

Paper delivered at the 10th Biennial Conference of the
Australian Population Association
POPULATION AND GLOBALISATION:
AUSTRALIA IN THE 21ST CENTURY
Melbourne 28th November to 1st December 2000
Melbourne Australia

Medicare, a source for migration statistics

Dr Badr Hanna
Australian Bureau of Statistics

Abstract

This paper investigates the possibility of using the Medicare enrolment file to derive internal migration for Statistical Local Areas (SLAs) in NSW. It discusses the data problems, the assumptions adopted and the methods used for estimating internal migration. With annual estimates of internal migration, it would be possible to use the component method to estimate the annual resident population of SLAs.

Two methods were used to estimate internal migration. The first was derived from all Medicare cards which were adjusted using 1996 Population Census data as a benchmark. As some addresses may have not been updated for several years prior to the reference year a second method was used. The later was based on the cards that have been updated and replaced in the previous 12 months. The migrants' estimates were then adjusted to remove the effect due to the lack of information on the date of migration.

The study also examined the possible use of the migration estimates for SLA estimates of the resident population. Two types of experimental population estimates were calculated using the component method, the first incorporated migration estimates based on all Medicare cards and the second incorporated migration estimates based on the replaced cards only. Both sets of experimental SLA population estimates were then compared with the ABS Estimated Resident Population .

The study shows the potential use of Medicare data as the best available source for internal migration statistics. The methods used provide consistent estimates of internal migration at SLA level, whether the migration estimates are derived from all Medicare cards and adjusted based on the Population Census data or based solely on the replaced cards. The analysis of population estimates shows that the component resident population estimates based on migration derived from the replaced cards were closer to the ABS Estimated Resident Population than those estimates based on migration derived from all Medicare cards.

Objectives

1. The resident populations for Statistical Local Areas (SLAs) for Australia are estimated by the Australian Bureau of Statistics (ABS) annually based largely on regression models that incorporate a number of administrative variables as symptomatic indicators of population change. The discrepancy between these estimates and the Population Census has been more likely to be proportionally greater for smaller populated SLAs than those with larger populations. In these smaller SLAs, internal migration is usually the main cause of population growth or decline and hence the discrepancy. Yet, it is the most difficult component to measure.

2. The objective of this paper is to investigate the possibility of using the Medicare enrolment file to derive internal migration statistics at the SLA level. It discusses the data problems, the assumptions adopted and the methods used for estimating internal migration. With annual estimates of internal migration, it would be possible to use the component cohort method which incorporates internal migration along with the registered births, deaths and overseas migration to estimate the resident population of SLAs. At the very least such estimates serve as a useful basis for validating the regression outputs.

Experience of other countries

3. In many countries, administrative data have been used to estimate internal migration. In the USA, taxation data are used to estimate migration for persons aged under 65 years, while estimates of child migration are prepared from school enrolments data. Social security data are used to provide sample estimates of migration for the elderly. In England and Wales, the main data source used to estimate internal migration is patient transfers details from the National Health Service Central Register in conjunction with electoral registration. Statistics Canada uses personal income tax data to produce finalised estimates of inter-provincial migration. Since the 1970s, Statistics New Zealand have used a variety of symptomatic indicators to prepare annual estimates of net internal migration at Territorial Local Authority level. The symptomatic indicator data are converted to net internal migration estimates by informal and subjective estimation practices.

4. In Australia, the ABS has in the past used administrative data sources such as electoral roll registrations and family allowance payments. Since 1996, estimation of post-censal interstate migration has been based on Medicare data as it is considered to be the best administrative data source available with regard to population coverage and timeliness.

Definition of spatial population movements

5. In theory, spatial population movements are defined on a space-time continuum framework. In practice, the definition of population movements is determined by the way they are measured. In Australia, the measurement of migration derived from Population Census data is restricted to two questions. Respondents are asked to report their usual address one year and five years prior to Census night. The measurement of migration derived from Medicare enrolment data is restricted by the way the data are collected and recorded to meet the administrative objectives. A comparison between the migration levels derived from the two sources is very difficult unless the time and space on which migration is defined, can be made comparable.

Sources of migration statistics

Census data

6. The Census of Population and Housing provides an almost universal coverage of migration of Australia's population. It defines precisely the time and space dimensions of population movement. The time is defined as 12 months (or 5 years) prior to the Census date. The space is defined by the geographical boundaries of SLAs, States and Territories and the national boundary of Australia. The Census defines usual address as the address at which the person has lived or intends to live for a total of 6 months or more. The concept of "usual residence" poses some measurement problems. In identifying the place of usual residence, the *de jure* approach leads to an under recording of population mobility as it excludes temporary movements. With the *de facto* approach, where actual residence is recorded, migration is over-estimated, since no distinction can be made between short term movements such as boarders and more permanent migration. Another problem relates to the accuracy of responses provided on previous place of residence. In the 1996 Census there were 184,611 records in NSW with undefined usual residence, representing 2.9% of the total population in NSW.

7. There are several stages of adjustment for Census data in order to get more accurate estimates of *de jure* population. Firstly, Census counts of usual residents are adjusted for under counting using a Census Post Enumeration Survey. Secondly, estimates of the number of Australian residents temporarily overseas are added to the appropriate SLAs. Thirdly, as Census night in recent years is in early August, a further adjustment is made to produce the estimates as at the 30th of June.

8. Another limitation is that migration data derived from the Census are based only on experience during the period preceding the Census. This experience is assumed constant throughout the intercensal years. Whilst at the aggregate level of States, population migration may not change substantially over time, at the smaller levels of SLAs greater variation may be experienced.

9. The Census measures migration by comparing place of usual residence at two points in time. It records the number of migrants who have experienced at least one move during the time interval, but it does not record the number of moves. People who move short distances are more likely to make a number of moves than those who move long distances. The 1996 Census revealed that only 1.2% of the NSW population was living in another State in 1995 while 6.5% of the population had migrated among different SLAs of NSW.

Medicare data

10. Medicare as a source of migration statistics has important advantages with regard to data content, timeliness, historical availability and electronic access. The eligible Medicare population includes all Australian residents. It also includes some non-Australian residents who have not been granted a permanent visa such as New Zealanders, but have the intention to stay. However, in reality the Medicare population excludes some Australian residents who are eligible to use Medicare but unlikely to do so because they have access to alternative health services. These people include defence force personnel, prisoners, some Aboriginal and Torres Strait Islanders and persons eligible for Department of Veteran's Affairs' Health Services. Nevertheless, the differences between ABS population estimates and Medicare enrolled members for the years 1996 to 1998 were less than 3% for NSW and tended to decrease in the more recent years.

11. In the early 1990s the Health Insurance Commission (HIC) introduced the Card Replacement Program to update their enrolment file. Since September 1999 the HIC has begun introducing other changes to improve the quality of the enrolment file. Key changes include linking the enrolment file to Australia's Registrars through the Fact of Death file, with the result that records

of deceased persons will be closed off on the Medicare file and linking to the Department of Immigration and Multicultural Affairs to check the number of permanent arrivals and departures.

12. Currently, the Medicare file does not include any reference to the time of the move. While a change in postcode implies that a move took place, and the direction of move is defined by the postcodes recorded in two consecutive years, there is no information as to whether this move took place within the reference period, or at any date prior to the reference period.

13. Postcodes in the Medicare file may refer to a business location rather than a residential location, as currently postal addresses are recorded in the file. Business post office boxes are more frequent in business areas and it is difficult to distinguish whether the postcodes relate to commercial or residential properties. But, the proportion of commercial postcodes is very small. Currently Medicare is in the process of rebuilding the enrolment file to provide both residential and postal addresses and possibly previous addresses. To convert postcodes into SLA codes, the ABS *Postal Area to Statistical Area Concordance* was used. In some cases, postcode boundaries are ill defined and do not match Census Collection Districts. In these cases the quality of the conversion for these areas could be affected.

