohort tracking mechanism. The ability to track a cohort enables 'age', 'period' and 'cohort' effects to be disentangled and validated. [1]

Everyone Output

By expanding the group feature to include everyone, the IOT output can be used to track every individual over their lifetime and validate their life path. Alternatively, by aggregating every individual, the simulation outcomes can be compared with national benchmarks.

Summary statistics

There are currently 672 summary statistics that are output each simulation cycle by APPSIM. This includes 22 population statistics (population by sex and age group), 480 labour force statistics (population by age group, sex, labour force status and highest education qualification) and 170 general statistics for the simulated year *t* (total population; total immigrants; average earned income; births to mothers in various age groups; women aged 30, 40, 50 with parity equal to 0, 1, 2, 3, 4, 5+; total superannuation guarantee contributions; total hospital admissions; etc). The summary statistics are output into a Microsoft Excel® spreadsheet at the end of the simulation run. The spreadsheet also contains external reference data and a baseline simulation output. The spreadsheet combines current simulation output with the baseline and reference data in a series of charts to allow easy checking and analysis of the simulation output.

4.2 Cross-Sectional Alignment

For alignment to occur in APPSIM, external benchmarking data must be available and an 'align' switch must be turned on. Most modules have such a switch. When alignment is switched on, every person in scope for the event is assigned a probability score. Then the in-scope population is ranked based on their score. The appropriate number of people to make the transition are then selected based their ranking to match the external benchmark. This strategy effectively ensures that correct number of persons with the right characteristics experience the event.

APPSIM also adds a degree of randomness into the selection process. The process described in the previous paragraph will never allow a low ranked person to make a transition. However, this does not reflect reality. A person with a low probability of

[1] The 'age' effects are those where we see a distinctive pattern over the life course. For example, fertility rates are low for women under 20 years, high for those in their 20s and 30s, low again in their 40s and zero above 50 years of age. 'Period' effects are those that reflect the conditions prevailing at different times. For example, the returns on investments and taxation rates vary over time. 'Cohort' effects refer to the way in which birth cohorts of the population behave differently. For example, people who lived through the depression and were in their fifties in the 1970s may well have behaved differently to how Gen Y will behave in their fifties in the 2020s.

becoming unemployed (for example, a married, degree-qualified, full-time employed male aged 45 years) will occasionally lose their job and become unemployed. To replicate this randomness, APPSIM allows a proportion of the probability scores to be inverted (that is the probability score used for ranking is subtracted from one [ranking probability = 1 – calculated probability]). With an inverted probability those that would be ranked very low are ranked very high and vice versa. By default the proportion to undergo inversion is set at ten per cent.

As noted earlier in the paper, in an ideal model alignment should not be necessary as it implies that the algorithms in the model have been inadequately or wrongly specified. However, in reality, it is impossible to completely specify a process even when sufficient estimation data are available. For dynamic modelling, the inadequacies of the process are exacerbated by the changing behaviour of individuals over time, the complexity of the modelling, insufficient estimation data and the interdependence of the functions. These shortcomings combine to make alignment an essential component of dynamic microsimulation.

A range of cross-sectional benchmarking data are available for comparing simulated and historical outcomes. These include broad macro or aggregate benchmarks, such as total population and total superannuation balance, and a number of disaggregated group benchmarks, such as number of women in a certain age range giving birth in that period. The majority of these benchmarks are available from the Australian Bureau of Statistics or other reliable sources that are accepted as being the ‘official’ number or rate for that period of time. In addition to historical data, some values or rates also include projections into the future. Again these are generally accepted as being the ‘official’ benchmark and, thus, provide useful standard settings.

The integration of these external benchmarks into the summary statistics makes the checking process quite straightforward.

