SHORT-TERM FORECAST ERROR OF AUSTRALIAN LOCAL GOVERNMENT AREA POPULATION PROJECTIONS

Tom Wilson
Northern Institute, Charles Darwin University, Darwin, NT, 0909, Australia.
Email: tom.wilson@cdu.edu.au.

ABSTRACT: Local government area population projections produced by state and territory governments are regularly subject to criticism for their supposed inaccuracy. This paper examines the 2006 round of such projections for five states, assessing their forecast accuracy after five years. It is demonstrated that, overall, the projections are quite accurate over this five year period relative to both user needs and simple extrapolations which constitute a basic benchmark. It is shown how the error distributions of these projections can be used to create approximate prediction intervals indicating the likely range of error in current local area projections.

KEY WORDS: Population projections; forecast error; prediction interval; local government area

DECLARATION: The author was responsible for producing the 2006-based New South Wales local government area population projections evaluated in this paper.

1. INTRODUCTION

Most State and Territory governments prepare population projections for the local government areas of their jurisdiction on a regular basis (e.g. New South Wales Department of Planning and Environment, 2014; Victoria Department of Transport, Planning and Local Infrastructure, 2014; Queensland Government Statistician's Office, 2013). The projections are used in a wide variety of planning, service provision, policy development and other related activities in both the public and private sectors (e.g. Diamond *et al.*, 1990; Taylor, 2015). For example, they inform state and local governments' regional and local planning strategies, future housing requirements, education demand, health service provision, transport modelling and infrastructure planning, market assessments and retail site selection. Collectively the projections influence investment decisions worth billions of dollars annually.

Local area projections often generate considerable attention from the media and councils when published. Some commentary is fairly neutral and concerned with describing and discussing the projected demographic changes; other comments are decidedly negative, usually when the projections for a particular local area indicate steady population numbers or decline (Johnstone, 2015). Projections which fail to align with a council's outlook for the future of its area tend to be criticised as inaccurate, and the projection models from which they are generated are often deemed flawed. Furthermore, projections of population decline are seen as dangerous because they risk discouraging investment and employment, thus becoming self-fulfilling prophesies.

Subsequent monitoring of projection accuracy through comparisons with the annual Estimated Resident Populations (ERPs) published by the ABS can generate further debate. Some commentators give mixed reviews of local area projections (e.g. Salt, 2014), but others are more critical. For the latter group, even small differences between a projection and an ERP in the first few years of the projection period are viewed as a failure of the projections, sometimes prompting questions of competency of those who produced them.

These sorts of criticisms, which are often directed at the projections for a single local government area, might give the impression that State and Territory governments in Australia do a poor job of forecasting local area populations. This paper examines whether this is the case, focusing specifically on the first five years of the projections. The reason for this short-term focus is three-fold: (i) most attention is placed on projections when they are the latest available set, and after five years usually an updated set of projections will have been produced, (ii) short-term errors provide a clue to longer-term errors because forecast error usually increases over time, and (iii) many stakeholders make use of short-term projections (e.g. Diamond *et al*. 1990; Carey, 2011; Salt, 2014; Dovey, 2015). Examining forecast error is also helpful in providing users with an indication of the likely range of error in a current set of population projections, and allowing demographers to assess whether there are obvious problems with projections that could be corrected in the future.

Specifically, this paper reports on an evaluation of the five-year forecast error of the 2006 round of local government area population projections from those states that published projections shortly after the necessary 2006 Census and ERP data became available. These jurisdictions are New South Wales, Victoria, Queensland, South Australia and Tasmania. Western Australia, the Northern Territory and the Australian Capital Territory are excluded, WA because its 2006-based projections were released after 2011, the NT because it did not produce local government area projections in the 2006 round of projections, and the ACT because it has no local government areas. Both total and age-specific populations are assessed in this analysis. Projected demographic components of change (births, deaths and migration) could not be assessed as they were not published.

What magnitude of error might be expected in the local government area projections? There is limited existing literature on this specific topic, but in an analysis of several past sets of Queensland local government area projections, Wilson and Rowe (2011) found average errors for the total population to be between 4 and 6 per cent after 5 years. Similarly, in an assessment of several rounds of county population projections for Florida, Smith and Rayer (2011) discovered average errors after 5 years to be also between 4 and 6 per cent. The limited amount of previous research on local area population forecast error has revealed that error generally increases as population size decreases, and in many cases areas with particularly high growth or decline in the recent past also tend to experience higher forecast errors (e.g. Rayer, 2008; Statistics New Zealand, 2008; Tayman *et al*. 2011). In Queensland, Wilson and Rowe (2011) found errors varied from just 3 or 4 per cent for areas with 50 000 or more people, 5 to 7 per cent for populations of 5 000 to 15 000, and around 9 to 11 per cent for those with fewer than 2 000 people. They also discovered that areas with relatively large shares of their populations in mining employment or identifying as Indigenous tended to experience larger than average errors. Few studies have examined local area age-specific forecast errors (although exceptions include Rayer and Smith, 2014; and Statistics New Zealand, 2008). Rayer and Smith's analysis of county population projections in Florida showed the highest errors to be found at the young adult and very oldest ages.