Intercensal population estimates

14. Currently, Estimated Resident Population (ERP) for SLAs during the intercensal years is calculated based on regression models. These models establish the relationship between changes in the proportion of an SLA population to the state population and changes in the proportion of an SLA's indicators to the state indicators. The symptomatic indicators used vary between States and Territories. With the knowledge of the growth in the indicators during the past year, population growth can be estimated.

Thus

$$\text{Pr } \text{SLA}_i{}_{t-1,t} = a + b \text{D}\text{SLA}_i{}_{t-1,t} + e \text{SLA}_i$$

where

$\text{Pr } \text{SLA}_i{}_{t-1,t}$ is the growth in the proportion of an SLA population to the state population between year t-1 and year t,

$\text{D } \text{SLA}_i{}_{t-1,t}$ is the growth in the proportion of an SLA's indicators to the state indicators between year t-1 and year t. In NSW, information on dwelling approvals and Medicare enrolments are used as symptomatic indicators. The regression coefficients 'a' and 'b' are estimated from data available for all SLAs and 'e SLA_i' is an error term.

15. All SLA population estimates that are produced from these models are scrutinised individually by ABS officers. Some estimates may need to be derived without the assistance of mathematical models. This approach requires extensive demographic and non-demographic knowledge, particularly of internal migration in the area under investigation.

16. The accuracy of these intercensal estimates can only be gauged when the results of the next Census are available. In 1996 almost 73 per cent of Australia's 1,330 SLAs were estimated to within five per cent of their 'true' value. Of the 27 per cent of SLAs which varied by more than 5 percent, the majority were for SLAs with small populations.

17. If migration at SLA level were available, this information, along with registered births, deaths and overseas migration could be used directly to estimate the resident population using the component method. The method takes the most recent Census population and adds births,

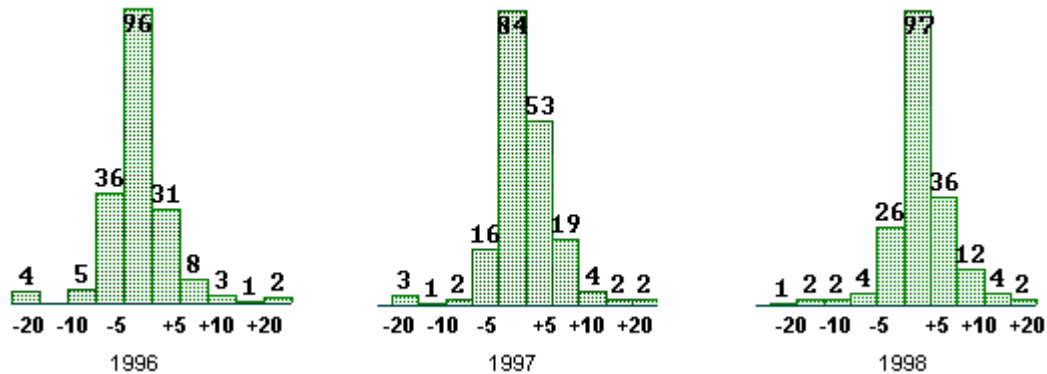
subtracts deaths and adds or subtracts net overseas migration and net internal migration that has occurred during the period.

Medicare population

18. The population for NSW Statistical Divisions (SD) as recorded in the Medicare enrolment file for 1996-98 is presented in Appendix A, The table also provides the 1996 ERP based on the Population Census and the 1997-99 ERP based on the regression model. The percentage differences between the ERP and the Medicare population for 1996-98 are calculated in the last three columns. It shows that at the SD level the differences between the two sources of data were less than 6%. At the State level the absolute differences were less than 3% . In recently obtained information for 1999, the NSW Medicare population was 6,389,000 compared with an ERP of 6,411,000, a difference of 0.35%. The table shows that estimates derived from Medicare data are lower than the ABS ERP in most cases, but the differences gradually declined over time. This could be a result of the HIC's program to improve the coverage of their file.

19. Figure 1 shows the distribution of the percentage differences between the ERP and the Medicare estimates for the 188 SLAs in NSW. It can be seen that the majority of the differences are clustered within the range of plus or minus 5%.

Figure 1 - Percentage differences between ERP and Medicare population estimates, NSW, 1996-98.



(The numbers above the columns indicate the numbers of SLAs)

Types of migration

20. The status of mobility in the Medicare population was defined by comparing the SLAs (converted from postcodes) of a record in two consecutive years. Between June 1998 and June 1999, migrants were defined in five different types as follows.

- a. Did not move
SLA99 = SLA98
- b. In-migrants from another SLA within NSW
SLA99 \neq SLA98, classified according to SLA99
- c. Out migrants to another SLA within NSW
SLA99 \neq SLA98, classified according to SLA98
- d. In migrants from another state or from overseas
SLA99 exists and SLA98 is missing, classified according to SLA99
- e. Out migrants to other state or to overseas
SLA99 is missing and SLA98 exists, classified according to SLA98

21. As it cannot be distinguished from the Medicare file whether a person moved over the NSW border or moved over the Australian border, migrants to/from other States and Territories and migrants to/from overseas are combined.

The main disadvantage of using Medicare data to estimate migration is that the records do not define the time of the move. For example, if the postcode at June 1996 is different from the postcode at June 1995, it implies that the persons has moved. However, it is not known whether the move took place in 1995-96 or at a date prior to June 1995. In other words, the number of migrants captured in one year's data might not represent the number of migrants who moved in the 12 month period, but also those who moved over a number of years prior to the reference date.

Estimates of migration

a) Estimates based on all Medicare cards

22. It is believed that the migration numbers derived from the Population Census are reliable and can be used as a benchmark for other migration estimates. Based on this assumption, the ratios of migrants estimated from the 1996 Medicare data to those estimated from the 1996 Census were used to adjust the preliminary Medicare migration estimates for the post Census years. These

adjustment factors were calculated separately for each SLA and for each of the five move types as well as the not-moved population.

Thus, for 1996 with respect to SLA i ‘in migrants’ who moved from SLAs, the adjustment factor is

$${}^{96}_{SLA\ i\ Cen\ In} / {}^{96}_{SLA\ i\ Med\ In}$$

Then the 1998 adjusted estimates for this category of migrants with respect to SLA i is as follows:

$$Adj\ {}^{98}_{SLA\ i\ Med\ In} = {}^{98}_{SLA\ i\ Med\ In} * ({}^{96}_{SLA\ i\ Cen\ In} / {}^{96}_{SLA\ i\ Med\ In})$$

Where

$Adj\ {}^{98}_{SLA\ i\ Med\ In}$ is the 98 adjusted Medicare records of ‘in-migrants’ from other SLAs

${}^{96}_{SLA\ i\ Cen\ In}$ is the 96 Census of SLA i in-migrants from other SLAs, and

${}^{96}_{SLA\ i\ Med\ In}$ is the 96 Medicare of SLA i in-migrants from other SLAs.

Similarly, the ‘out migrants’ (out) and ‘never moved’ (Nm) population can be adjusted as follows:

$$Adj\ {}^{98}_{SLA\ i\ Med\ Out} = {}^{98}_{SLA\ i\ Med\ Out} * ({}^{96}_{SLA\ i\ Cen\ Out} / {}^{96}_{SLA\ i\ Med\ Out})$$

$$Adj\ {}^{98}_{SLA\ i\ Med\ Nm} = {}^{98}_{SLA\ i\ Med\ Nm} * ({}^{96}_{SLA\ i\ Cen\ Nm} / {}^{96}_{SLA\ i\ Med\ Nm})$$

23. These factors adjust the preliminary Medicare migration estimates for the possible under estimation or over estimation of Medicare’s population due to those who have died, those who have left the country permanently without notifying Medicare or those who are not covered under Medicare. The application of these factors to adjust the data for post Census years was based on the assumption that the types, level and direction of errors remain similar across the years. This assumption was reasonable for two reasons. Firstly, the level of overseas migrants to/from Australia did not vary substantially during the period under study. Secondly, the Health Insurance Commission did not initiate any major changes to their file during this period.