4.3 LONGITUDINAL ALIGNMENT

Longitudinal alignment and validation require quite different data to the cross-sectional benchmarking data. The longitudinal validation may be considering rates of change over time, the frequency of transitions or the number of transitions in a lifetime. To validate data of this type requires monitoring of individuals or groups of individuals over periods of time. The benchmarking data must also be of this form. The cost and complexity of collecting this type of data means that it is often not available, so other less reliable sources often are used. For example, to establish the number of childbirths a woman has in her lifetime would involve a panel survey lasting 35 years (from age 15 years to age 50 years). However, the tracking and collecting of data for this length of time would generally be prohibitively expensive. The data may also have limited value. As Kelly and King (2001) noted ‘the use of functions based on data which finish before the simulation begins increase the probability that the behaviour predicted by the simulation will not align with actual data’. For example, because a male typically stayed in full-time employment for almost all of their working life in the past, does not mean that this behaviour will still be typical in the future. A model that projected male labour force patterns based on historical behaviour could seriously misrepresent the future situation. In the example of childbirth, the function derived from 35 years of history is probably not a good predictor of the behaviour of women now - and probably of even less value in predicting future behaviour.

Despite these limitations, panel surveys do provide considerable insight into the life courses of individuals and groups of individuals. In Australia, the only major longitudinal survey is the Household, Income and Labour Dynamics in Australia Survey[2] (HILDA). The majority of the transition probabilities and logistic regression equations in APPSIM have been derived using data drawn from this survey.

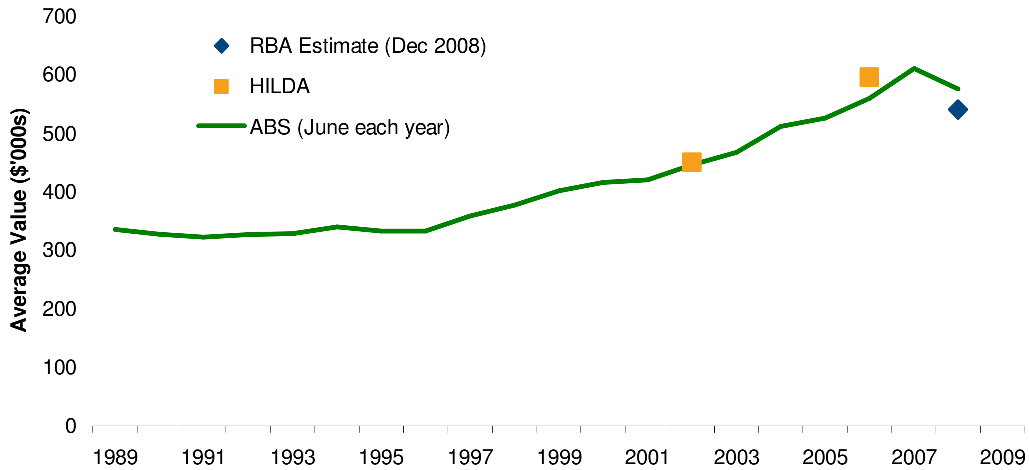
The HILDA Survey is a household–based panel study which began in 2001. HILDA is funded by the Australian Government for at least 12 waves and is managed by the Melbourne Institute of Applied Economic and Social Research. While HILDA enables a range of longitudinal issues to be considered and modelled in APPSIM, it is limited by its small size (wave 1 consisted of 7,682 households and 19,914 individuals); the short life of the survey (only seven waves available in 2009); and that some topics are not surveyed every year (for example, household assets were only surveyed in 2002 and 2006).

A major issue with using the HILDA Survey relates to its short life to date. The small number of waves makes it difficult to separate out the age, period and cohort effects (see footnote 1) from the underlying trends and behaviour. A clear example of difficulty of separating these effects from the underlying trends is in average household net worth. Figure 1 shows a comparison of estimates of average household net worth with a series from the Australian Bureau of Statistics, an estimate for December 2008 from the Reserve Bank of Australia, and 2002/2006 estimates from HILDA. The estimates for various points in time are surprising close.

---

[2] However, there have been a few smaller surveys that collect historical information from respondents. As the historical approach relies on the respondent accurately recalling when an event occurred, it is not generally as accurate as annually surveying participants.

