NATSEM
National Centre for Social and Economic Modelling
• University of Canberra •

APPSIM - Modelling Migration

Sophie Pennec and Marcia Keegan

Working Paper No. 5
September 2007

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, Employment and Workplace Relations, Treasury, and Education, Science and Training.

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

ISSN 1834-7630

ISBN 978-1-74088-274-3

National Centre for Social and Economic Modelling
University of Canberra ACT 2601 Australia
170 Haydon Drive Bruce ACT 2617

Phone + 61 2 6201 2780 Fax + 61 2 6201 2751
Email natsem@natsem.canberra.edu.au
Website www.natsem.canberra.edu.au

Title Modelling Migration in the APPSIM Dynamic Microsimulation Model: a proposal

Author Sophie Pennec and Marcia Keegan

Series Working Paper No. 5

Abstract

This paper is aimed at guiding the development of the migration module of the Australian Population and Policy Simulation Model (APPSIM). It examines the international experience in modelling migration flows and stocks within dynamic microsimulation models and makes recommendations for a suitable modelling approach for APPSIM, given the goals of the project and the data and resources available.

Author notes

Dr Sophie Pennec is a researcher at the Institut National d'études Démographiques in France and is currently a visiting researcher at NATSEM.

Marcia Keegan is a doctoral candidate at the National Centre for Social and Economic Modelling at the University of Canberra.

Acknowledgments

The authors would like to acknowledge comment and input from Ann Harding and Simon Kelly. The authors would also like to gratefully acknowledge the funding and support provided by the Australian Research Council (under grant LP0562493), and by the 13 research partners to the grant: Treasury; Communications, Information Technology and the Arts; Employment and Workplace Relations; Health and Ageing; Education, Science and Training; Finance and Administration; Families, Community Services and Indigenous Affairs; Industry, Tourism and Resources; Immigration and Citizenship; Prime Minister and Cabinet; the Productivity Commission; Centrelink; and the Australian Bureau of Statistics. We would also like to acknowledge our two partner investigators, Professors Jane Falkingham and Maria Evandrou of the University of Southampton.

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

1 Introduction

The Australian Population and Policy Simulation Model (APPSIM) is currently being developed by the National Centre for Social and Economic Modelling (NATSEM) at the University of Canberra, with the assistance of two international partner researchers and 13 industry partners. APPSIM is to be a dynamic microsimulation model (DMSM) of the Australian population, which will project one percent of the Australian population from 2001 to 2050 and beyond. The aim is that APPSIM will be used by the Commonwealth Government and its agencies to assess the future distributional and revenue consequences of changes in tax and outlay programs in an ageing population context.

Rather than simply creating *aggregate* economic and social projections (as non-DMSMs do), APPSIM will project the lives of approximately one percent of the Australian population, including such life events as birth, education, marriage, employment, earnings, taxation, wealth generation, housing, migration, disability, use of health care and aged care services, retirement and death. Because APPSIM will model individuals, it can be used to project the distributional effects of future policy, as well as aggregate effects (NATSEM 2006).

Migration has long been an important component of Australian population growth. Since 1998-99, the number of permanent visas issued has increased from 79 290 to 157 080. The bulk of this increase is from the greater number of skilled migrant visas issued (35 000 in 1998-99 rising to 97 340 in 2005-06 (DIMA 2007)). Emigration has nearly doubled from 1998-99 to 2005-6, from 35 181 to 67 853 persons (DIMA 2006). Slightly more than half of all emigrants in 2005-06 were Australian-born, and a majority of all overseas-born emigrants returned to their country of birth (DIMA 2007).

This paper focuses on modelling migration within APPSIM. It contains a review of international literature on migration modelling in a microsimulation framework, a brief description of Australia's migration system and a proposal for how APPSIM will model migration. In this paper we focus on migration and the characteristics of the migrants. The potentially different behaviour of migrants in the areas of demographics, labour force participation, housing and so forth will be discussed in separate papers on these other modules.

2 International examples of migration modelling

Dynamic microsimulation modelling has become more prevalent throughout the developed world, as increasing computer power makes computation-hungry computer models more feasible and ageing populations create a greater need to address the distributional acts of policy, particularly pension and aged care policies (Harding and Gupta 2007). When embarking upon a large project such as building a new DMSM, it is prudent to review existing models to compare and contrast modelling methods for each model sector.

The international experience shows a number of common methods used to model migration, depending on the data available on the number and characteristics of migrants and the importance of migration to the country's population growth. Some of the most recently constructed DMSMs, such as the SAGE model in the UK, do not model migration at all (Cheesbrough and Scott 2003, p 2). In other cases, countries with limited data on migration (and particularly emigration), such as France and Norway, have used a *net immigration* equation to add immigrants to the population without removing any of the base population from the model. Countries with greater data on the number of migrants, such as Australia, Sweden and Canada, have modelled immigration and emigration separately. The flows of migrants are in most cases an exogenous figure reflecting the migration policy of the country.

In addition, new immigrants must be imbued with characteristics for their incorporation into the model. This has been done by either 'cloning' migrants from random samples of the general population, or from data on the characteristics of immigrants where such data are available.

The following pages review the migration modules of other DMSMs, from Australia and overseas. Table 1 provides a summary of these models.

Table 1: **Modelling migration – international examples**

	Data source	Migration flows	Emigration explanatory variables	Source of immigrant characteristics	Can emigrants re-enter population
DESTINIE (France)	1998 Financial Assets Survey, 1999 census	Net migration only	n/a	Population projections, birth cohort distributions	No
LifePaths (Canada)	1921-1998 censuses, birth and immigration records	Immigration and emigration flows; international and interprovincial movement	Age, sex, province of residence, year, immigrant status, year of immigration	Immigration data	Yes
DYNAMOD (Australia)	1986 census, LSIA, emigration records	Immigration (by visa status) and emigration flows, net long term migration	Age, sex, marital status, Australian or overseas-born	LSIA, cloned from census data	No
MOSART (Norway)	1993 National Insurance Scheme data	Net migration only	n/a	General Norwegian population characteristics, except education	No
Sverige (Sweden)	Immigration data from 1989-90	Immigration and emigration flows	Age, sex, education level, previous immigration, country of origin, labour force participation, income	Immigration pool of 60 000 individuals	No
SESIM (Sweden)	Linda	Immigration and emigration flows	Swedish or foreign born, number of children, adults in household, highest education and oldest age in household, time since immigration	Linda	Yes

Source: See below

2.1 DESTINIE

DESTINIE is a stochastic dynamic microsimulation model built by the French National Institute for Statistics and Economic Studies (INSEE) to analyse the effects of pension reform on retirement patterns and pensions in France, given the heterogeneity of labour force and retirement preferences. It has been used to simulate the impact of the 2003 French Pension Reform on the age of retirement, the replacement rate, the long-term number of retirees and the financial equilibrium of the pension system. Although DESTINIE's primary focus is the simulation of pensions, the model can be used to assess the distributional aspects of other public policies especially if one is interested policy impacts over time.