Disadvantages of the Census based adjustment

24. In the above method the 1996 Census was used as a benchmark to adjust the Medicare migration estimates in post Census years. The adjustment factors brought the pattern and level of migration in line with the Census. The main disadvantage of this approach is that the migration estimates are based on the Medicare enrolment file, part of which has been edited every year under the card replacement program (see paragraphs 27-28). Thus, the denominator of the adjustment factor consists of a mixed bag of records, some of which were updated in the previous year and some of which were updated two, three or four years prior to the reference year. Furthermore, the proportion of records edited in 1996 was different from the proportion of records edited in 1997 or 1998.

25. Therefore, an alternative methodology was developed in which migration estimates are based on the cards which were updated in the previous year as part of the card replacement program. This methodology does not rely on benchmarking the Medicare data to Census data, and will therefore avoid the problems related to Census data being outdated.

b) Estimates based on the Medicare replaced cards

26. With an increased reliance on direct billing by doctors, the proportion of patients who actually process their Medicare claims, and therefore have an opportunity to report their change in address, is falling. These out of date addresses will then be translated into an incorrect migration pattern. In 1994, the Health Insurance Commission began a cycle of replacing member's cards in an effort to maintain the quality of the data in their file. All cards would be replaced in a five to seven year cycle, meaning that approximately 14-20% of records would be updated each year. The program does not pick up cards that are issued to visitors because such cards have an end date not an expiry date. Those cards are usually issued for quite short periods of time.

27. There are four stages to the card replacement program:

Medicare cards that are going to expire at the end of each month are detected and checked against the Patient Claims. When a patient claims a benefit at a Medicare office or sends a claim in by post, they have to declare their current address. For members who have completed a claim in this manner in the last 9 months, a replacement card is automatically sent to the most recent address. The remaining cards are then run past the enrolment file. Where a change of address within the last 18 months is detected, this address is considered reasonably accurate and a replacement card is automatically sent to that address.

The remaining cards are then run past Telstra's electronic White Pages. If a match of surname and address is found, a replacement card is automatically sent to that address.

It is likely that the remaining cards belong to people who use direct billing, do not make frequent claims or do not have a phone account in their name. For these cards, the HIC sends a letter to the last known address inviting the cardholder to apply for a replacement card. If no reply is received, the card record is closed off.

28. Migration was estimated based upon a subset of records whose cards were replaced in the previous year. Due to the process of the card replacement program, as described above, it is believed that the addresses in these records are more accurate. The estimation of the migratory population based on the card replacement program was processed in five steps.

Step 1- Calculating replacement ratios

29. The replacement ratio was calculated for each SLA by dividing the replaced cards by all cards in that SLA.

Therefore, the replacement factor for SLA i in 1996 is:

$${}^{96}_{SLA\ i\ Med\ R} / {}^{96}_{SLA\ i\ Med\ All}$$

Where;

${}^{96}_{SLA\ i\ Med\ R}$ is the number of replaced cards in SLA i and

${}^{96}_{SLA\ i\ Med\ All}$ is the number of all cards in SLA i .

Step 2 - First estimates of the number of migrants

30. For SLA i an estimate of the population in each type of move was calculated by dividing the number of cards replaced in each move category by the SLA specific replacement factor.

Thus, the 1996 estimate of in-migrants from other SLAs is equal to:

$${}^{96}_{SLA\ i\ Med\ R}^{In} / \text{the SLA } i \text{ replacement factor}$$

31. Having completed the first two steps the number of not-moved population was underestimated, and represents only 71% of the not-moved population derived from all Medicare cards. But, the number of migrants estimated from the replaced cards was over-estimated for every type of migration. The estimates were double that of in-migrants and out-migrants within SLAs, 77% higher than out-migrants to other States or overseas and five times higher than in-migrants from other States or overseas (see table below). This magnitude of over estimation was expected as the card replacement program while identifying migrants, does not determine the time of move, and therefore may identify not only migrants who might have moved during the previous year, but those who moved at anytime since their card was last replaced five years ago.

1996	Not Moved	In migrants from SLAs	Out migrants to SLAs	In migrants from States and overseas	Out migrants to States and overseas	Total cards
Estimated	3,780,177	904,403	917,266	1,202,887	256,035	6,149,934
All records	5,309,401	458,370	458,370	237,781	144,381	6,149,933
Ratios of the estimated figures	0.712	1.973	2.001	5.059	1.773	1.00

Step 3 - Excluding ex-migrants who moved prior to the reference year

32. Unlike the migrant's status, the 'not-moved' status is not subject to space and time dimensions. The not-moved are a stable population whose status remains constant, independent of the time when their not-moved status was initiated and irrespective of their previous historical migration. The 1996 Census recorded 5,170,310 persons as not-moved during the year prior to the Census. During the same period, the Medicare all cards data set recorded 5,309,401 persons as not-moved. The difference between these two estimates was small (2.6%) an indication that the not-moved population in the Medicare all cards data set was nearly accurate. Based upon this assumption, the difference between 5,309,401 (the not-moved population from the all cards data set) and 3,780,177 (the not-moved population estimated from the card replacement data set) most likely represents migrants who moved prior to the reference year. Therefore, the difference of 1,529,244 persons (5,309,401 - 3,780,177) was removed from the migratory population and was added to the not-moved population proportionally by SLA. The assumption being that all migration categories contributed proportionally to the over-estimation of migrants. The table below shows the result of this adjustment.

1996	Not Moved	In migrants from SLAs	Out migrants to SLAs	In migrants from States & overseas	Out migrants to States & overseas	Net migration	Total Cards
Excluding the over estimated migrants	5,309,401	488,650	492,497	631,075	139,145	488,083	6,570,195

Step 4 - Excluding extra cards from "in-migrants from other States or overseas"

33. Having established the not-moved population from the replaced card data set at the same level as the not-moved population from the all cards data set and adjusted the migratory population to exclude migrants who may have moved prior to the reference year, the total number of

records increased from 6,149,933 to 6,570,195. This excess in the total NSW persons must have been as a result of the in-migrants from other States and overseas, which was five times the size of the category in the all cards file. This could be a bias introduced by the new cards issued for newly arrived persons from overseas or new cards issued to replace lost or damaged cards or even additional cards issued to members of the same family for convenience. The difference of 420,262 (6,149,933 - 6,570,195) was removed from the in-migrants from other States and overseas, reducing this category to 210,813 migrants. The table below shows the result of this adjustment.

