Australian Population Association

12th Biennial Conference

population and society: issues, research, policy

15-17 September 2004 - Canberra, Australia

Models of the income, expenditure and assets of Australians

Richard Cumpston (Richard_Cumpston@cumsar.com.au)
Cumpston, Sarjeant, Truslove, Pty Ltd

Abstract

Income, expenditure and asset models are needed on a national basis for many policy questions, such as the adequacy of retirement incomes. Similar models can be needed for small areas, for example to plan service networks. Often these models are needed for households as well as individuals.

After allowing for non-reporting and under-estimation, census data provide a reasonable picture of incomes and housing expenses, for areas down to collection districts. ABS income and expenditure surveys provide approximate estimates, for much broader geographic regions. Taxation statistics provide income details for postcodes. Unimproved site values and dwelling sale prices are available from state authorities, at varying geographic levels and prices. Superannuation assets are available from APRA statistics, with more detail available from industry surveys.

Census sample files provide a reasonable basis for projections dealing with national policy issues. To help local decisions, synthetic individuals and households can be created from ABS census collection district community profiles. Ideally, households should be located with Cartesian co-ordinates.

As data can be very expensive, it is important to maximise the use of existing data. Job and household changes have many causes, and are largely unrecorded. The available cross-sections and fragments can be used to estimate behavioural models, allowing objective decisions about the value of further data. Models need to be validated by comparing projections with recent experience.

This paper suggests some of the potential users of model outputs, and ways in which their needs can be met.

Paper prepared for the 12th Biennial Conference of the Australian Population Association, 15-17 September 2004, Canberra.

Needs for models

Some of the needs for models are

- Policy issues
- Service networks
- Investment decisions
- Marketing strategies

These needs can have different time and geographic perspectives, and require different types of data. Some examples follow.

Retirement incomes policy

Dunsford & Rice (2004) discussed 12 problems in Australia's retirement incomes system, and suggested a variety of solutions. Desirable characteristics of a model intended to evaluate these proposals include

- ability to model assets and incomes for individuals and couples
- estimates for separate geographic regions, at least down to capital/other regions in each state
- ability to look at different life stages, from first employment through to residential care for the aged
- estimates of the effects on individuals, employers and governments
- realistic assumptions about changes in individual occupations and earnings
- replication of present superannuation assets and retirement incomes
- allowances for known future changes (such as increasing age pension entitlement ages)
- ability to simulate the likely behaviour of individuals in response to tax and social security changes
- credibility from public exposure and practical applications
- ability to update baseline data with new data from censuses and other sources
- long time scales – Treasury's Intergenerational Report 2002-03 is legislatively required to be for 40 years, but a longer period would be helpful.

Service networks

Many different organizations have distributed service networks – schools, religions, banks, police, fast food chains, hardware stores. Some of the issues in managing service networks are

- optimal use of staff and buildings
- good accessibility to clients
- co-location with other relevant services
- minimum client numbers for viability
- responsiveness to future changes.

The development of Canberra provides some good examples of service network change. Much of Canberra's growth occurred through the planned release of land, each area intended to house about 3000 persons, with associated schools and local shopping centres. The initial residents were largely young families, and as their children grew up

the need for local school places sometimes fell to about one-fifth of the peak demand. Many schools and local shops have been converted to other uses.

Good geographic detail is essential for projections to help service network planning. Postcodes and census collection districts are unlikely to provide sufficient detail, particularly for developing areas. Latitude and longitude may provide the necessary accuracy, and remain immune to boundary changes.

Investment decisions

Some of the investment decisions which may be helped by models of the future are

- dwelling purchases
- commercial developments (shopping centres, tourism facilities)
- resource-based projects (such as biomass to methanol)
- infrastructure (trains, roads, sewerage).

Fine levels of geographic detail, and long time scales, may be relevant to most of these decisions.

Marketing strategies

A broad process used in many types of marketing is

- target particular prospects
- choose appropriate incentives
- select promotional channels.