DESTINIE is based on individual data derived from the 1998 Financial Assets Survey collected by INSEE. The original sample has been reweighted to reflect the demographic structure of the whole French population, as shown in the 1999 French Census. The initial sample of DESTINIE contains about 20 000 households and 50 000 individuals. In order to study intergenerational relationships, individuals from the sample are connected to each other through the artificial imputation of kinship ties. DESTINIE is modularised and operates with demographic, labour market and earnings submodules. The model works on an annual basis and cycles through the submodels in sequence(Afsa and Buffeteau 2007).

The demographic events are simulated in the following order: immigration, death, couple dissolution, couple formation and births (for couples only).

Official population projections of net migration are used to simulate migration. The characteristics of the migrants are also based on the official projections. It is assumed that an equal number of men and women will be migrants. However, the age distribution in the model cannot use the official age distribution of the migrants as there are age groups within which net immigration is negative (in particular for newborns). The changes are made in order to have a profile that is as close as possible to the official one without pockets of negative net immigration, through compensation with the nearest age-groups. The age at end of school attendance of the migrants is estimated using the distribution of their birth cohort (Afsa and Buffeteau 2007).

The family characteristics are age-driven and fairly simple. However, the data available only relate to immigrants and so may not reflect the characteristics of net migrants. Since equal numbers of men and women are assumed to migrate, they are all grouped into pairs (one man and one woman). Each pair created is allocated one of four family situations: two singles without children or ties between them, a couple with one child, couple with two children, or a couple with no children. The woman's age is then estimated according to the family situation, and their partner's age allocated based on the woman's age.

2.2 LifePaths

LifePaths, built by Statistics Canada, is an open, continuous time model designed to simulate life histories of Canadian birth cohorts dating back to 1878. It models births, deaths, international and interprovincial migration, marital history (including common-law unions), education, employment and presence of children at home. It is used to analyse government policies with a longitudinal component and whose nature requires evaluation at the individual or family level (such as post-secondary education costs and public pension sustainability (Gribble, et al. 2003). It can also be used to explore a variety of societal issues of a longitudinal nature, such as intergenerational equity.

LifePaths creates synthetic individuals based on data from birth records, censuses and immigration records. These are aligned with median population projections from 2000-2062 modelled by Statistics Canada. When creating a new individual, sex, place of birth (in a specified province or territory or outside Canada) and date of birth are assigned. Given that the earliest birth occurs in 1872, LifePaths can simulate a complete population aged 0 to 99 from 1971 onwards. The population represented does not include non-permanent residents, a large number of which are foreign students.

There are four types of migration modelled: immigration, emigration, former Canadian residents returning to Canada and interprovincial migration (Statistics Canada 2006).

Immigrants: When a non-native born Canadian is generated by the LifePaths model, he or she is allocated an immigration age and a province of entry. Immigrant characteristics are generated using official population estimates.

Emigration: All individuals in the model who are currently living in Canada face a risk of emigration. Factors affecting emigration risk include age, sex, province of residence, calendar year, immigrant status and the year of immigration (for immigrants).

Returning citizens: Canadian citizens who have emigrated may return to Canada. Factors affecting a return to Canada include age, sex, year, last province of residence, immigration status and year of original immigration (for those born outside Canada). Individuals who were born in Canada or lived in Canada for at least three years before leaving may return to Canada. However, within Lifepaths individuals who immigrate to Canada and leave before staying three years do not return (Statistics Canada 2006).

Interprovincial migration: LifePaths models the movement of families from one province to another. The risk of interprovincial migration is estimated by family types — namely husband-wife couples, common-law couples, lone parents, other single males and other single females — with the effects of age and education on interprovincial migration incorporated into the risk. A base probability of moving between each pair of provinces is used as the benchmark for estimating the risk of interprovincial migration (Statistics Canada 2006).

All forms of migration are based around the family unit. If an individual is selected to migrate, his or her spouse and dependent children will migrate as well.

2.3 DYNAMOD-2

DYNAMOD-2 is a dynamic microsimulation model of the Australian population developed in the 1990s by the National Centre for Social and Economic Modelling (NATSEM). Originally built to assess the impact of income-contingent student loans, it was designed to project characteristics of the population over a period of up to 50 years. The model operates with a one per cent sample of the 1986 census — about 150,000 records — and generates lifetime profiles of people in terms of demographic events (fertility, mortality, couple formation and dissolution, and overseas migration), education participation and attainment, labour force activity and earnings (King, et al. 1999).

In DYNAMOD-2, six distinct overseas migration flows are modelled. One of these is permanent emigration, four are based on the different methods of entry available to permanent immigrants (King, et al. 2002), and the other is based on net long term movement and category jumpers. The reason why users of these entry methods are modelled separately is because of the very different characteristics of the groups of immigrants and the use of eligibility category by governments in setting immigration targets. Permanent arrivals (also known as settler arrivals) are grouped into the following four categories according to visa class: skilled migrants, family migrants, humanitarian migrants and non-program migration (mostly New Zealand citizens). Long term immigrants, emigrants and 'category jumpers' (individuals who change their intended duration of stay while in Australia or overseas), are modelled as a single flow of long term net migration (King, et al. 2002).

The modelling of the immigration flows is a four step process:

1. creating pools of potential immigrants;

2. setting the characteristics of the pools;
3. selecting immigrants from the pools; and
4. introducing the selected cases into the model population (King, et al. 2002).

Establishing the pools of potential immigrants involves creating new records to represent the new arrivals. This can be done by cloning existing cases in the model population or by constructing new cases. A combination of these two approaches is used in DYNAMOD-2. Pools of permanent visa migrants (skilled, family, special purpose and humanitarian), are created with unit record data from the Longitudinal Study of Immigrants to Australia (LSIA) as the main data source. The pool for other permanent immigrants, which largely comprises arrivals from New Zealand, is cloned from appropriate cases over the previous ten years (1976-1986) of census data. In the absence of any reasonable alternative, and on the basis of a comparison of key characteristics, the skilled immigrant pool is used as a proxy for the pool of net long-term migrants. This means that net long term migrants will have a similar distribution of characteristics as skilled migrants. The individual cases in the pools are grouped in families.