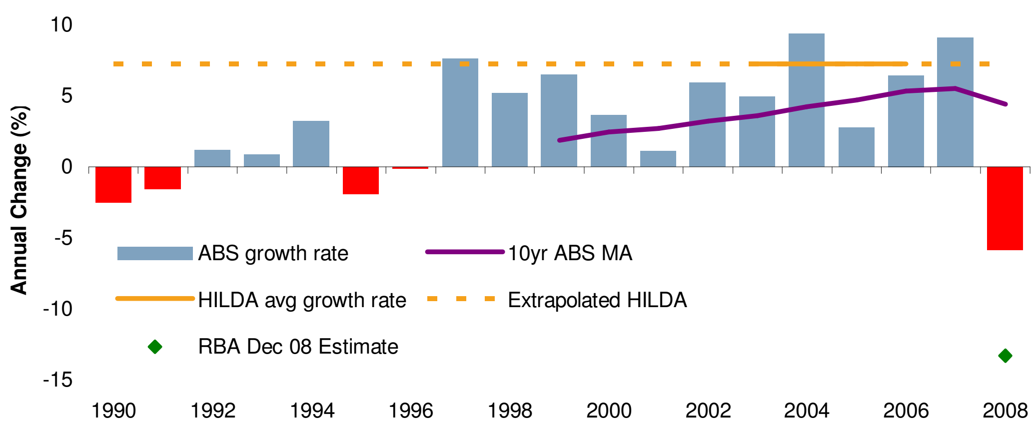
Figure 1 Household net worth, Australia, September 2006 prices



Source: Bloxham and Betts 2009, Graph 1

However, if we now look at the growth rates that produce these historical averages, clear differences appear. Figure 2 shows the annual percentage change in net worth and a 5-yr moving average based on ABS data, the average annual growth rate based on the two HILDA estimates, and the current estimate for the year ending December 2008 from the Reserve Bank of Australia.

Figure 2 Growth rates of household net worth, Australia



Source: Author's calculations based on Bloxham and Betts 2009

As Figure 2 shows, Australia underwent strong growth in the period from 1997 to 2007 averaging 5.7 per cent in real terms. The period between 2002 and 2006 of the HILDA surveys has a real growth rate of 7.2 per cent as it covers some high growth years. It is clear that the use of either a long term average (either the 10-yr moving average or the HILDA estimate) is going to overestimate the growth in net worth in the next decade.

In addition to the difficulties of selecting an appropriate aggregate growth rate, dynamic microsimulation models are trying to capture the changes in underlying distributions. For the distribution of some household assets, the levels of ownership and values can be highly skewed whereas, for other assets, the distribution is reasonably even. A dynamic microsimulation model must capture this variation. This requires projected differential rates for each asset type and projected changes in distributions of ownership and asset values. As Figure 2 shows (even without the complication of different distributions for each asset type), there are significant hurdles to be overcome.

APPSIM attempts to overcome these obstacles by using a large number of disaggregated alignment benchmarks. In the example above, benchmarks would be used (where available) for proportion of households owning each asset type and the average value by age, gender, income and household type, growth rates for asset type by value, and location. Aggregate benchmarks would also be used to validate the overall outcomes.

The summary statistics and the group feature of the IOT enable the multitude of output parameters to be monitored and validated.

5 Modelling Superannuation and Social Security in APPSIM

In this section details are provided of the functioning of the superannuation and social security modules of APPSIM. These explanations will enhance the understanding of the example of the usefulness of a validated and aligned dynamic microsimulation model in the next section. Explanations of the overall operation of APPSIM and more details on the APPSIM modules can be found in the APPSIM Working Paper series available at www.natsem.canberra.edu.au.

5.1 Modelling Assets in APPSIM

Previous research (Kelly 2007a) has shown the dominant forms of accumulating household wealth in Australia are through the family home and superannuation. As residential property accounts for 60 per cent of the total net worth and superannuation accounts for a further 15 per cent (Bloxham and Betts 2009), capturing just these two assets will capture the majority of the average household's net worth. Other assets that will be included in the final version of the APPSIM model include cash holdings, share portfolios, investment properties and net business assets. These additional assets do not have a major impact on the average levels of wealth but, due to the uneven distribution of some assets, they do have significant values for some groups within the Australian community. For example, the median level of equities owned was $18,000 in 2006 - but only eight percent of the poorest one-fifth of households owned equities and the median value for those with equities was $3,000. At the other end of the spectrum, 72 per cent of households in the richest quintile own equities with a median value of $55,000 (Bloxham and Betts 2009). It will be important to capture these skewed asset distributions in the final model.

Other assets, such as vehicles and contents of the family home, can be significant but the modelling of these assets is somewhat more complex and does not impact significantly on the distribution of wealth. The complexity comes from the continually changing nature of these assets. In addition, they are assets that often depreciate rapidly with time and are of little importance to the building of long term wealth.