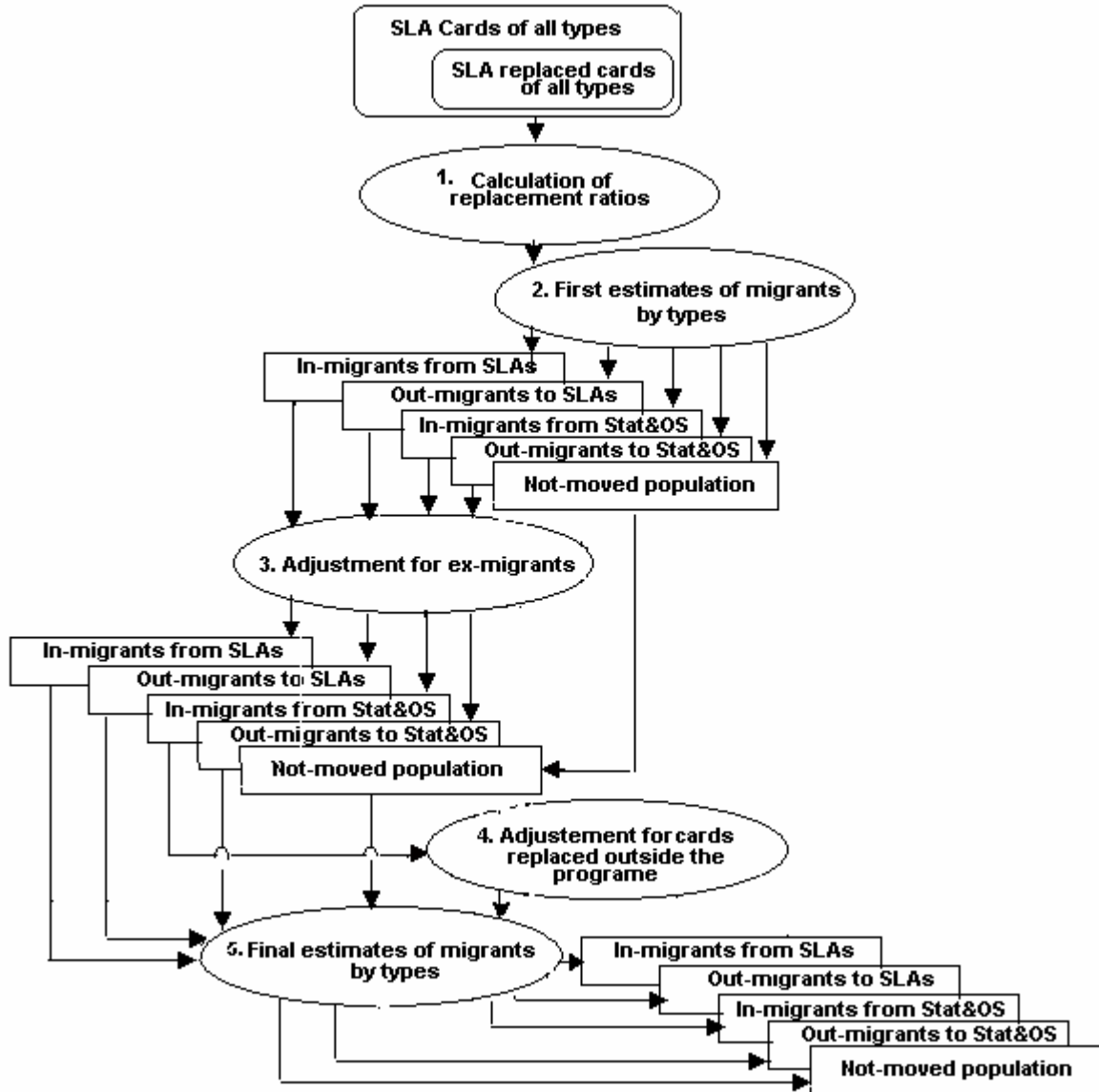
1996	Not Moved	In migrant s from SLAs	Out migrant s to SLAs	In migrants from States & overseas	Out migrants to States & overseas	Net migrati on	Total Cards
Excluding possible new cards	5,309,401	488,650	492,497	210,813	139,145	67,822	6,149,933

Step 5 - Imputing for missing values or outliers

34. In a few SLAs, the number of migrants was either very small, zero, or outliers after the previous adjustments were made. The number of migrants in these SLAs, was then imputed on the basis of its weight in the statistical division where it is located in the following years. The estimates were then pro-rated to the NSW totals.

35. The following chart summarises the flow of these five steps

Figure 2 - Flow chart of the procedure of migration estimates based on the replaced Medicare cards



Further development of the Medicare Card Replacement Method

36. There are two main issues to be further developed for the card replacement method. The first is that the replaced cards might include cards which have been newly issued for various reasons. These cards having been flagged as cards in the replacement program, to ensure they are not selected for replacement in the following year. It is important to filter these cases from the replaced

cards so that the card replacement program can reasonably serve as a statistical sample. Fortunately, it is likely to be possible to identify these cases with the use of a particular source office code which is part of the card replacement program. The second is that the absence of information on the time of the move poses another problem since the captured movers are not necessarily the movers during the reference year. However, currently the HIC is in the process of redeveloping the operation system to allow the recording of previous addresses and if changes of address are combined with the date of migration there will be complete records of migration history .

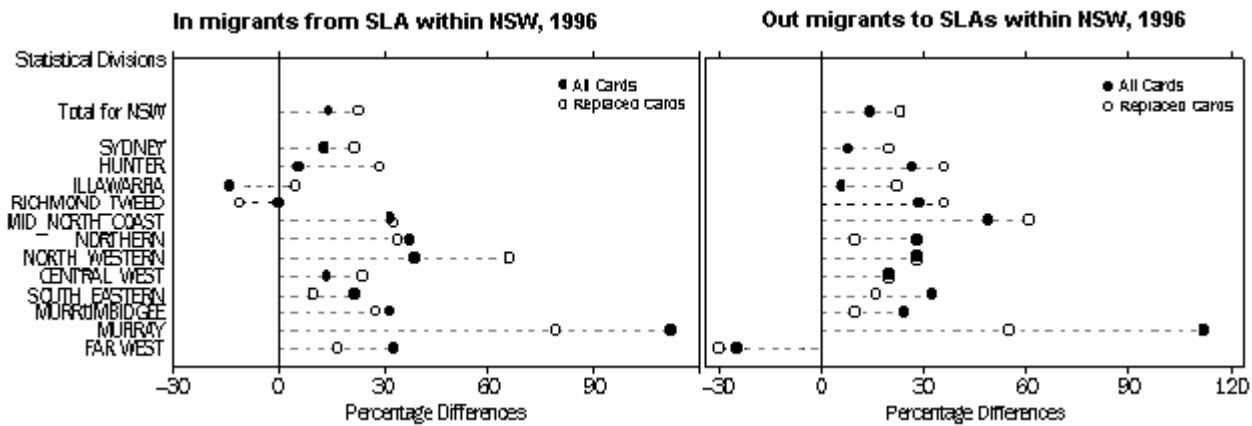
37. The Medicare enrolment file is being linked with the Registrar Generals Fact of Death file and the Department of Immigration and Multicultural Affairs records. With the increasing inclusion of Aboriginal communities in the Medicare enrolment file there will be comprehensive coverage of population that should match closely the ABS Estimated Resident Population for SLAs.

Analysis of results

38. The migration estimates based on all cards adjusted using the 1996 Census are given in Appendix B. The migration estimates are provided for different SDs by the five types of migration. Similarly, the migration estimates based on the replaced cards are given in Appendix C. Graphs 3-5 present the results of the migration estimates using the two methods. The Census adjusted estimates are referred to as 'all cards' and the card replacement program estimates are referred to as 'replaced cards'. The differences between each of these two sets of estimates and the 1996 Census figures were calculated for the five types of migration, in-migrants from SLAs, out-migrants to SLAs, in-migrants from other states and overseas, out-migrants to other states and overseas, and not-moved population.