The total number, characteristics and category of immigrants are determined for each year of a simulation by the cases selected from each of the five pools. The extent to which the characteristics of the pools differ affects the structure of the immigrant population. The model also provides the opportunity to alter the structure of the populations within each pool. For example, there may be a need to model a changing age structure of family immigrants. This facility is achieved through a combination of reweighting the members of the pools and the specific method of selecting cases as immigrants (King, et al. 2002).

The 'successful' immigrants are then selected from the pools of potential immigrants so that the chance of selection is related to a case's weight, thereby allowing selection to match target numbers and structure. When an individual is chosen as an immigrant, his or her immediate family immigrates as well. The selection of cases from the pools of potential immigrants is undertaken once for each year of the simulation.

Modelling the emigration flows is simpler and involves selecting cases to remove from the model population. Individuals aged 18 years and over are allocated a propensity to emigrate based on age, sex, marital status and country of birth. The Monte Carlo method is used to select the required number and distribution of emigrants. When a case is selected for emigration, other immediate family members are also selected and the unmet target adjusted accordingly. As with immigration, the emigration selection is undertaken on an annual basis, though actual emigration from the model population is spread across the months of a year (King, et al. 2002).

2.4 MOSART

The development of the Norwegian model MOSART started in 1988. The first version comprised demographic events, marriage, education and labour force participation. The second version of the model added public pension benefits and labour market earnings. The current version added household formation, income from non-work sources, savings, wealth and taxation, and upgraded some of the underlying assumptions relating to net immigration, life expectancy and propensity to study (Fredriksen 2003).

The starting population of the MOSART model is a 12% sample of the Norwegian population from National Insurance Scheme data in 1993. This dataset includes information on marriage, birth histories, education, pension status and entitlements, rehabilitation schemes, wealth and household status. The latter includes some information on spouses, parents and children (Fredriksen 2003). Most of the information is annualised data from 1985 onwards, and pension entitlements are represented as annual labour market earnings from 1967 onwards.

MOSART includes some simplifications of migration flows. Rather than model immigration and emigration separately, it only models total net migration — and includes an assumption that total net migration must always be greater than zero (i.e. that the total number of immigrants exceeds the total number of emigrants). In years with net emigration, the negative net immigrants are transferred to the closest year with positive net migration. Each year, an exogenous number (presumed immigration minus emigration) of immigrants are added to the model. If net emigration occurs in any particular age group, a random selection of individuals of that age group are removed from the model (Fredriksen 1998).

Immigrants are classified by age and sex upon arrival, based on 1989 immigrant profiles, and are allocated a similar distribution of characteristics as are found in the general Norwegian population. The only exception is educational attainment, which is set to 'unknown' for new immigrants. (Most individuals who have an 'unknown' education status in education registers are immigrants, and almost all others with unknown education status did not complete nine years of schooling. Thus an education status of 'unknown' is associated with higher risk of disability and lower labour force participation than the average population (Fredriksen 1998).)

2.5 Sverige

The Sverige spatial microsimulation model is the work of the spatial modelling centre in Sweden. It is based on the CORSIM model, with the addition of time geography and agent based modelling.

The Sverige model explicitly models both emigration and immigration. The emigration process is estimated at the household level. The probability a household will emigrate is based on the age and partnership status of the head of the household, with four classifications: 16-21 years, couples where one partner is aged 22-58, singles aged 22-58 and over 59s (Holm, et al. 2002).

The decision to emigrate is influenced by age, sex, education level, previous immigration, duration since immigration if applicable, country of origin, labour force participation and level of income for the household of working age.

The immigrants are picked from an immigration pool of 60 000 individuals (Holm, et al. 2002). The demographic distribution of immigrant household heads is made according to age, sex, marital status and country of origin. A demographic index of the head of the household is determined using age and marital status, and is used to draw the household head from the immigration pool. Stratification according to country is also made. Given the characteristics of the head, a person or a family with all its members will be cloned from this pool. Another transition matrix is used to assign the labour market participation.

The two matrices are determined from the characteristics of the more than 100 000 immigrants who arrived in Sweden in the years 1989 and 1990.

2.6 SESIM

SESIM is a dynamic microsimulation model developed by the Swedish Ministry of Finance in collaboration with Swedish academics. Its development began in 1997, and its first use was an evaluation of the Swedish study allowance system in 1998. Since year 2000 the focus has shifted from education to evaluating the financial sustainability of the new Swedish pension system. SESIM has since evolved into a more general DMSM that can be used for a broad set of analyses.

SESIM's main data source is Linda, a longitudinal dataset drawn from administrative records, representing about 3.5 percent of the Swedish population. A 1999 sample of some 100,000 Swedish citizens is used as a simulation base. Additional observations are sampled from individuals living abroad who have retained Swedish pension rights (Flood, et al. 2005). This is significant because the retention of records for individuals within SESIM who have emigrated allows for emigrants to re-enter the model at a later date.

Emigration: Separate equations are estimated for the foreign-born and the Swedish-born population. Emigration is estimated at household level. The probability that a household will emigrate is based on the age of the oldest person in the household, number of children, number of adults, highest education level in the household and for those who were previously immigrants, the time since immigration is added as a covariate.

Immigration: As with emigration, two different equations are estimated; one for previous emigrants and one for first time immigrants. For the former, the time since emigration is included as a covariate. For the latter, the number is deduced from the official immigration statistics and forecasts. Like emigration, immigration is modelled at the household level. New immigrants are cloned from the overall population and the selection of the cloned household is done such that the immigrated household population achieves a reasonable composition according to the highest age and size of the household (Flood, et al. 2005).

3 Migration in Australia

Arrivals and departures from Australia are treated as short term (less than 12 months), long term (12 months or more) and permanent (no required date for departure/re-entry). Short term arrivals and departures are not considered migrants by the Department of Immigration and Citizenship (DIAC) and ABS (ABS 2006a) and would not be picked up by APPSIM's annualised demographic module, so APPSIM will focus solely on permanent and long term arrivals.

3.1 The immigration programs

Permanent or settler arrivals come under one of five categories: skilled migrants, family reunion migrants, special eligibility (these three together form the Migration Program), the humanitarian program, and non-program migration.

3.1.1 The migration program

The migration program grants skilled migrant visas, family visas and special eligibility visas. The number of visas granted each year is subject to planning figures.

The skilled migration category provides visas to migrants who are sponsored by an employer because of the migrant's specific skills and migrants who can demonstrate sufficient points on the General Skilled Migration test. Within this category, migrants sponsored by State and Territory governments, Australian-based businesses, successful businesspeople wishing to establish themselves in Australia, and applicants with particular distinguished talents, are demand driven and are not subject to capping. If the number of these skilled migrants exceeds planned levels, their number will usually be offset by reductions in unsponsored skilled migrants.