Projecting the current value of an asset requires estimates to be made at specific points during the simulation. First, initial values must be assigned to the starting population and to immigrant families as they enter the simulation. Second, on an annual basis, changes in the ownership of each asset and valuations need to be estimated. Finally a range of special events need to be handled — for example, the transfer and dispersion of assets on the death of a person or the new ownership of joint assets after the divorce of a couple.

In this prototype version of the model, only housing and superannuation are fully modelled. Other assets have not been modelled in this prototype version; however, investment income has been included to ensure that pension entitlements can be estimated.

The modelling of superannuation assets is discussed in the following paragraphs.

Superannuation

For this IMA-prototype version of APPSIM, only 'defined-contribution' or 'accumulation fund' superannuation is implemented. (This means, for example, that we have not simulated any post-2001 increases in the 'defined benefit' superannuation rights of those still working - particularly significant for some public servants.) As 85 per cent of superannuation funds in Australia are of this type, the omission of defined benefit funds is not believed to be a significant limitation to the broad findings of the paper.

The simulation of the accumulation or drawdown of superannuation in APPSIM is done by estimating contributions, investment returns and draw-downs and applying these to each individual account. For each simulated year up to 2006, the superannuation balances for every person aged 15 and above are estimated and imputed using two different techniques. Both techniques – regression equations and probability tables – use the ABS 2005-06 Survey of Income and Housing as the source data. For those that are not retired, two regression equations were developed – one for men and one for women. The non-retired regression equations are influenced by age, earnings, whether the person holds a degree, labour force status, the person's household tenure, and whether the person is self-employed.

For those that are retired, a probability table was shown to give a better distribution of superannuation. The reason for this was the skewed distributions of superannuation, with large numbers of people having a zero balance and a few having very large balances. The tables were broken down by sex, whether the person had a degree, and age. Estimated values were provided for the percentiles of 0%, 25%, 50%, 75%, 90%, 95% and 100%.

The equations and tables estimate superannuation balances in 2006. For years before 2006, a deflator is applied to the estimated balance. However, superannuation coverage is not reduced in the years prior to 2006 and, as a result, the model will slightly overestimate coverage, and as a consequence the aggregate level of superannuation, prior to 2006.

In all years, superannuation guarantee (SG) contributions were estimated based on annual earnings and added to the superannuation account. The SG rate is a user -defined variable within APPSIM, but is set to nine per cent by default (the current mandatory rate set in legislation). For every individual aged over 15 years and below the maximum contribution age (currently 75 years) with earnings above a minimum value ($450 per month), an employer contribution is calculated using their annual earnings and the SG rate.

In Australia, a number of employers (particularly government) contribute at higher rates than are mandated by the SG and, to simulate this, one-fifth of SG superannuation contributions have been doubled. The selection of individuals to have the higher SG contributions is decided using an assigned lifetime random number. This ensures that the same individuals continue to receive higher contributions.

To simulate extra voluntary or 'member' contributions in addition to the SG, a probability distribution based on ABS superannuation member contributions is applied (ABS 2008). The probability tables are based on age and sex. They range from 4.3 per cent of females aged 15-24 years to 37.6 per cent of males aged 65-69 years making extra contributions. The proportion of their income that they contribute is again based on ABS survey data and, based on age, varies from 8.5 per cent to 28.1 per cent.

As people are less likely to make voluntary contributions into superannuation when the funds are performing poorly, a feature of the model is that member contributions are reduced when superannuation returns in a simulated year are negative or zero. The reason all member contributions are not set to zero is that a number of funds have mandatory member contributions. In the baseline scenario, where historical and long term averages are used, a negative return only occurs in 2008 (so all other years are positive). The model also has a one year 'spike' in member contributions for the year 2007, to replicate the actual spike that occurred due to changes in superannuation taxation in that year.

The return on all superannuation funds is assumed to be the same and set to historical net return values for a 'balanced' fund up to 2008 and then at the 10 year average (4.63 per cent) for projections. The returns were all adjusted using the CPI to give real 2009 values. This return was added to all superannuation balances on an annual basis.