Australia's major political parties have detailed computer data bases of the enrolled voters in each marginal electorate. They may try to make contact with each new arrival, and will have a good idea of the age and socio-economic composition of the electorate. Rod Cameron referred on 13/9/04 to

> "the big money spent on building up databases of targeted voters (neatly classified by demography, affiliation and interests)"

Local issues may be identified by surveys, and policy proposals tested on focus groups. Promotion will be through national and local media, together with local meetings, and direct communication through email, phone and post. A policy such as a tax rebate for older workers may be put in national terms, but is likely to have been framed for the marginals.

Data sources

Some of the potential data sources are

- Australian Bureau of Statistics censuses and surveys
- Australian Taxation Office
- Family & Community Services
- Health Insurance Commission

- Reserve Bank of Australia
- Australian Prudential Regulatory Authority
- State property sale data
- Planning authorities and councils
- Longitudinal surveys – LSAY and HILDA
- Special purpose surveys
- Client data.

ABS data charges

Cells	A	B	C	D	E
1000	40	40	40	40	9
3000	100	40	40	40	9
6000	140	40	76	40	45
25000	520	116	304	116	273
50000	895	191	529	191	498
100000	1095	231	649	231	618
500000	1895	391	769	391	738
2000000	3395	691	1069	691	1038
Maximum	15500	15500?	15500	15500	15500

The above table, taken from a "Charging matrix from automated system for determining charges applicable to ABS customised tables", shows five of the data cell charge bases in use by ABS in March 2004:

A	Consultancy
B	Standard infrastructure formula
C	Census community profile 1991 & 1996
D	International trade infrastructure formula
E	Census community profile 2001

In addition, ABS labour charges, depending on the staff involved, are between $44 and $160 an hour, for short term consultancies. A minimum charge of about $195 applies for any data request.

Government policy on cost recovery by information agencies

The Productivity Commission recommended in August 2001 that

> "Information agencies and the Government should together define a basic product set according to: public good characteristics, significant positive spillovers and other Government policy reasons. The basis product set should be funded from general taxation revenue. Additional information products should be classified into three broad categories and priced accordingly:
>
> - Dissemination of existing products at marginal cost;
> - Incremental products (which may involve additional data collection or compilation) at incremental (avoidable) cost; and
> - Commercial (contestable) products according to competitive neutrality principles."

This recommendation was accepted by the Government in December 2002, in its "Commonwealth cost recovery guidelines for information agencies", which said

> "Incremental products are those for which additional work has been undertaken to modify taxpayer funded information to meet the demands of a specific client group. The additional work could involve:
> - extending a data collection
> - expanding research to cover new issues; and/or
> - undertaking additional analysis or manipulation of the information.

The guidelines recommend that cost recovery for incremental products be based on the incremental cost of the additional product.

It is difficult to see how the long-standing ABS "infrastructure charges", calculated in various non-linear ways from numbers of data cells, can have any connection with incremental costs. Given the great unhappiness within government, business and universities about ABS fees, this issue will resurface after the election.

Other data costs

A whole range of costing policies seem to be in use by Commonwealth and state agencies. The Australian Taxation Office seems to have a policy of charging $800 for moderate data requests, but will sometimes provide data free. The Health Insurance Commission provides many useful statistics for free on its data base, but recently quoted $3,065 for an analysis of five years of Pharmaceutical Benefits Scheme expenditure by age and sex. State agencies are used to charging fees for individual data requests on properties, and can quote extremely high fees for bulk data. Some data may be restricted to particular users, such as real estate agents.
Across the Commonwealth and states, there seems to be an accidental policy of setting fees at the level a few key users will bear. A special purpose survey, put by phone to 2000 adult Australians, may cost about $35,000 to $40,000 (less if added to an existing survey, more if the questions are open-ended).

Maximising use of existing data

As data can be very expensive, it is important to maximise the use of existing data. For example, household and job changes reflect many different events, for which no data are routinely collected (apart from limited data on births, deaths, marriages and divorces). But the Australian census does capture a picture of almost every household each 5 years, as well as data on changes in address in the last year and the last 5 years. For research purposes, confidential unit record files provide a 1% sample of the census, and unit records from some regular surveys. The biennial labour mobility survey estimates the number and broad nature of job changes, and the duration of current employment. These cross-sections and fragments can be used to estimate behavioural models, which may allow objective decisions about the value of further data.