The family category grants visas to spouses, dependent children (including biological and adopted children, and orphaned relatives), fiancés, interdependants (i.e. same sex partners) and other relatives. Visas for spouses and dependent children are uncapped and subject to demand. Visas for fiancés and interdependants may be capped subject to demand for spouse and child visas, and other family visas may be capped subject to demand in other categories. Priority is given to parents who can demonstrate that they can contribute significantly to their costs in Australia, over parents who may need to be supported.

The special eligibility stream grants visas to former Australian residents who have kept ties with Australia, and certain groups of people subject to resolution of status

3.1.2 The humanitarian program

The humanitarian program provides visas for refugees and people who have suffered substantial discrimination in their home country. Its target numbers are set by government, but unfilled places in any one year can be rolled over to the next year and places can be borrowed from subsequent years.

3.1.3 Non-program arrivals

In addition to migrants who arrive on a visa, certain persons are allowed to come to Australia without applying for a visa. These include New Zealand citizens (who make up the bulk of non-visaed arrivals), plus overseas born children of Australian citizens, residents of Cocos (Keeling) Islands and Norfolk Island and persons granted Australian citizenship while living overseas.

3.1.4 Onshore and offshore arrivals

Prospective permanent migrants can apply for permanent residency in Australia in their own countries, or come to Australia on a short or long term visa and apply for permanent residency while in Australia. The former are referred to as offshore visa recipients, the latter are onshore visa recipients. Since the turn of the century, onshore visa grants have increased significantly, with many foreign students who have graduated from Australian universities and long-term skilled visa recipients, applying for permanent residency while physically present in Australia.

3.1.5 Long term arrivals

A further class of migrants of interest to APPSIM are long term arrivals. These include individuals who visit Australia for 12 months or more, and Australian residents who return home after more than 12 months overseas. These individuals are counted as part of Australia's population and are included in estimates of net overseas migration. Long term visas may be granted for many reasons, but the main recipients are students, businesspeople and young people on working holidays (DIMIA 2006).

Certain categories of long term arrivals are of particular interest to APPSIM because of their potential impact on Australia's demographic. For example, '457' visas, which are granted to skilled employees from overseas who have a job offer in Australia, are being used more extensively as Australia suffers from skills shortages. In addition, a long term visa can give migrants a chance to gain training and employment in Australia and allow them to apply for permanent residency. In 2005-06, fourteen percent of long term business visa holders and eleven percent of student visa holders applied for some form of permanent residency (John Ryan, personal communication). Long term arrivals of any visa type may also apply for permanency if they find an Australian partner or have a child while in Australia.

3.2 Patterns of immigration

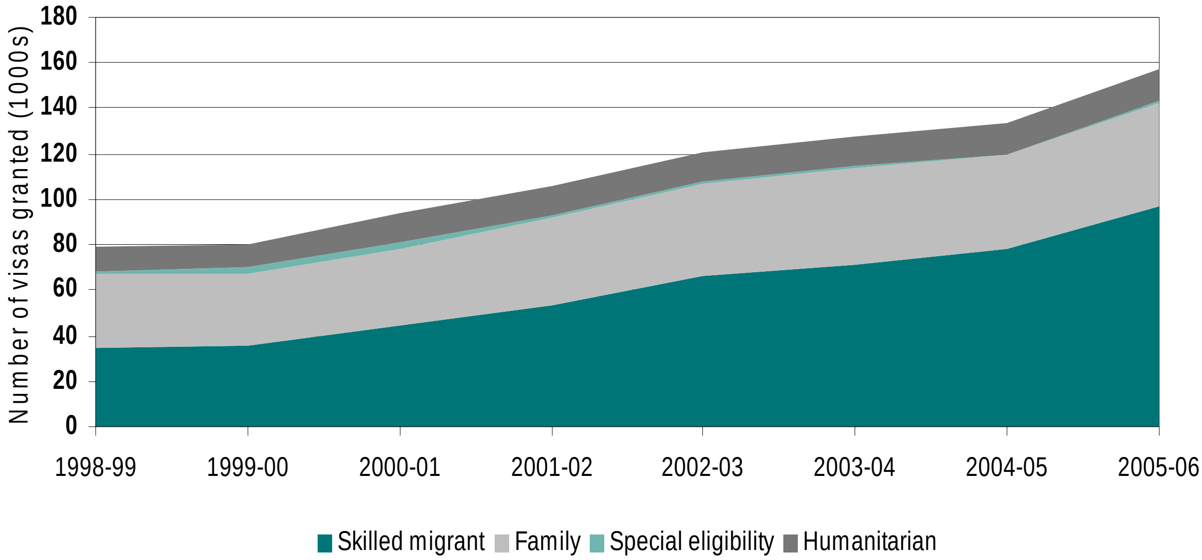
3.2.1 Trends in immigration

In 2005/2006, Australia gained 48 865 skilled migrants, 45 290 family migrants and granted 14 140 humanitarian migrants. Of over a million arrivals from New Zealand, 32 325 of these intended to stay in Australia long term or permanently. There were also 325 821 long term arrivals, of which 32 percent were Australian residents returning home.

Total immigration tends to be positively correlated with GDP — that is, immigration increases in boom times. The most noticeable trend over the last ten years has been the increase in the number of skilled migrant visas issued from 35 000 in 1998-99 in 1995-96 to 97 340 in 2005-06, representing over sixty percent of permanent visas issued (DIMA 2007). Of these permanent arrivals, the largest source country was the United Kingdom (17%) followed by New Zealand (11%.)

Figure 1 shows the makeup of immigrants arriving in Australia over the last eleven years.

Figure 1: **Permanent visas granted 1998-99 to 2005-6**



Data source: Population Flows: Immigration Aspects (DIMA 2007).

A significant change has occurred in the number of people applying for permanent residency whilst present in Australia, as opposed to applying from overseas. Historically, the measure of permanent immigration was permanent settler arrival data; people arriving in Australia with permanent visas, plus New Zealanders intending to stay permanently. Changes in migration policy have made it easer for temporary entrants to apply for permanent residency after arrival in Australia, particularly those on student or business visas. As a result, permanent settler arrivals now account for only 73 percent of all permanent additions to Australia's population (DIMA 2007).

3.2.2 Characteristics of immigrants

Characteristics of immigrants to Australia depend a great deal on the type of visa they are granted. The most significant nation of origin for holders of skilled migrant visas was the UK (26%) followed by India (14%) and China (13%). Skilled migrant visa applicants must meet certain threshold criteria for skills, work experience (except for international students applying to remain in Australia after they finish their degree), age and English language ability, so these migrants tend to be more skilled, more likely to be employed and have better English than other migrants.