For those in retirement, an annual drawdown also takes place. According to the Westpac ASFA Retirement Standard Calculator (https://info.westpac.com.au/rlscalculator/) a single person currently requires between $18,500 (modest) and $36,500 (comfortable) per annum depending on the living standard required. Using these estimates as a guide, the model assumes that those under the minimum age pension age drawdown $20,000 per annum. For those over the minimum pension age, the drawdown is assumed to be only $10,000, as it will generally be used to supplement the government pension. For those over pension age and with a high superannuation balance, a more comfortable living standard might be expected while funds are available – and, hence, the drawdown is the minimum of five per cent of the superannuation balance or $10,000. The value of five per cent was selected as it would see the superannuation balance lasting at least 20 years and is approximately in line with the government minimum drawdown guidelines.

For the scenario modelled with the IMA prototype version of the model, the global financial crisis is assumed to have produced nominal annual returns from superannuation of zero in the years 2009, 2010, 2011 and 2012.

5.2 Modelling Social Security in APPSIM

STINMOD is NATSEM's static microsimulation model of the Australian income tax and cash transfer system (Vu, 2008). This publicly available model provides estimates of the distributional, revenue and expenditure impacts of taxation and transfer policies on Australian individuals and families. STINMOD is now a standard model used by Australian government departments for their analyses of existing and possible policy options in the tax/transfer field (e.g. Bremner, 2005) and it is consistently used by NATSEM to examine the immediate impact of policy change (Lloyd, 2007; Harding et. al., 2005).

For this IMA prototype, an innovative approach was taken of generating a 20 per cent sample of output records for the base year (2009) and four future 10 year periods (2019, 2029, 2039 and 2049) from APPSIM and then running that sample against the STINMOD income support and taxation code. The STINMOD income support code (which is written in SAS) includes a replication of the current Australian age pension payment parameters and implements the age pension income test and age of eligibility test for singles and couples.[3] This version of the modelling assumes that those of age pension age meet the residency test. A simplified version of the age pension assets test is also simulated. If a person's assets are below their relevant asset test cut-off point (which varies by couple/single and homeowner/non-homeowner status), then they are assumed to pass the assets test and only their income is used to determine how much age pension they are simulated to receive. In the out-years, the rate of pension and all other pension parameters are held fixed at their 2009 level (as earnings and other income sources are similarly held fixed in constant 2009 dollars).

6 Using APPSIM

In this section, a policy example is used to demonstrate the usefulness of dynamic microsimulation and the value and operation of alignment procedures.

The usefulness of the modelling is totally reliant on the outcomes being considered to be credible. As the paper has also noted, simulating a policy change is often the only way that the individual and joint effects of the all the processes in the model can be tested. In this case the baseline outcomes are used to provide projections that should be both valid and align with official estimates.

[3] The modelling of the Age Pension system in this paper was undertaken before the changes announced in the May Australian Government Budget to the pension rates, eligibility age and taper rates.

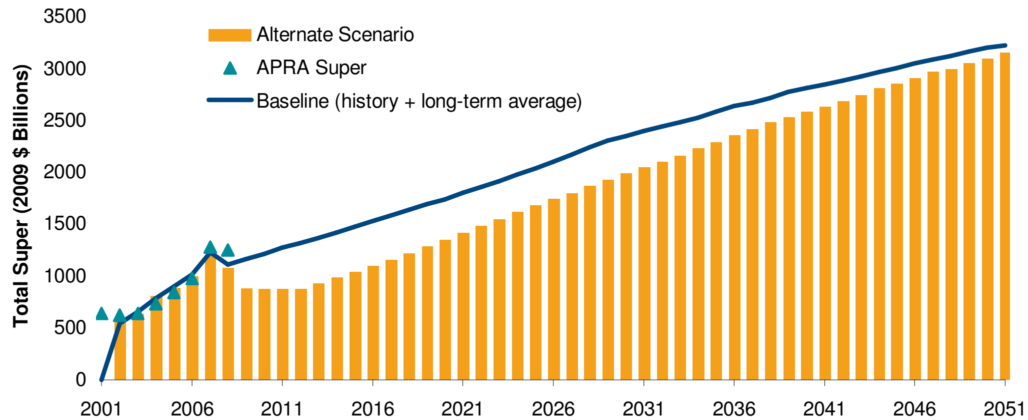
To ensure the baseline scenario outcomes are aligned, contributions and returns from superannuation are aligned with historical values (2001-2008) and projections from 2009 onwards are based on long term averages. In addition, the population, number of people qualifying for an Age Pension and Age Pension entitlements are also modelled and aligned within the simulation. The baseline scenario includes the negative 6.4 per cent (nominal) performance of the superannuation funds in the 2008 financial year.