Deterministic versus stochastic projections

Population projections by ABS, the Australian Treasury and state planning authorities are deterministic. For example, assumptions are made about mortality rates in future years, and used to estimate the proportion of persons of each age dying each year. For example, in some of the population projections it published in 2003, ABS assumed that

the mortality rate of females aged 60 in 2047 would be about 0.004. For every 1000 females age 60 in 2047, 996 were assumed to be still alive in 2048.

An alternative approach is to randomly simulate the various possible outcomes for each person in each time period. For each female aged 60 in 2047 a random number between 0 and 1 is generated, and the female assumed to die if the random number is below 0.004. Also known as microsimulations or "Monte Carlo" simulations, stochastic simulations are widely used by actuaries looking at insurance mechanisms. Their use for socio-economic projections was suggested by Orcutt in 1957, and implemented for 4,580 households by Orcutt, Greenberger, Korbel & Rivlin (1961).

The same probabilities are needed for deterministic and stochastic population projections, and they should yield approximately the same results. The results of each stochastic projection differ randomly, and many projections may be needed to get a reasonable picture of the range of possible outcomes. Our actuarial firm routinely simulates 5000 insurance years for proposed schemes, but demographic events have less extreme distributions than insurance claims, and far fewer simulations are likely to be needed.

The great advantage of stochastic projections is that they provide much more flexibility. There is no need to calculate and store all the possible combinations of each variable for each individual. Household relationships can be readily included. Walker (2003) describes NATSEMS's DYNAMOD-2 model

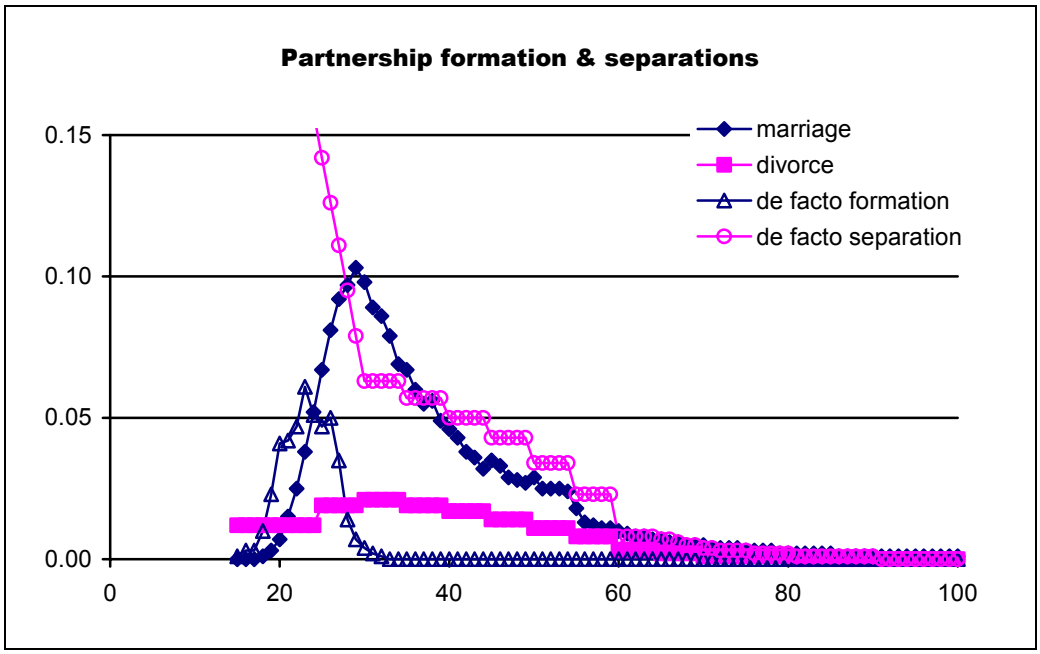
> "The simulations account for future events occurring in the lives of persons in the model's base population – such as couple formation, birth of a child, education, leaving home, migration, divorce, being employed, income from work and government, wealth accumulation, becoming disabled, recovering from disability and death."