On the matter of terminology, a population *projection* is strictly any numerical statement about the future of a population whether plausible or not. A population *forecast* on the other hand is a projection deemed to describe the most likely demographic future. However, in this paper the data under assessment are labelled 'population projections' because this term is widely used by State and Territory governments. But given that most users regard them as *forecasts* they are evaluated as such, and thus reference is made to their 'forecast error' (following Smith, 1987).

2. DATA AND METHODS

Population Projections Data

Local government area population projections were obtained shortly after their publication from the websites of the relevant government departments responsible for projections in New South Wales, Victoria, Queensland, South Australia and Tasmania (NSW Department of Planning, 2009; Victorian Department of Planning and Community Development, 2009; Queensland Department of Infrastructure and Planning, 2008; South Australian Department of Planning and Local Government, 2011; Tasmanian Demographic Change Advisory Council, 2008). Total populations were obtained for all five of these states, whilst projections for five year age groups were available for all states except Queensland. Projections were 2006-based for New South Wales, Victoria, Queensland and South Australia and 2007-based for Tasmania, and the projections data extracted for analysis were those for 2011 (four states) and 2012 (Tasmania). All projections data refer to 30th June of the reference year.

The study examined projections for 152 LGAs in New South Wales, 77 in Victoria, 60 in Queensland, 68 in South Australia and 29 in Tasmania (totalling 386). Only 18 LGAs had to be excluded from the analysis because of either boundary changes or missing data – the latter being a few small Indigenous councils in Queensland for which projections were not published. All remaining 386 local government areas did not undergo boundary changes over the period, or did so with negligible population changes.

All sets of projections were produced from cohort-component models using fertility, mortality and migration assumptions. For urban areas in NSW, Victoria, and Queensland data on anticipated dwelling growth were also taken into account, often via a housing-unit model used in concert with the cohort-component model. Projections of total populations for Queensland LGAs outside South East Queensland were prepared using a ratio-share method to which were constrained age-sex projections from the cohort-component model. Methodological information was not published for the South Australian LGA projections, although some form of cohort-component model will have been used.

Benchmark Projections Data

The state's local government area projections were compared against a benchmark set of projections created from a basic, naïve and automated method. Naïve 2006-based projections of total population were generated from a linear extrapolation of population change over the decade 1996-2006. A comparison of official projections against these naïve projections indicates the extent to which state government demographers provide greater value (or not) than a basic, but very quick and low-cost, projection method.

Population Estimates Data

Projections were assessed against ABS Estimated Resident Populations (ERPs) for local government areas (ABS, 2013a). These ERPs take into account the 2011 Census and are final up to 2011 and provisional for later years. However, these ERPs are problematic in that they are inconsistent with the ERPs available at the time the projections were produced. Following the 2011 Census ABS decided to 'recast' its ERP series back to 1991, superseding all previous ERP data (ABS, 2013b). During the 2011 Census evaluation ABS used a new and improved method of estimating census net undercount, finding that in earlier years it had over-compensated for undercount in creating its ERPs. Nationally, the ERP for 2011 was about a quarter of a million lower than it would have been if the old undercount adjustment had been applied. Without making allowance for the ERP recasting, any evaluation of the projections would be completely unreliable. The solution applied here is to calculate error measures which allow for the discontinuity (described below).

Error Measures

Forecast error is defined as the population forecast minus the ERP, and is often expressed as a ratio of ERP to standardise for population size. This is Percentage Error (PE):

$$PE_t = \frac{F_t - ERP_t}{ERP_t}\ 100\% \qquad (1)$$

where F denotes the population forecast and t a forecast year. To allow for ABS's ERP recasting a modified error measure used by Keilman

(1999) is applied. This is Corrected Percentage Error (CPE), and it removes the difference between the old and recast 2006 (or 2007) jump-off ERPs. It is calculated as:

$$CPE_t = \frac{F_t - ERP_t^{recast} - (ERP_0^{old} - ERP_0^{recast})}{ERP_t^{recast}} 100\% \qquad (2)$$

where 0 denotes the jump-off year. Corrected Percentage Error is not a perfect solution, however, because it does not make allowance for the projected fertility, mortality and migration assumptions which might have been prepared if the producers of the projections had had the recast ERPs available to them at the time.