In the alternate scenario, APPSIM is used to project the impact of the current global financial crisis on retirement savings and related government Age Pension outlays. In this alternate scenario, called the *Financial Crisis* scenario, voluntary contributions to superannuation and the returns received from superannuation are depressed for a five year period from 2008 to 2012. The actual superannuation return is used in 2008 (-6.4% nominal) and zero (nominal) is used for the 2009-2012 years. From 2013 superannuation returns are set to the long term average.

The scenarios presented are hypothetical ones. Firstly, APPSIM is still in development and a number of its outputs have not been completely validated. It is worth noting that since these scenarios were run, a number of changes and improvements have been made to the model. Secondly, the scenario is an obvious over-simplification of the real environment in Australia. While superannuation investment return parameters have been altered in this modelling, the impact on earnings, unemployment levels, returns of other asset types and the negative returns projected for 2009 has not been modelled. Finally, the impact on the Government's Age Pension outlays are only considered in regards to changes in superannuation assets and income. The general declines in asset values have not been included in the modelling.

Initial validation and alignment of the baseline scenario is done by comparing the simulated total population, number of age pensions, government age pension outlays, SG contributions and total superannuation balances with historical data and projections from the Australian Bureau of Statistics, the Australian Prudential Regulation Authority (APRA) and the Department of Families, Housing, Community Services & Indigenous Affairs. Individual superannuation balances and contributions are validated by age, gender, type of employment, labour force status and education against the 2005-06 ABS Survey of Income and Housing data file.

Figure 3 Estimated total superannuation under a baseline scenario and a 'financial crisis' scenario, Australia, 2001-2051



Source: Preliminary output from APPSIM, see text for details.

Figure 3 shows that the simulated superannuation data aligns reasonably well with historical ARPA estimates. It also shows that four more years of poor returns (2008 was another bad year but is included in both the baseline and the financial crisis scenarios) will have an impact on the Australian population for at least the next forty years.

Table 1 Simulated changes in Age Pensioner numbers, Government Age Pension outlays and Total Superannuation, Australia, selected years

	2009	2019	2029	2039	2049
Age Pensioners – Baseline	100.0	131.0	162.1	181.7	206.5
Age Pensioners – Financial Crisis	99.5	133.7	162.7	183.9	202.9
Difference	***-0.5***	***2.7***	***0.7***	***2.2***	***-3.7***
Gov't AP Outlays – Baseline	100.0	134.3	166.6	187.5	218.5
Gov't AP Outlays – Financial Crisis	100.7	137.7	168.5	204.1	216.5
Difference	***0.7***	***3.5***	***1.9***	***16.6***	***-1.9***
Total Super – Baseline	100.0	154.2	226.7	283.0	328.2
Total Super – Financial Crisis	75.8	117.2	187.1	269.6	318.0
Difference	***-24.2***	***-37.0***	***-39.5***	***-13.5***	***-10.1***

Note: The values shown are a percentage of the 2009 Baseline value.
Source: Preliminary output from APPSIM, see text for details.

Under the financial crisis scenario, total superannuation is immediately reduced by an estimated 24 per cent due to the negative returns on the assets already invested and the lower than normal member contributions (Table 1 and Figure 3). The poor performance combined with its compounding effect results in the greatest deficit (almost 40 per cent) occurring 20 years after the crisis. However, the superannuation deficit does not result in large increases in Age Pensioner numbers. The reason for this is that compulsory superannuation was only introduced in 1992 and by 2009 superannuation balances are generally quite low. This means a change in the superannuation balance is unlikely to drive a person over the 'free' asset threshold of the pension means test.[4] The superannuation balances of these baby boomers and its impact on the government outlays are the topic of many research papers (for example, see Kelly and Harding 2004, 2006, 2007).

Table 2 Simulated differences in average superannuation outcomes under the baseline and financial crisis scenarios by age group, 2009 and 2019

Age Group	2009	2019
	%	%
15-24	**-1.7**	1.4
25-34	**-5.1**	**-16.8**
35-44	**-7.4**	**-19.2**
45-54	**-2.0**	**-25.4**
55-64	0.2	**-25.9**
65-74	4.4	**-20.3**

Note: The values shown are the difference between the financial crisis value and the baseline value.
Source: Preliminary output from APPSIM, see text for details.