The most significant nation/country of origin for holders of family migrant visas was the UK (15%) followed by China (11%) and India (6%). Of family visas granted 80%, or 36 370, were partner visas. Of these, 82% were for spouses, 17% were for prospective marriage partners and the remainder were for same-sex couples. Most partner visas were for people aged 20-39

In 2005-06, just over half of all offshore humanitarian visa grants were to Africans, primarily Sudanese. 17% were granted to Iraqis, and a further 14% were granted to Afghanis.

Skilled visa holders were the most likely to be employed after arriving in Australia. Humanitarian visa holders were the least likely to be employed. However, the longer an immigrant remained in Australia, no matter what their visa class, the more likely they were to be in employment (DIMA 2007).

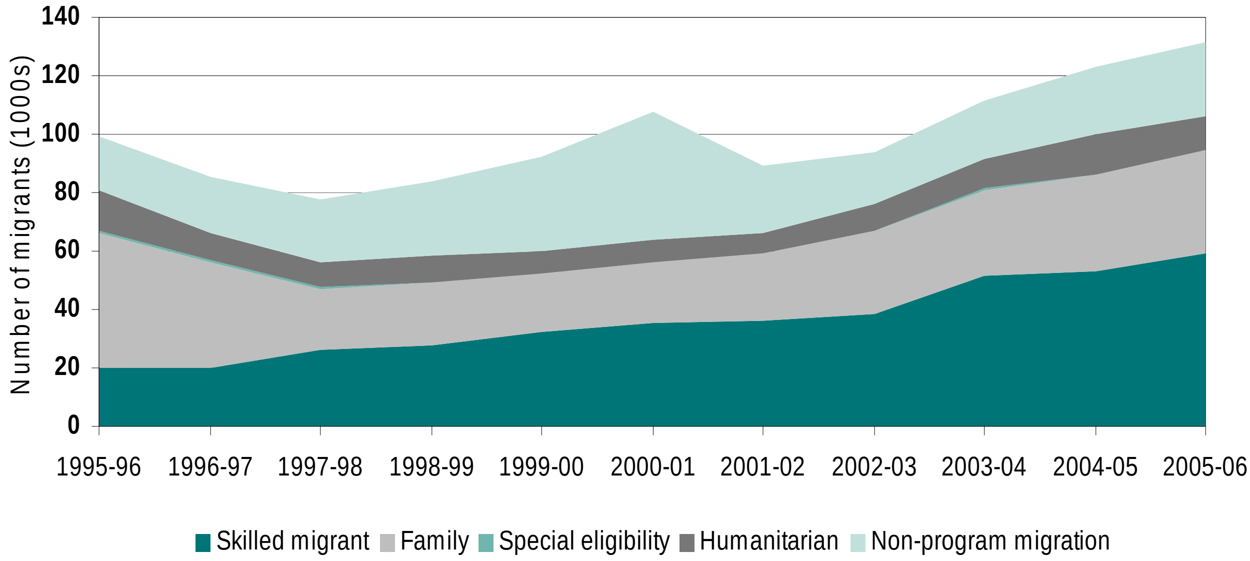
3.3 Patterns of emigration

3.3.1 Trends in emigration

With international travel becoming cheaper and skills becoming more transferable across countries, permanent emigration has nearly doubled over the period from 1998-99 to 2005-06, increasing from 35 181 to 67 853.

Figure 2 shows the increase in the number of Australian-born and overseas-born emigrants over this time period.

Figure 2: **Permanent emigration patterns 1998-99 to 2005-06**



Data source: Settler Arrivals (DIMA 2006)

3.3.2 Characteristics of emigrants

The majority of overseas-born emigrants were departing for their country of birth (DIMIA 2006). Seventy three percent of emigrants classified themselves as skilled workers (i.e. employed in managerial, administrative, professional, associate professional occupations or as tradespeople.) The majority (69.3 percent) of overseas-born emigrants had lived in Australia for at least five years prior to departing.

Australian-born emigrants were most likely to be young families, and overall had a lower age distribution than overseas-born emigrants. This may be because Australian emigrants are more likely to be motivated by economic or employment factors, which primarily motivate a younger demographic; while overseas-born immigrants are more likely to be motivated by factors such as homesickness, widowhood, divorce and desire to retire to one's birth country, which may have a greater impact on older people (DIMIA 2006).

4 Proposed approach to modelling migration within APPSIM

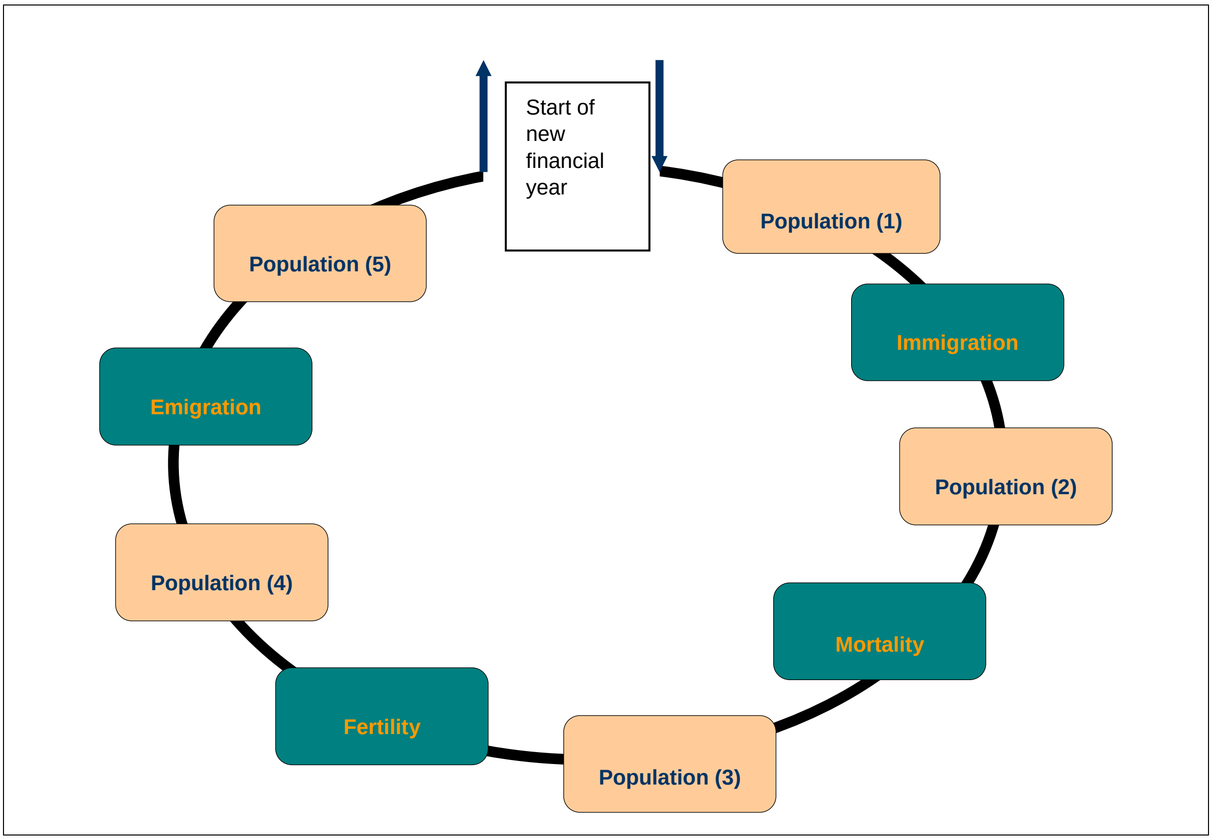
Within the structure of APPSIM, migration is a component of the demographic module. One major difference between migration and the other events simulated in APPSIM is that immigration and emigration are not considered at an individual level but at a family level. So the decision to migrate will be simulated for the whole family (that is the person, his/her spouse and dependent children).