Data for stochastic models

Walker noted that DYNAMOD was based on a 1 per cent sample from the 1986 census. The sample file from the 2001 census provides data for 48 geographic areas, each of at least 250,000 persons (except for the Northern Territory). This level of geographic detail may be adequate for many policy applications. If finer geographic detail is needed, then it is feasible to synthesise individuals and households from census data for census collection districts. These synthetic objects are not intended to match any real individuals, but should provide statistically valid starting points for long-term projections. For each local area, the synthetic individuals and households should replicate the observed totals for the area, and should broadly follow national patterns (for example, of age differences between partners). Surprisingly, data synthesis becomes quicker and more reliable as area sizes are reduced.

Stochastic models need assumptions about the probabilities of different types of events. Some examples follow.

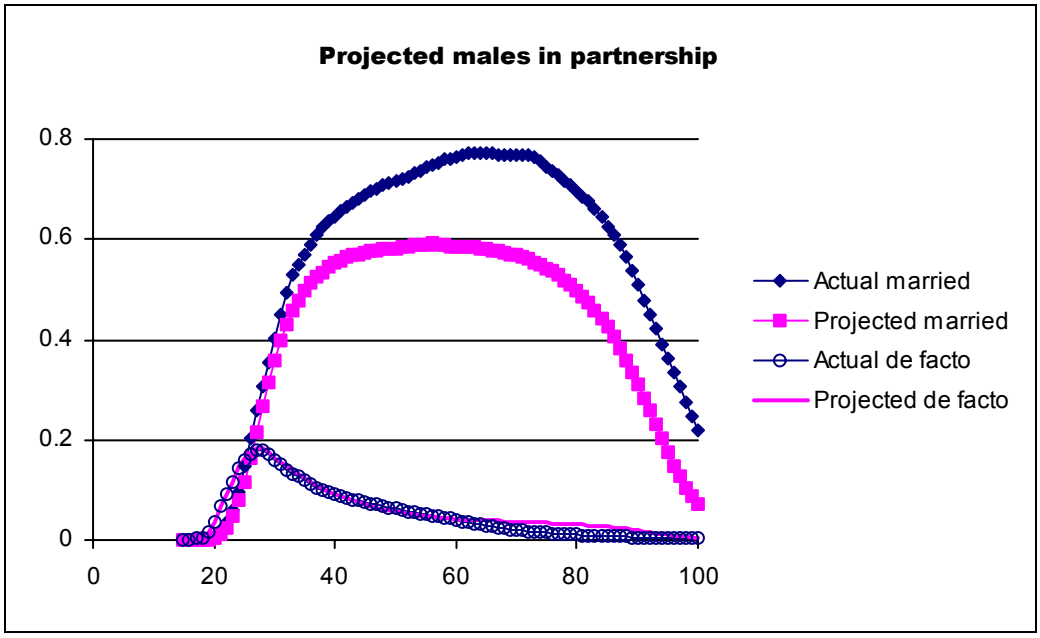
Partnership formation and separation



Sources for the above partnership and separation rates are

- Marriage and divorce rates from ABS (2002)
- De facto formation rates to approximately replicate the observed numbers of persons in de facto partnerships
- Guesses at de facto separation rates.