Table 2 and Table 3 compare the outcomes between the two scenarios for a range of individuals. In Table 2 the financial crisis scenario outcome is compared with the baseline scenario in 2009 and 2019 for a range of age groups. The baseline and financial crisis scenarios differ in their investment returns over the period 2009-2012 and the impact is already beginning to be evident in most age groups in 2009 and is clearly evident in all age groups (except the 15-24 age group) in 2019. Table 2 shows that superannuation balances are down by 17 to 26 per cent seven years after the financial crisis has finished.

Table 3 Simulated differences in accumulated average superannuation by age 55-64 under the baseline and financial crisis scenarios

Age in 2009	Birth Cohort	Difference at age 55-64
		%
15-24	Born 1985-1994 Generation Y	1.9
25-34	Born 1975-1984 Generation X and Y	**-3.7**
35-44	Born 1965-1974 Generation X	**-20.1**
45-54	Born 1955-1964 Baby Boomers	**-25.9**
55-64	Born 1945-1954 Baby Boomers	0.2

Note: The values shown are the difference between the financial crisis value and the baseline value.
Source: Preliminary output from APPSIM, see text for details.

Table 3 considers the impact of the simulated financial crisis on the pre-retirement superannuation balances of different birth cohorts. The table shows that the worst impact will be felt by those Baby Boomers aged 45-54 years in 2009. The simulation estimates that

---

[4] In March 2009, the free threshold under the assets test for a single homeowner is $171,750; for a couple homeowner it is $243,500; for a non-homeowner single it is $296,250; and, for a non-homeowner couple it is $368,000.

this cohort will have a pre-retirement superannuation balance that is one quarter (25.9%) lower than under the baseline scenario. The Generation-Xers born between 1965 and 1974 are marginally better, being down 20 per cent.

7 CONCLUSION

The results produced by dynamic microsimulation must be seen as credible if they are to be accepted and used by researchers and policymakers. This paper presents the validation and alignment steps being taken in the development of APPSIM to ensure that the output is credible.

The paper detailed the tools and techniques that are used within the model to ensure that the outcomes are valid. Then the paper noted that simulating a policy change is often the only way that the individual and joint effects of the all the processes in the model can be tested and problems which otherwise would not be apparent, can be seen. Accordingly, the paper used the simulation of an economic downturn to test APPSIM in this way.

The first step was to show the value of alignment as a validation tool by allowing a baseline scenario to be benchmarked to previous projections. For APPSIM, the very early results from this new model were able to show strong congruence between the simulated baseline scenario aggregate and individual outcomes and the known benchmarks. The second step was to compare of the baseline with an alternative policy scenario. In this case the alternate was changed superannuation and pension outcomes under a financial crisis. Through this scenario alignment and modelling it was possible to show the power of the dynamic microsimulation. In this case, the modelling demonstrated it is possible to isolate the effects of the policy shock by using two future worlds that aligned perfectly except for the introduced change.

In this demonstration we showed the impact of the hypothesized shock as causing a loss in taxation revenue and an increase in government outlays due to more people qualifying for the Age Pension. The results also showed the projected impact at an individual level on different groups in the community.

Finally, the methodology used in the demonstration could just as effectively have been used to show the different outcomes under a range of alternative policy scenarios.

REFERENCES

Australian Bureau of Statistics 2008, *Employment Arrangements, Retirement and Superannuation, Australia*, Cat No. 6261.0, Canberra, November.

Bacon B. and Pennec S. 2007a, *APPSIM – Modelling Family Formation and Dissolution*, APPSIM Working Paper No. 4, National Centre for Social and Economic Modelling, University of Canberra, September.

Bacon B. and Pennec S. 2007b, *Model Validation and Alignment: Microsimulation and Macrosimulation*, APPSIM Technical Note No. 4, National Centre for Social and Economic Modelling, University of Canberra, unpublished.

Bloxham P. and Betts T. 2009, 'The composition and distribution of household wealth in Australia', *RBA Bulletin*, Reserve Bank of Australia, Sydney, April, p.5-23.