As shown in Figure 3, during a one-year cycle, immigration occurs first, followed by fertility (births), mortality (deaths) and then emigration. This order allows the following to occur:

- Newly-arrived migrants can die, take up a job, marry, buy a house or have children as soon as they arrive;
- A woman can have a baby and then emigrate in the same year (the family-based nature of migration in APPSIM means the woman would take her baby plus any spouse or other children with her); and
- An individual or family may immigrate to Australia in a given year on a permanent visa, but may leave in the same year.

To take into account the fact that immigrants can arrive any time during the year, it will be assumed that they arrive, on average, in the middle of the year. They will then be subject to six months of 'life' in the model, eg six months of labour force transitions, risk of marriage, childbirth etc.

Figure 3: **The events and their order in the demographic module**



4.1 Data sources

The data sources that will be used in the migration module are mainly the 2001 population census, ABS overseas arrivals and departures data, data on visas issued from DIAC and the different Longitudinal Surveys of Immigrants to Australia (LSIA1 to LSIA3).

4.1.1 Census

As well as other demographic and socio-economic information, the population census identifies the country of origin and the date of arrival in Australia for all foreign born Australian residents. It will be the main source of data for the stock of potential emigrants, and it contains some data of the number of Australian citizens born overseas and non-Australian citizens born in Australia.

4.1.2 Longitudinal Survey of Immigrants to Australia

The LSIA survey is a set of longitudinal panel surveys of primary visa applicants and their households conducted by the Department of Immigration and Citizenship

(DIAC). The surveys detail the migrant's origin, background and experiences settling into Australia. The first panel LSIA1, interviewed three times, at different points in time, a sample of immigrants arriving between 1 September 1993 and 31 August 1995. The second panel LSIA2, conducted two waves of surveys of immigrants arriving between 1 September 1999 and 31 August 2000 (DIMIA 2002). The third panel survey, LSIA3, followed people who arrived between December 2004 and March 2005. The first wave was conducted in August 2005 and the second one 12 months later. A third wave is foreseen for 2008. The results of wave 2 have been released in June 2007. (Appendix A).

LSIA3 has the most recent data and it incorporates the effect of recent changes to the skilled migration program. However, LSIA3 does not survey immigrants who arrived under the humanitarian visa program, while LSIA1 and LSIA2 do.

The LSIA will be studied and used to implement the demographic, family and labour force participation characteristics of migrants and their family members.

4.1.3 Arrivals and departures data

When a person enters Australia, they are required to fill out an arrival card which contains identifying information on the individual and how long they plan to stay in Australia. Persons leaving Australia must fill out departure cards indicating how long they intend to remain overseas. These form the basis of Australia's arrivals and departures data. These data provide the flows of migrants and include information on the age, sex, occupation, reason for travel and expected duration of stay, and it is the only source of data on emigrants.

However, these cards only indicate the person's current intentions, which can change over time. An Australian who intends to only spend a month overseas may stay for over a year; likewise a person may arrive in Australia with a long term visa and leave after less than a year. For this reason, data on past years' long term and short term arrivals and departures are updated to take into account these changes in travel plans (ABS 2006a).

4.2 Modelling emigration

Each year, the model will define the number of emigrants required based on overseas departures data, grouped by age, sex, marital status (married or unmarried) and country of birth (Australian or overseas). For each grouping, a propensity to emigrate will be calculated and a random number generated for each individual to determine who will emigrate.

Emigration will be modelled around the family unit; that is, if an individual is chosen to emigrate, his or her spouse and any non-adult children will also emigrate.

Total emigration levels will be set exogenously and alignment with ABS emigration projections/population projections (ABS 2006b) will ensure that stochastic variation does not interfere with these macro emigration targets being reached.

4.2.1 Re-entry of emigrants

Some consideration was given to whether emigrants should be 'put aside' for possible re-entry into Australia/APPSIM. This option would involve taking some or all of the individuals selected to emigrate and modelling their lives apart from other individuals within APPSIM. They would earn money, build wealth, marry, separate, have children, change health status and so forth. Individuals could then be selected from this group for return to Australia. Essentially, simulations would need to be run separately from APPSIM for these expatriates.

The advantage of this option is that allows APPSIM to reflect the fact that emigrants, particularly long term emigrants, do return home. It would allow the migration module in APPSIM to more accurately reflect migration flows.

There are two main difficulties associated with this option. First, there is no data on how emigrants change while overseas. We have no way of knowing how many emigrants come home with a spouse, children, more or less money, a new qualification etc. It could be assumed that an individual's life course follows a similar pattern when they are overseas to what it would have been had they remained in Australia. However, there is no evidence on how reasonable or otherwise this assumption would be. Secondly, a spouse with a full array of characteristics would need to be generated for each emigrant who marries while overseas and returns to Australia, introducing a new level of complexity into the model. If an emigrant has children while overseas, then a lifecourse will need to be generated for the children as well.

It should be noted that of the international DMSMs reviewed in Section 2, only SESIM and Lifepaths modelled emigrants returning home. Lifepaths could do this because it is based on synthetic data. In the Swedish SESIM case, because Sweden's pension system gives Swedish nationals rights to a Swedish pension while overseas, their base data source, Linda, contains some data on the lifecourses of emigrants. Thus the creators of SESIM had access to some data on the life transitions of emigrants. Neither of these situations applies to APPSIM.