Projected males in partnerships



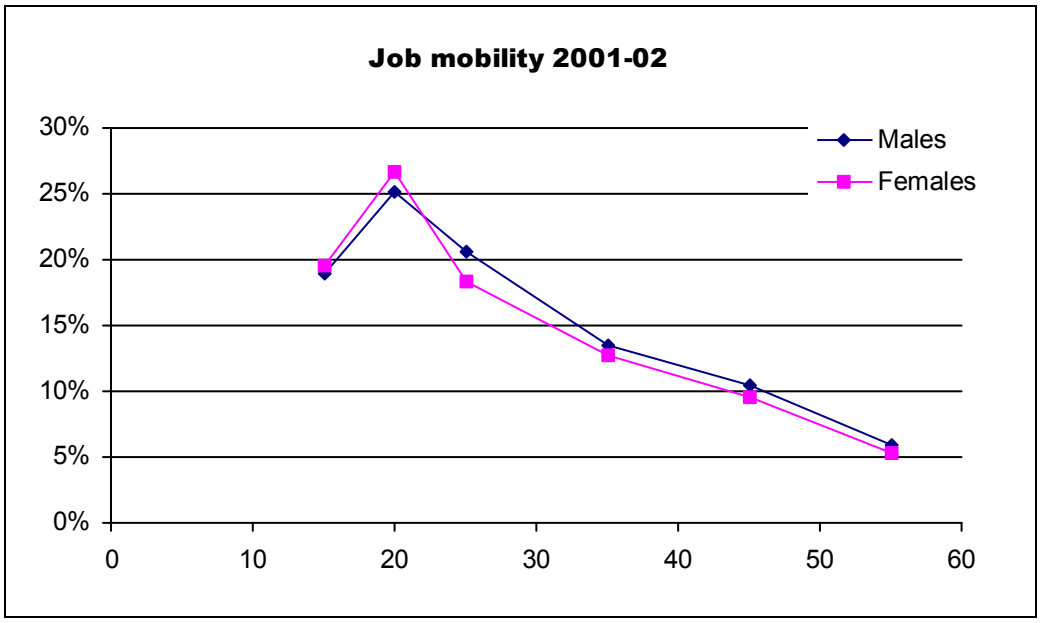
If present marriage and divorce rates continue, only about 59% of males aged 60 will be in registered marriages, compared with 77% in 2001. It is not clear that there will any compensating increase in de facto partnerships at older ages.

Mortality rates of married persons

Numbers of deaths are from "Deaths Australia 2002", and similar earlier ABS publications. Deaths with "not stated" marital status were treated as not married. Expected deaths were estimated by multiplying numbers in registered marriages at the 2001 census by mortality rates from the “Australian Life Tables 2000-2002" (Australian Government Actuary 2003).

The lower mortality of married persons is clear, particularly for young adults. Persons considering marriage are clearly very effective at selecting healthy spouses, and marriage itself may have health benefits. The erratic ratios of actual to expected deaths at ages 80+ may reflect deficiencies in the data on marital status, particularly for females.

Job mobility



The above graph, taken from table 6 of "Labour mobility Australia February 2002" (ABS 2003) shows that job mobility peaks for young adults, then declines steadily with age. This simple pattern conceals many different types of job movement, with a minority of workers changing jobs far more often than others.