Bremner, K. 2005, 'Net tax thresholds for Australian families', *Economic Roundup*, Winter, Department of the Treasury, Canberra (available from www.treasury.gov.au).

Harding A. 2007, *Challenges and Opportunities of Dynamic Microsimulation Modelling*, Plenary paper presented to the 1st General Conference of the International Microsimulation Association, Vienna, 21 August.

Harding A. and Gupta A. 2007, 'Introduction and Overview' in Harding, A and Gupta, A (eds), *Modelling Our Future: Population Ageing, Social Security and Taxation*, Chapter 1, International Symposia in Economic Theory and Econometrics, Volume 15, Elsevier BV, Amsterdam, pp. 1-29.

Harding A, Vu Q. N, Percival R and Beer G. 2005, 'Welfare-to-Work Reforms: Impact on Sole Parents', *Agenda*, Volume 12, Number 3, pp.195-210.

Harding A., Kelly S., Percival R. and Keegan M. 2009, *Population Ageing and Government Age Pension Outlays - Using Microsimulation Models To Inform Policy Making*, Paper for the Economic and Social Research Institute International Collaboration, Cabinet Office of Japan.

Keegan M. 2007, *APPSIM – Modelling the Labour Force*, APPSIM Working Paper No. 6, National Centre for Social and Economic Modelling, University of Canberra, December.

Keegan M. and Thurecht L. 2008, *APPSIM – Modelling of Earnings*, APPSIM Working Paper No. 11, National Centre for Social and Economic Modelling, University of Canberra, June.

Kelly S. 2007a, 'Self Provision in Retirement? Forecasting Future Household Wealth' in Harding, A and Gupta, A (eds), *Modelling Our Future: Population Ageing, Social Security and Taxation*, Chapter 15, International Symposia in Economic Theory and Econometrics, Volume 15, Elsevier BV, Amsterdam, pp. 387-405.

Kelly S. 2007b, 'The Australian DYNAMOD Model', in Gupta, A and Harding, A (eds), *Modelling Our Future: Population Ageing, Health and Aged Care*, Model 2, International Symposia in Economic Theory and Econometrics, Volume 16, Elsevier BV, Amsterdam, pp. 439-442.

Kelly S. 2008, *Australian Trends in Wealth Inequality: Household Projections Using Microsimulation*, VDM Verlag, Germany.

Kelly, S. and Harding A. 2004, 'Funding the retirement of Baby Boomers', *Agenda*, Vol. 11, No. 2, Australian National University, Canberra, pp. 99-112.

Kelly, S. and Harding A., 2006, "Retirement Incomes and Australia's Ageing Population", in Bhajan Grewal and Margarita Kumnick (ed), *Engaging the New World: Responses to the Knowledge Economy*, Chapter 15, Melbourne University Publishing, Victoria, pp. 357-373.

Kelly, S. and Harding A. 2007, 'Baby Boomers – Doing it for Themselves', *Australian and New Zealand Institute of Insurance and Finance Journal*, Vol 30 (4), August/September, Matheson Publishing, Victoria, pp. 16-19.

Kelly S. and King A. 2001, 'Australians over the coming 50 years: providing useful projections', Brazilian Electronic Journal of Economics, Vol. 4, No. 2, December, pp.1-23.

Lloyd, R. 2007, 'STINMOD: Use of a Static Microsimulation Model in the Policy Process in Australia', in Harding, A. and Gupta, A. *Modelling Our Future: Population Ageing, Social Security and Taxation*, (eds), International Symposia in Economic Theory and Econometrics, Elsevier, Amsterdam.

Montgomery D. 1991, *Introduction to Statistical Control Second Edition*, John Wiley & Sons, Canada.

Morrison R. 2008, *Validation of Longitudinal Models: DYNACAN Practices and Plans*, APPSIM Working Paper No. 8, National Centre for Social and Economic Modelling, University of Canberra, March.

Pennec S. and Bacon B. 2007, *APPSIM – Modelling Fertility and Mortality*, APPSIM Working Paper No. 7, National Centre for Social and Economic Modelling, University of Canberra, December.

Vu, Q. N. 2008, *STINMOD Interface User Guide V06b*, National Centre for Social and Economic Modelling, University of Canberra (available from www.natsem.canberra.edu.au)