It is therefore proposed that permanent emigrants be removed from APPSIM completely. Long-term emigrants will be allowed to return within two years of departure, with no change to their circumstances except for an increase in their age. Although this will not reflect true migration flows in Australia, in which emigrants return after several years, given the lack of data, a more complex model of emigration would not add enough value to APPSIM's projection ability to justify the extra costs involved.

4.3 Modelling immigration

Modelling immigration will be conducted in greater detail than emigration. More data are available and this event is particularly prone to policy monitoring. APPSIM must be able to take into account any policy changes that may occur in the future (for example, changes to immigration targets for certain visa classes).

Several data sources will be used and five categories of immigrants will be assessed separately. As with long term emigration, modelling long term immigration is subject to further discussion with DIAC and ABS.

It is proposed that modelling immigration in APPSIM will involve the following steps:

1) creating pools of potential immigrants/families and their characteristics based on visa status; the characteristics of the individuals of each may be altered during the simulation period if needed to take into account potential changes in the characteristics of migrants in a particular sub-group.

2) choosing the required number and type immigrants from each pool; and

3) incorporating the chosen immigrants into the model population.

4.3.1 Creating pools of potential immigrants

Immigrants are classified according to their visa status: family, skilled migrant or humanitarian visa, non-programme migrants (mostly New Zealand citizens) or long term arrivals. The respondents to LSIA are the immigrant pools for family, skilled worker and humanitarian visas. Pools of permanent visa arrivals are based on LSIA data.

Permanent, non-program migrants (mostly New Zealand citizens) will be 'cloned' from similar populations within the 2001 census (that is, those who have arrived in Australia since 1991).

Long term arrivals include foreigners visiting Australia for more than a year, and Australians who, having previously emigrated, are returning home. They are not covered by the LSIA and it is difficult to identify them from the census data. However, LSIA3 classifies visa holders as onshore or offshore applicants. Many onshore applicants would have originally arrived on long term visas. A pool of long term arrivals will be created, including former emigrants and foreigners. The characteristics of the latter will be based primarily on onshore visa recipients in LSIA3, and if this sample size proves insufficient, on migrants from the family and skilled worker visa categories. Having a separate pool will enable us to alter the characteristics of these migrants if some data sources show that it is necessary.

Long term arrivals who are former Australian residents will have had their records saved, and will be simply added to the population. For simplicity and due to lack of data, it will be assumed that the only change in characteristics a person experiences from their stay overseas is to their age.

LSIA provides weights that reflect the estimated prevalence of the individual or family in the population. This step allows the user to make any necessary adjustments to weights for each potential immigrant. Adjusting these weights allows the module to be used to test the effects of changing the characteristics of the immigrant population (for example, increasing the age structure of immigrants).

4.3.2 Selecting immigrants from each of the pools

This requires not only selecting the target number of immigrants from each group, but also selecting the desired composition of immigrants. The number of immigrants for each year in the model is drawn from either arrivals data by visa class or ABS immigration projections. Selection occurs by random systematic sampling so that the probability of selection is proportional to the weight in the population of potential immigrants.

ABS migration projections do not classify immigrants according to visa class. However, APPSIM requires estimates of future immigration by visa class to impute characteristics to immigrants to be added to the population. A proportion or a number of immigrants according to each visa class/category of migrants will be applied to future immigration flows. This proportion can rely on the distribution observed the previous year or be chosen by the user and his knowledge of the policy regulations. This process will be undertaken in consultation with DIAC.

4.3.3 Incorporating the chosen immigrants into the model population

Immigrants will be selected by family unit. The total number of immigrants will usually slightly exceed the target number because family units continue to be selected until the target number of individuals is reached. However, the overall effect on the model population will be small.

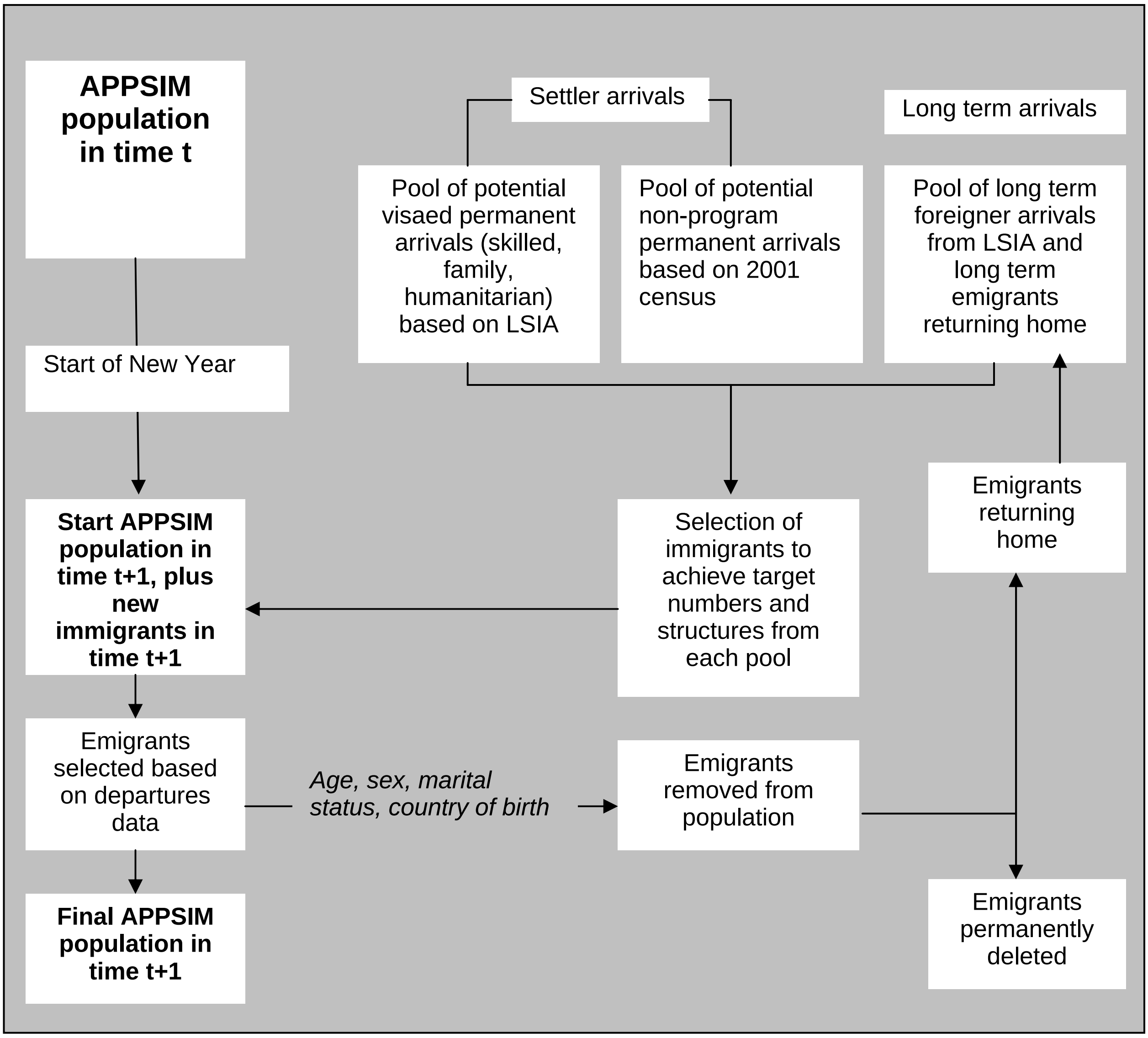
The long term visa arrivals will have a flag placed on their record (excluding Australians returning home). These long term arrivals will be incorporated into the base population along with other immigrants, but they will face a much higher propensity to emigrate in the first few years of their stay in Australia. This will reflect the fact that most long term immigrants will return home a within a few years of arriving in Australia, and the longer they stay, the more likely they are to apply for and receive permanent residence. The propensity of a long term immigrant to leave Australia will be based on visa class, and the proportion of that visa class who apply for permanent residency while in Australia.