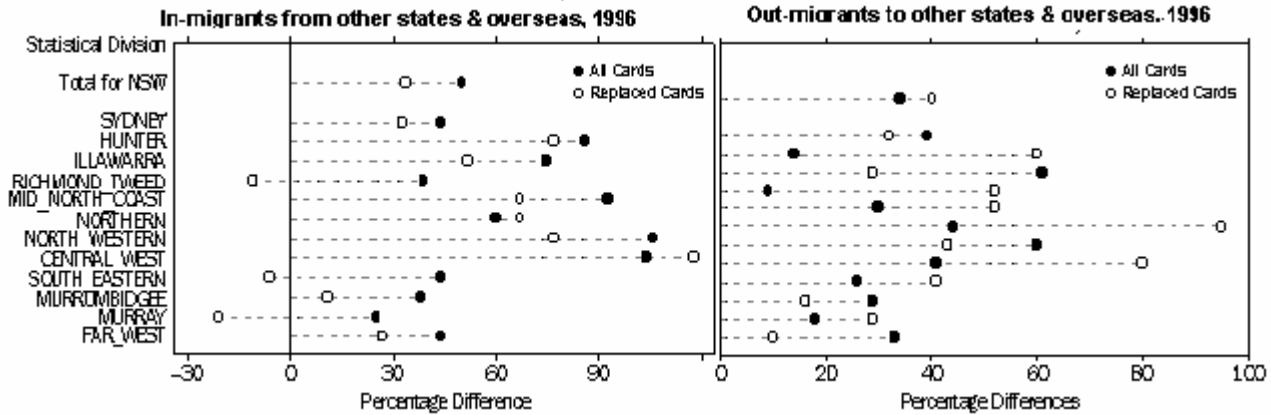
39. The positive differences shown in the graphs indicate that in most SDs, more migrants were estimated from the Medicare data than from the Census. This was case for both in-migrants and out-migrants, whether they moved within SLAs or between NSW and other states or overseas. However, there was no common pattern among the different types of migration. Figure 3 shows that estimates of in-migrants from SLAs based on 'all cards' were closer to the Census figures in Sydney, Hunter, Illawarra, North Western and Central West SDs while 'replaced cards' estimates were closer to the Census figures in the remaining SDs. Estimates of out-migrants to SLAs based on 'all cards' were closer to the Census figures in Sydney, Illawarra, Hunter, Richmond-Tweed and Mid-North Coast SDs while 'replaced cards' estimates were closer to the Census figures in the remaining SDs.

Figure 3. Migrants within NSW, percentage differences between Medicare estimates and Census figures by SDs, 1996.



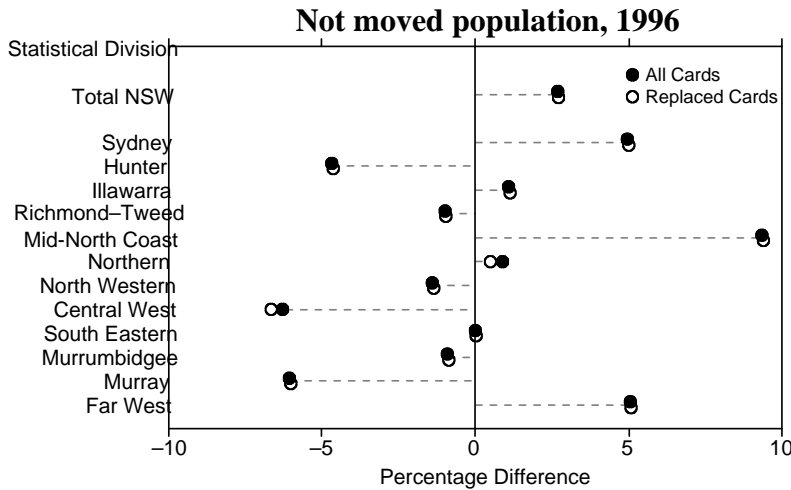
40. Figure 4 shows that in ten out of twelve SDs' estimates of in-migrants from other States and overseas, based on 'replaced cards', were closer to the Census figures than the estimates based on 'all cards'. Estimates based on 'all cards' were closer to the Census figures in seven out of the twelve SDs. In no SDs were the Medicare estimates of out-migrants to other States or overseas below the Census figures.

Figure 4. Migrants across NSW, percentage differences between Medicare estimates and Census figures by SDs, 1996.



41. Figure 5 shows the not-moved category. The similarity between estimates based on 'all cards' and the 'replaced cards' was the product of the methodology. The differences between them and the Census figures were relatively small due to the large base of the not-moved population and were equally distributed around zero.

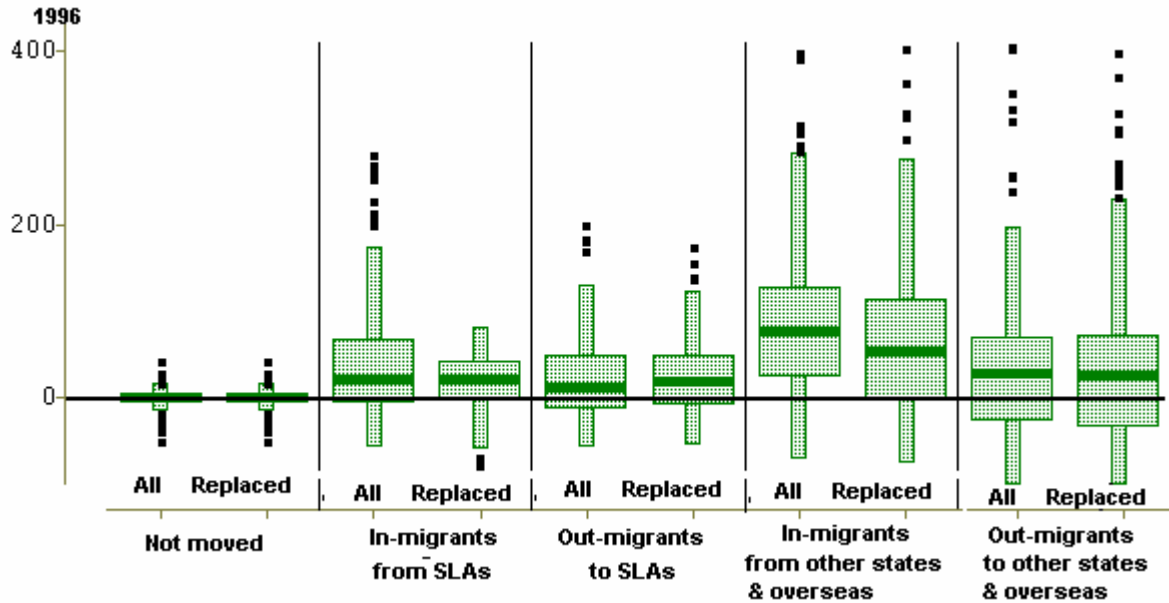
Figure 5. Not moved population, percentage differences between Medicare estimates and Census figures by SDs, 1996.



42. The box plots in Figure 6 represent the SLA distribution of the percentage differences between Medicare estimates and Census figures for each type of move. The solid line in the middle of the box plot marks the median of differences, or 50th percentile. The top and bottom edges of the wide box mark the quartiles, or the 25th and the 75th percentiles. The narrow boxes extending to the top and bottom are called whiskers. Whiskers extend from the quartiles to the farthest observation not farther than 1.5 times the distance between the quartiles. Any more extreme observation, are plotted with individual markers. The position and shape of the box plots compare the percentage

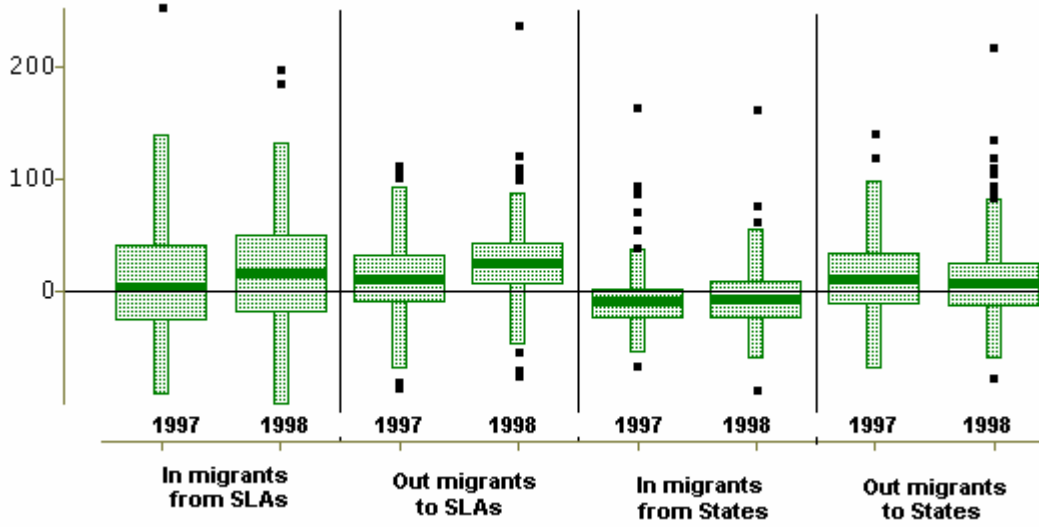
distribution of the differences between each of the Medicare estimates and the 1996 Census figures of the same movement type. The location of the box plots in relation to the line drawn at zero point indicates the positive and negative differences across the 188 SLAs.