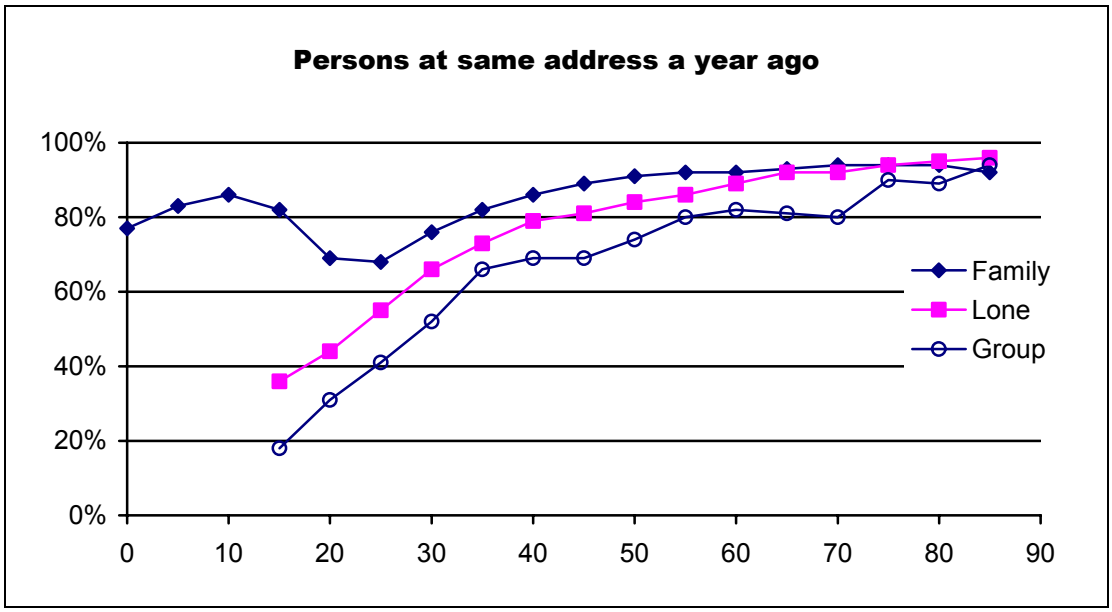
Sources of job mobility

- Changes in industry
- Changes in occupation
- Changes in employer
- Changes in location with same employer
- Movements to and from unemployment
- Children and study
- Immigration and emigration
- Short-term travel
- Young entrants to the workforce.

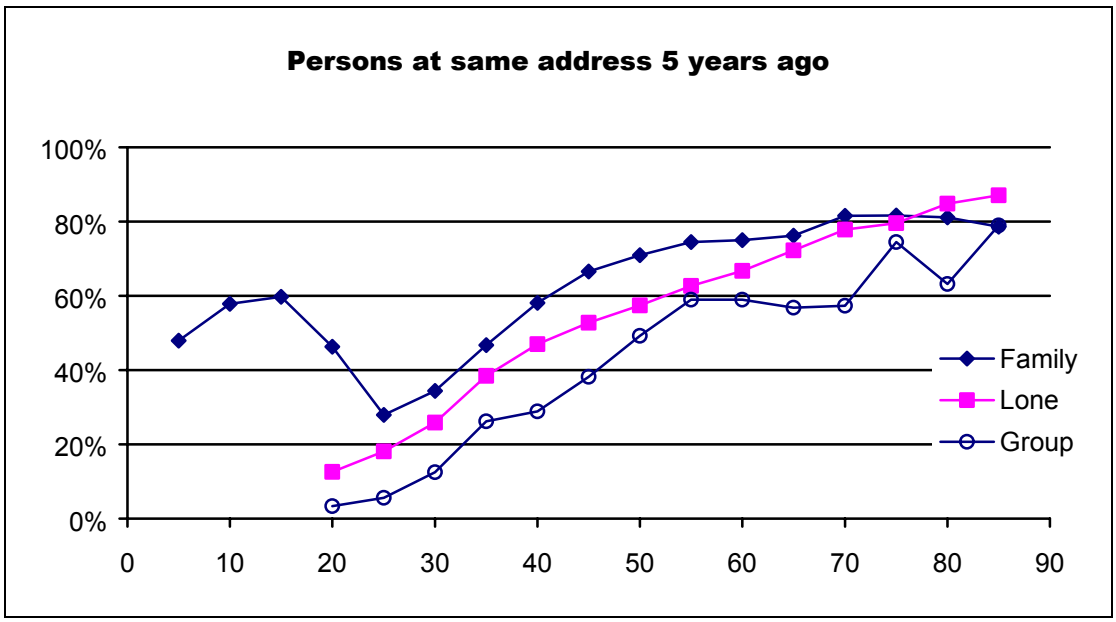
There were 9.061m persons working in February 2002, and of these 1.317m had changed jobs. Of these job changes, 41% involved a change in industry, and 33% a change in occupation. In spite of these many job changes, the average durations of employment of persons in some industries are high:

- 11.9 years in agriculture, forestry and fishing
- 7.8 years in manufacturing
- 5.5 years in retail
- 3.8 years in accommodation, cafes and restaurants
- 9.2 years in education
- 7.5 years in health and community services.

Mobility of persons in different types of household



The above percentages were obtained from the Household Sample File, a 1% sample of the private households from the 2001 census. At younger ages, group households show high mobility, and family households moderate mobility, with lone persons intermediate. At older ages persons in all three types of household show little mobility.