4.4 Alignment

A macro model will be created to set annual target immigration levels for each immigrant category. The model will be based on the data used in the ABS population projections (ABS 2006b), specifically, annual total immigration and emigration projections by age and sex. No predictions are available for migration by visa category or permanent/long term migration. It is intended that annual immigration levels can be set by the user if desired.

Figure 4 summarises the proposed method of modelling migration flows in APPSIM.

Figure 4: **Flowchart of proposed migration module for APPSIM**



5 Conclusion

Migration is an important feature of Australian society and it is thus appropriate to model it in considerable detail.

The proposed framework will involve modelling immigration and emigration separately, with five immigration classifications modelled. It will be based on data from the 2001 census, specifically non-visaed immigrants who arrived from 1991-2001, LSIA data (a combination of waves 2 and 3), and arrivals and departure data.

Once developed, the migration equations will be incorporated into the population module.

References

ABS 2006a, *Australian Demographic Statistics*, 3101.0, ABS, Canberra.

ABS 2006b, *Population Projections Australia, 2004 - 210. Reissue*, 3222.0, ABS, Canberra.

Afsa, C. and Buffeteau, S. 2007, DESTINIE ("modèle Démographique Economique et Social de Trajectoires Individuelles Simulées i.e. Demographic, Economic and Social model of Simulated Individual Trajectories "), in *Modelling our future: Population Ageing, Social Security and Taxation*, Harding, A. and Gupta, A. (eds)**.**

Cheesbrough, S. and Scott, A. 2003, '*Simulating Demographic Events in the SAGE Model*', SAGE Technical Note no. 4, London School of Economics, UK.

DIMA, 2006, 'Settler Arrivals 2005-2006', Department of Immigration and Multicultural Affairs, Canberra.

DIMA, 2007, 'Population Flows: Immigration Aspects 2005-06 edition', Department of Immigration and Multicultural Affairs, Canberra.

DIMIA 2002, User Documentation: Longitudinal Survey of Immigrants to Australia, Department of Immigration, Multicultural and Indigenous Affairs, Canberra.

DIMIA, 2006, 'Population Flows: Immigration Aspects 2004-05 edition', Department of Immigration and Multicultural and Indigenous Affairs, Canberra.

Flood, L., Jansson, F., Pettersson, T., Sundberg, O. and Westerberg, A. 2005, *SESIM III - a Swedish dynamic microsimulation model*, Ministry of Finance, Sweden. http://www.sesim.org/

Fredriksen, D. 1998, *Projections of Population, Education, Labour Supply and Public Pension Benefits: Analyses with the Dynamic Microsimulation Model MOSART*, Social and Economic Studies, Statistics Norway, Oslo. http://www.ssb.no/emner/02/03/sos101/sos101.pdf

Fredriksen, D. 2003, The MOSART model - a short technical documentation, *Paper presented at the International Microsimulation Conference on Population, Ageing and Health: Modelling our Future*, Canberra, 7-12 December.

Gribble, S., Hicks, C. and Rowe, G. 2003, The Lifepaths Microsimulation Model, *Paper presented at the International Microsimulation Conference on Population, Ageing and Health: Modelling our Future*, Canberra, 7-12 December.

Harding, A. and Gupta, A. 2007, Introduction and Overview, in *Modelling Our Future: Population Ageing, Social Security and Taxation*, Harding, A. and Gupta, A. (eds), International Symposia in Economic Theory and Econometrics, North Holland, Amsterdam**.**

Holm, E., Holme, K., Makila, K., Mattson-kaupi, M. and Mortvik, G. 2002, *The Sverige microsimulation model - content, validation and example applications*. Series, Kulturgeografiska Institutionen/SMC, University of Umea.

King, A., Baekgaard, H. and Robinson, M. 1999, '*DYNAMOD-2: An Overview*', Technical Paper No. 19, National Centre for Social and Economic Modelling, University of Canberra, Canberra.

King, A., Walker, A. and Baekgaard, H. 2002, '*Modelling Overseas Migration in DYNAMOD-2*', Technical paper no. 22, National Centre for Social and Economic Modelling, University of Canberra, Canberra.

NATSEM 2006, The APPSIM Project, National Centre for Social and Economic Modelling, University of Canberra, Accessed 15 February 2007

Statistics Canada 2006, The Lifepaths Microsimulation Model: An Overview.
http://www.statcan.ca/english/spsd/LifePaths.htm,

A Appendix

Table 2: **LSIA2 sample by age**

	15–24 years	*25–34 years*	*35–44 years*	*45–54 years*	*55–64 years*	*65 years and over*
Preferential family	3544	8608	2363	1181	506	844
Concessional family	132	1584	1155	330	99	
Business/ENS	19	208	757	833	76	
Independent	399	5101	2312	159		
Humanitarian	145	702	895	460	145	73
Total	4220	16231	7466	2922	649	974

Source: User documentation: Longitudinal Survey of Immigrants to Australia (DIMIA 2002)

Table 3: **LSIA2 sample by sex**

	Males	*Females*
Preferential family	6407	10453
Concessional family	1945	1351
Business/ENS	1577	300
Independent	5500	2471
Humanitarian	1471	940
Total	16856	15559

Source: User documentation: Longitudinal Survey of Immigrants to Australia (DIMIA 2002)