Figure 6. Percentage differences between Medicare estimates and the 1996 Census figures by types of move at SLA level.



43. Figure 7 below highlights the difference between estimates produced by the ‘all cards’ and ‘replaced cards’ methods for 1997 and 1998 data. While the distribution of differences among different types of moves were marked, these differences were relatively consistent across the two years.

Figure 7. Percentage differences between Medicare estimates based on all cards and replacement cards by types of move at SLA level, 1997 and 1998.



SLA population estimates using the component method

44. SLA population estimates based on the component method were compared with those derived from the regression model. The component method uses vital statistics for births and deaths and Medicare data for the migration component. In doing so, it is important to take into account the fact that the benchmark used was the ERP figures which were estimates in themselves. All estimates are based on a financial year ended 30 June. The analysis examined the differences between the population estimates using the component and regression methods and whether the differences varied between larger and smaller SLAs.

45. Using the component method, the estimate of the population of SLA i at year t , is as follows:

Population SLA i t =

Population SLA i $t-1$ + Births SLA i $t, t-1$ - Deaths SLA i $t, t-1$ ± Net migration SLA i $t, t-1$

Where

$t, t-1$ is the period between 1 July year $t-1$ and 30 June year t , and

SLA i is the Statistical Local Area i .

46. To keep the component model as independent as possible, the base year for calculation is population estimates produced by the regression model for 1996, which have been adjusted to the 1996 Census figures. The population of SLA i as calculated for 1997, was then used as the base for calculating the population of SLA i for 1998. The analysis is based on the percentage differences between the estimates produced by the component method and the regression model. Two sets of estimates were produced, using the component method, one set of estimates used migration components based on the 'all cards' methodology while the other set used migration components based on the 'replaced cards' methodology.

That is

$$(P_{\text{Reg}} - P_{\text{Com All}}) / P_{\text{Reg}} * 100 \quad \text{or} \quad (P_{\text{Reg}} - P_{\text{Com Rep}}) / P_{\text{Reg}} * 100$$

Where

P_{Reg} is the ERP produced by the regression model,

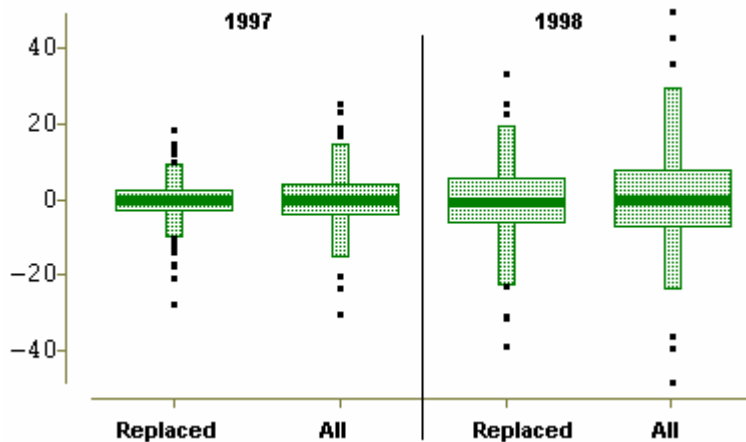
$P_{\text{Com All}}$ is the ERP produced by the component method using ‘all cards’ for the migration component, and

$P_{\text{Com Rep}}$ is the ERP produced by the component method using ‘replaced cards’ for the migration component.

The difference in SLA population estimates using two approaches

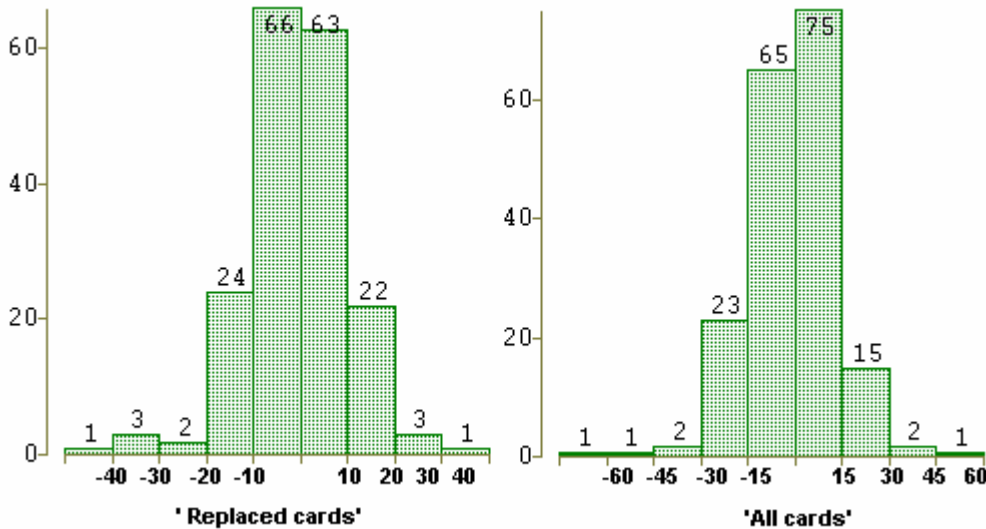
47. Figure 8 below shows box plots of the differences between P_{Reg} and $P_{\text{Com All}}$ or $P_{\text{Com Rep}}$ in 1997 and 1998. The population estimates calculated using the ‘replaced cards’ showed a narrower range of differences, in both 1997 and 1998, when compared with the population estimates based on ‘all cards’ migration.

Figure 8. Percentage differences between SLA populations estimated by regression and by component methods, 1997 and 1998.



48. Figure 9 shows that the distributions of the percentage differences between component estimates and the regression estimates were close to the normal distribution. Both distributions have a similar mean of -0.4 but the differences for the ‘replaced cards’ estimates had a standard deviation of 10.9 compared with a standard deviation of 14.7 for the distribution of differences for the ‘all cards’ estimates. This indicates that the estimates produced using the ‘replaced cards’ method were more consistent with the regression estimates than those estimates produced using the ‘all cards’ method.

Figure 9. The SLA distribution of the two sets of percentage differences between component estimates and the regression estimate, 1998.



49. The following table shows the mean and the standard deviation of the two distributions. The estimates based on the 'replaced cards' method produced less extreme values than the estimates based on 'all cards' method.

	Replaced cards	All cards
Mean	-0.4607	-0.4816
Std Dev	10.9648	14.7051

50. From the observations above, it is possible that estimates of migration based upon 'replaced cards' provide more consistent estimates than those based on 'all cards'. This conclusion can be made because the births and deaths components were the same across the two sets of estimates.