The percentages of persons at the same address 5 years ago are not as low as might be expected from the one-year percentages. Although in part this may be due to persons returning to households, it seems likely that some persons change households frequently, while others are more stable. As with job mobility, there are many sources of household mobility, including dwelling changes, partnership formations and separations, children leaving home and groups forming and breaking up.

Household tenure and stability

Tenure	1-year stability	5-year stability
Owned	91.5%	68.6%
Rented	85.4%	30.1%

Households renting dwellings show less stability than owners (where stability is defined as no change in address for any of the household members). Given that higher proportions of younger persons rent, the differences between owners and renters are not marked.

Common aspects of job and household mobility

- High short-tem mobility
- Low long-term mobility
- Low skills a source of job mobility
- Non-ownership a source of household mobility
- Job and household changes may occur together
- Both job and household mobility require complex models
- Trial and error fitting of models may be needed.

Validation of model assumptions

Checks are needed to ensure that model assumptions yield sensible results. Abello, Kelly & King (2002) suggested several reasons for the substantial differences between projected and actual births, marriages and divorces in the first version of DYNAMOD. The assumptions were based on 1971 to 1986 behaviour, and may not have properly captured subsequent behaviour. Unknown marriage and separation dates were assumed to be recent. Comparing the first few projected years with recent experience can help detect model misspecifications – these have also occurred with my regional population projections. To help validate a model of job transitions, I am currently trying to project Australian workers forward from 1961 to 2001.

Helping model users

As with any consulting task, understanding by the client is vital for acceptance. Formulas and tables may be poorly understood, even by the well educated. Maps may be very helpful in showing projections, and in obtaining vital user input on local conditions. In organizations with decentralised decision-making, much better results are likely if persons at each level understand and accept the process.

It may be feasible to add planning aids to projection outputs. For example, client numbers could be calculated for each designated service centre, or optimal service centre locations automatically chosen.

References

Abello A, Kelly S & King A (2002) "Demographic projections with DYNAMOD-2", National Centre for Social & Economic Modelling, Technical Paper no. 21, May, vi + 50 pages

Australian Bureau of Statistics (2002a) "Marriages and divorces Australia 2001", catalog no 3310.0, Canberra 22/8/02, 122 pages

Australian Bureau of Statistics (2002b) "Labour mobility Australia February 2002", catalog no 6209.0, Canberra 2/10/02, 39 pages

Australian Bureau of Statistics (2003a) "Population projections Australia 2002-2101", catalog no, 3222.0, Canberra 2/9/03, 186 pages

Australian Bureau of Statistics (2003b) "Deaths Australia 2002", ABS catalog no 3302.0, 2/12/03.

Australian Bureau of Statistics (2004) "Charging matrix from automated system for determining charges applicable to ABS customised tables", obtained under FOI 3/3/04, 5 pages

Australian Government Actuary (2003) "Australian Life Tables 2000-02", Canberra, 48 pages + CD

Australian Treasury (2002) "Intergenerational report 2002-03", Canberra, 2002-03 Budget paper No 5, 14 May, vi + 94 pages

Cameron R (2004) "Marketing the key to victory", Financial Review 13/9/04 p10

Department of Finance (2002) "Cost recovery by government agencies", Finance Circular 2002/02, www.finance.gov.au, December, 4 pages

Department of Finance (2002) "Commonwealth cost recovery guidelines for information agencies, December, 36 pages

Dunsford G & Rice M (2004) "Retirement incomes integration - superannuation, social security and taxes", paper presented to the Institute of Actuaries of Australia, February, 87 pages

Orcutt GH (1957) "A new type of socio-economic system", Review of Economics and Statistics, 58, pp 773-797

Orcutt GH, Greenberger M, Korbel J & Rivlin A (1961) "Microanalysis of socioeconomic systems - a simulation study", Harper & Brothers, New York, vxiii + 425

Productivity Commission (2001) "Cost recovery by government agencies", Report no 15, Melbourne, 16/8/01, lii + 217 pages + attachments

Walker a (2003) "Narrower health inequalities in Australia: Impacts simulated using a dynamic microsimulation model", paper prepared for the 7th Nordic Conference on Microsimulation Model, Helsinki, Finland, June 12-13, 36 pages