Comparison between SLAs with different size of population

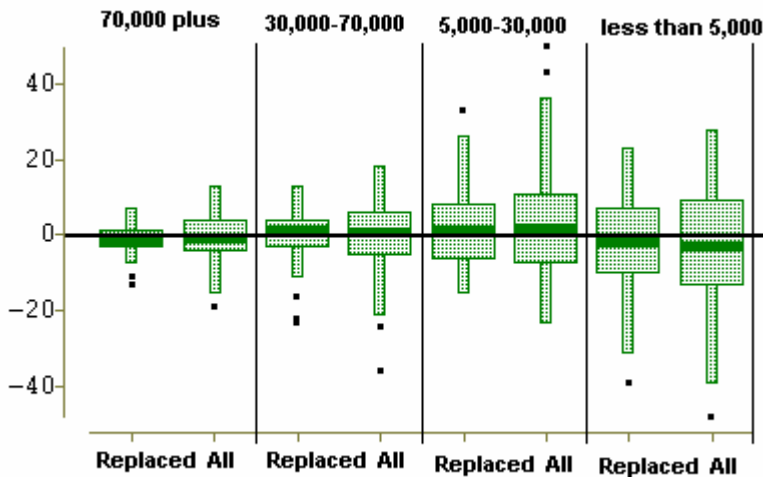
51. The regression model used to estimate population has recently been shown to be less accurate for the smaller populated SLAs than the larger populated SLAs. This analysis was carried out to investigate whether the estimates based on the component method tend to have similar results to the regression estimates. However, the only real way to investigate these differences is to compare the population estimates produced by each method against results from the upcoming 2001 Census.

52. The table below shows the mean and the standard deviation of the differences between each of the two component estimates and the regression estimates for SLAs with various population size. The differences between the component estimates and the regression estimates were the smallest for the larger populated SLAs. It can also be seen that the component estimates based on 'card replacement' were more consistent with the regression estimates in all SLAs with different population size than the component estimates based on 'all cards'.

Population size	'All Cards' Mean	'All Cards' Standard Deviation	'Replaced Cards' Mean	'Replaced Cards' Standard Deviation
5000 or less	-0.96	8.51	-0.44	8.19
5000 to 30,000	0.67	6.36	0.61	4.13
30,000 to 70,000	0.13	5.13	0.34	3.53
70,000 plus	0.71	3.03	-0.78	1.64

53. The shape of the box plots in figure 10 confirms the results that the percentage differences between the components estimates and the regression estimates were closer to zero in the large populated areas than in the small areas. Also, the percentage differences between the component estimates based on the replaced cards and the regression estimates had less variation than the percentage differences between the component estimates based on the all cards and the regression estimates.

Figure 10. The SLA distribution of the percentage differences between the two sets of component estimates and the ERP regression estimates, 1998



Conclusion

54. This paper examined the possibility of using the Medicare enrolment file to derive internal migration statistics. It discussed the data problems, the assumptions adopted and the methods used for estimating internal migration. The study showed the potential use of Medicare data as the best available source for internal migration statistics. The methods used provided consistent estimates of internal migration at the SLA level, whether the estimates are based on 'all cards' and adjusted using Census data as a benchmark or based solely on the 'replaced cards'. The study also examined the possible use of migration estimates for the SLA estimates of resident population using the component method. The analysis showed the component population estimates based on 'replaced

cards' migration data were more consistent with the ABS population estimates based on the regression model than those estimates based on 'all cards' migration data. It also showed that the percentage differences between both components estimates and the regression estimates were smaller in large populated SLAs than in smaller areas.

55. The Health Insurance Commission is currently embarking on a development program to improve the quality and coverage of Medicare data. It will undoubtedly lead to improvement in the internal migration estimates. Along with the encouraging results achieved in the first phase, the project is continuing into the second phase using 1999/2000 and 2000/2001 Medicare data with the aim of further developing the methodology to improve the quality of migration data outputs. The planned third phase involves a comprehensive review to evaluate these results against the 2001 Population Census.



APAPaper.12

Note

Views expressed in this paper are those of the author and do not necessarily represent those of the Australian Bureau of Statistics. Where quoted or used, they should be attributed clearly to the author.

Appendix A

Population by SD as estimated by ABS and as recorded in the Medicare file

	<i>ERP</i>				<i>Medicare</i>				<i>% Difference between ERP and Medicare</i>			
	1996	1997	1998	1999	1996	1997	1998	1999	1996	1997	1998	1999
Statistical Divisions												
Sydney	3881136	3933724	3982315	4043298	3801199	3896221	3976947	4062547	2.10294	0.96255	0.13498	-0.4738
Hunter	555154	561687	566911	573276	530203	539658	548319	556494	4.70593	4.08203	3.39073	3.01567
Illawarra	372860	377022	380408	385167	359916	366015	371890	378801	3.59639	3.00725	2.29046	1.68057
Richmond-Tweed	200542	203658	206334	209059	192142	196223	200789	205335	4.37177	3.78906	2.76161	1.81362
Mid North Coast	262441	265145	268042	271024	256635	260891	265988	270522	2.26236	1.63057	0.77222	0.18557
Northern	178579	177149	175773	174649	174755	174677	175276	175704	2.18821	1.41518	0.28355	-0.6004
North Western	117263	117378	117257	117463	113488	114075	115167	115660	3.32634	2.89546	1.81476	1.55888
Central West	172438	172498	172604	173092	165512	166712	167932	169347	4.18459	3.47066	2.78208	2.21144
South Eastern	178939	179803	180109	181207	172869	174602	176659	179102	3.51133	2.97877	1.95291	1.17531
Murrumbidgee	149150	149046	148516	148765	143009	143811	144078	144912	4.29414	3.64019	3.08028	2.65886
Murray	110882	110786	110655	110478	104933	105539	106251	106338	5.66933	4.97162	4.1449	3.89325
Far West	25344	24888	24591	24202	24864	24675	24845	24789	1.9305	0.86322	-1.0223	-2.368
New South Wales	6204728	6272784	6333515	6411680	6039525	6163099	6274141	6389562	2.73536	1.77971	0.94633	0.34616

Appendix B

Outcomes of the Census adjusted method

1. Adjusted of migrants within the SLAs of NSW

Statistical Divisions	In Migrants from SLAs			% Change		Out migrants to SLAs			% Change	
	1996 Census estimates	1997	1998	-1	-1	1996 Census estimates	1997	1998	-1	-1
Sydney	270305	301711	280015	11.62	-7.19	276878	309135	285672	11.65	-7.59
Hunter	34747	34777	33069	0.09	-4.91	31552	30690	30057	-2.73	-2.07
Illawarra	17444	18325	16893	5.05	-7.81	14951	14921	14457	-0.2	-3.11
Richmond-Tweed	12410	12184	10893	-1.82	-10.6	9952	10953	9780	10.06	-10.71
Mid North Coast	14446	15019	16460	3.96	9.6	12524	12770	14509	1.96	13.62
Northern	11790	11373	9631	-3.54	-15.32	13130	13407	11504	2.11	-14.19
North Western	6943	6860	6032	-1.2	-12.08	8240	7893	6885	-4.21	-12.77
Central West	11799	10600	9328	-10.16	-12	11613	11787	9796	1.5	-16.89
South Eastern	8939	8796	7672	-1.61	-12.78	9052	7971	7253	-11.95	-9
Murrumbidgee	7507	7106	6330	-5.34	-10.91	8043	8119	6809	0.95	-16.14
Murray	3780	3784	2733	0.11	-27.78	3930	3864	2693	-1.67	-30.31
Far West	639	531	534	-16.95	0.64	878	864	987	-1.61	14.29
New South Wales	400749	430462	398450	7.41	-7.44	400743	430454	398444	7.41	-7.44

Appendix B

Outcomes of the Census adjusted method

2. Adjusted of migrants between SLAs and other States and overseas

Statistical Divisions	In Migrants from States & overseas			% Change		Out migrants to States & overseas			% Change	
	1996 Census estimates	1997	1998	-1	-1	1996 Census estimates	1997	1998	-1	-1
Sydney	109579	110271	99217	0.63	-10.02	53074	54252	51776	2.31	-4.56
Hunter	8437	8577	9057	1.66	5.59	8165	10404	10210	27.54	-1.87
Illawarra	6508	5598	5637	-13.99	0.7	5605	5007	5093	-10.59	1.73
Richmond-Tweed	7170	8239	7352	14.91	-10.76	7424	7917	7578	6.74	-4.28
Mid North Coast	4377	4287	4234	-2.05	-1.25	5378	5564	5500	3.54	-1.15
Northern	3256	2865	2875	-12.02	0.38	4818	4259	4013	-12.38	-5.79
North Western	1726	1490	1690	-13.7	13.45	2223	2054	1973	-7.53	-3.93
Central West	2495	2192	1639	-12.13	-25.25	3329	3155	2888	-6.48	-8.47
South Eastern	5752	4789	4963	-16.75	3.64	6132	5467	5204	-10.77	-4.8
Murrumbidgee	3851	3490	3411	-9.36	-2.28	4745	4032	4188	-14.95	3.89
Murray	5299	4941	4528	-6.75	-8.37	5849	5862	5276	0.31	-10
Far West	558	652	697	16.83	6.87	927	883	789	-4.74	-10.58
New South Wales	159008	156985	145474	-1.27	-7.33	107670	108308	103983	0.59	-3.99

Appendix B

Outcomes of the Census adjusted method

3. Adjusted not moved population and net migration in NSW

	<i>Not Moved Population</i>			<i>% Change</i>		<i>Net Migrants</i>		
	1996 Census estimates	1997	1998	-1	-1	1996 Census estimates	1997	1998
Statistical Divisions								
Sydney	3166895	3226177	3344761	1.87	3.68	49932	48546	41736
Hunter	494427	502757	517678	1.68	2.97	3466	2250	1849
Illawarra	325639	332390	339894	2.07	2.26	3396	3990	2974
Richmond-Tweed	169343	173243	179044	2.3	3.35	2204	1546	881
Mid North Coast	207325	210993	208068	1.77	-1.39	921	967	680
Northern	153279	153568	155015	0.19	0.94	-2902	-3392	-2975
North Western	100731	101255	103568	0.52	2.28	-1794	-1599	-1139
Central West	156323	157894	162070	1.01	2.64	-648	-2108	-1679
South Eastern	152051	155153	158421	2.04	2.11	-492	142	172
Murrumbidgee	128587	129795	131968	0.94	1.67	-1430	-1558	-1260
Murray	95493	96860	100017	1.43	3.26	-700	-1006	-713
Far West	22035	21901	21837	-0.61	-0.29	-608	-565	-546
New South Wales	5172128	5261987	5422340	1.74	3.05	51344	47213	39981

Appendix C

Outcomes of the card replacement method

1. Adjusted of migrants within the SLAs of NSW

Statistical Divisions	In Migrants from SLAs			% Change		Out migrants to SLAs			% Change	
	1996	1997	1998	-1	-1	1996	1997	1998	-1	-1
Sydney	328,843	399,129	396,414	21.77	-0.40	331,479	401,045	397,879	20.60	-1.07
Hunter	44,956	48,053	47,504	7.23	-0.87	42,797	44,948	45,761	4.69	1.52
Illawarra	18,241	22,395	23,284	23.17	4.26	18,240	19,526	20,825	6.70	6.35
Richmond-Tweed	11,068	14,957	13,814	35.58	-7.39	13,562	13,242	12,726	-2.68	-4.17
Mid North Coast	19,227	21,767	22,387	13.57	3.14	20,185	22,081	23,618	9.04	6.66
Northern	15,853	16,686	16,131	5.59	-3.06	14,491	17,954	16,122	23.50	-10.46
North Western	11,559	11,770	11,346	2.16	-3.33	10,586	11,544	10,916	8.70	-5.70
Central West	14,667	14,296	14,812	-2.21	3.90	13,888	15,228	14,918	9.29	-2.31
South Eastern	9,807	9,128	9,444	-6.63	3.75	10,537	10,589	10,764	0.16	1.37
Murrumbidgee	9,594	8,235	8,032	-13.89	-2.19	8,864	10,374	9,929	16.66	-4.56
Murray	6,782	6,520	5,751	-3.56	-11.54	6,099	6,321	5,688	3.30	-10.26
Far West	750	676	927	-9.48	37.43	617	760	700	22.75	-8.11
New South Wales	491,346	573,612	569,846	17.12	-0.38	491,346	573,612	569,846	16.37	-0.94

Appendix C

Outcomes of the card replacement method

2. Adjusted of migrants between SLAs and other States and overseas

	<i>In Migrants from States & overseas</i>			<i>% Change</i>		<i>Out Migrants from States & overseas</i>			<i>% Change</i>	
	1996	1997	1998	-1	-1	1996	1997	1998	-1	-1
Statistical Divisions										
Sydney	146,231	148,814	142,478	1.77	-4.26	70,091	74,590	71,206	6.42	-4.54
Hunter	14,898	15,280	15,037	2.56	-1.59	13,100	14,386	13,301	9.82	-7.54
Illawarra	9,882	9,630	9,695	-2.55	0.67	7,231	7,430	7,697	2.75	3.59
Richmond-Tweed	6,413	5,464	6,702	-4.50	9.42	11,307	10,296	9,533	-12.60	-3.54
Mid North Coast	7,298	6,959	7,233	-4.65	3.93	8,169	9,194	9,484	12.54	3.16
Northern	5,427	4,723	4,517	-12.98	-4.34	9,418	7,226	6,299	-23.27	-12.82
North Western	3,048	3,237	3,246	6.20	0.29	3,181	3,455	3,375	8.59	-2.30
Central West	5,432	4,342	3,910	-20.06	-9.96	5,988	5,482	4,818	-8.45	-12.12
South Eastern	5,399	5,531	5,601	2.43	1.28	8,622	7,974	7,057	-7.52	-11.50
Murrumbidgee	4,293	4,162	4,010	-3.05	-3.65	5,497	5,646	5,404	2.71	-4.29
Murray	4,163	3,794	3,626	-8.86	-4.44	7,542	6,431	5,850	-14.73	-9.04
Far West	706	640	755	-9.43	18.02	1,019	929	832	-8.85	-10.45
New South Wales	213190	212575	206,810	0.02	-3.01	151,165	153,039	144856	0.97	-5.09

Appendix C |

Outcomes of the card replacement method

3. Adjusted not moved population and net migration in NSW

Statistical Divisions	Not moved population			% Change		Net Migrants		
	1996	1997	1998	-1	-1	1996	1997	1998
Sydney	3,323,645	3,385,861	3,510,315	1.87	3.68	73,504	72,308	69,808
Hunter	471,307	479,248	493,471	1.68	2.97	3,959	4,446	3,479
Illawarra	329,176	336,000	343,585	2.07	2.26	2,651	5,270	4,457
Richmond-Tweed	167,664	171,526	177,269	2.30	3.35	-7,388	-2,981	-1,742
Mid North Coast	226,683	230,694	227,496	1.77	-1.39	-1,829	-2,338	-3,482
Northern	153,992	154,273	156,371	0.18	1.36	-2,629	-3,605	-1,773
North Western	99,320	99,836	102,117	0.52	2.28	839	120	301
Central West	145,856	147,322	151,218	1.01	2.64	222	-1,930	-1,014
South Eastern	152,045	155,146	158,414	2.04	2.11	-3,953	-3,810	-2,776
Murrumbidgee	127,420	128,617	130,770	0.94	1.67	-475	-3,534	-3,291
Murray	89,709	90,994	93,959	1.43	3.26	-2,696	-2,377	-2,162
Far West	23,147	23,006	22,939	-0.61	-0.29	-180	-366	150
New South Wales	5,309,964	5402523	5,567,924	1.74	3.06	62,025	61,203	61,954