The principal measure used to report average error across all local government areas in each state is the Weighted Mean Absolute Percentage Error (WMAPE). This is a weighted mean of the absolute (in this case Corrected) Percentage Errors, with weights defined as each local government area's share of all local government area populations for the forecast year (Siegel, 2002; see also Wilson, 2012). It is also known as the Mean Percentage Absolute Difference (MPAD) (Murdock *et al*., 1984). WMAPE may be calculated as:

$$WMAPE_t = \sum_i \left(\left| CPE_t^i \right| \frac{ERP_t^{i,recast}}{\sum_i ERP_t^{i,recast}} \right) \qquad (3)$$

where i is a local government area. WMAPE is relevant when there is a wide range of local area population sizes and preferable in such cases to the Mean Absolute Percentage Error (MAPE), which effectively weights all observations equally:

$$MAPE_t = \sum_i \left(\left| CPE_t^i \right| \frac{1}{n} \right) \qquad (4)$$

where n is the number of observations. Population forecast errors tend to form right-skewed distributions where a small number of high errors result in the mean being regarded as unrepresentative of 'average' error (Tayman and Swanson, 1999). An alternative is the Median Absolute Percentage Error (MedAPE) which is the middle of a set of ranked absolute CPEs. Although WMAPE is the preferred measure in this paper, both MAPE and MedAPE are reported alongside WMAPE in Table 1 to facilitate comparison with other studies.

In addition to reporting average errors, the distributions of errors across LGAs are also presented. Low average errors are impressive, but if there are a few highly erroneous outliers then there remain problems for users and producers of the projections. For each local government area absolute values of Corrected Percentage Errors are categorised into those (i) below 5 per cent, (ii) 5-10 per cent, (iii) 10-20 per cent and (iv) 20 per cent or more. These are regarded here as (i) good, (ii) acceptable, (iii) poor, and (iv) bad on the basis of Tye's (1994) finding that most users consider errors up to 10 per cent as acceptable.

Forecast bias is also briefly reported. This refers to whether projections were too high or too low on average across all LGAs. The measure used is Mean Percentage Error (MPE) which is simply the mean of all Corrected Percentage Errors. A positive value indicates the projections were too high overall; a negative value shows they were too low.

3. RESULTS

Error in Projecting Total Populations

Table 1 presents the average errors of local government area population projections after five years. All averages as measured by WMAPE are below 3 per cent, indicating fairly accurate projections overall. The fifth column of the table reports WMAPE errors from the naïve linear extrapolation. It can be seen that in all cases the states' projections outperformed the naïve model – as shown by the right-hand most column of the table. In the cases of Queensland, New South Wales and South Australia the official projections achieved less than half the average error of the naïve projections. The results confirm that the efforts and expense made by the states in preparing their projections paid off in terms of greater accuracy. Across all five states the average error of the states' projections was 2.4 per cent compared to 5.3 per cent for the naïve projections.

Table 1. Average Error of Local Government Area Population Forecast Errors after Five Years.

	States' projections			**Naïve projection**	**Additional value**
State	**WMAPE (A)**	**MedAPE**	**MAPE**	**WMAPE (B)**	**(B) – (A)**
	per cent				
NSW	2.7	2.6	3.1	5.5	2.8
Vic	2.4	2.2	2.4	4.1	1.8
Qld	2.5	2.9	4.8	7.5	5.1
SA	1.4	2.0	3.8	3.1	1.7
Tas	2.0	1.8	2.7	2.6	0.6
All	2.4	2.3	3.3	5.3	2.9

Source: the Author.

The naïve projection errors also play another role: they can be viewed as a measure of the degree of difficulty in producing accurate projections in each state. Although the average errors for South Australia and Tasmania's projections were the lowest, their small naïve errors indicate that projecting local government area populations in these states was relatively easy. Population growth tends to be relatively steady and predictable. Conversely, while the average error for Queensland's local government area projections wasn't especially low, the naïve error indicates that this was the most challenging jurisdiction for which to produce local area projections.

Figure 1 presents an alternative perspective on forecast error by showing the distribution of absolute errors across LGAs. For example, in NSW 84 per cent of LGAs were projected with absolute CPE under 5 per cent after 5 years (good), 15 per cent had errors of between 5 and 10 per cent (acceptable), and just 1 per cent experienced errors of 10-20 per cent (poor). In Victoria and Tasmania no LGAs had errors exceeding 10 per cent. Queensland and South Australia experienced some errors between 10-20 per cent while a small proportion had errors exceeding 20 per cent. Across all five states 83 per cent of LGAs had absolute errors under 5 per cent, 12 per cent had errors of between 5 and 10 per cent, 4 per cent between 10 and 20 per cent, whilst 1 per cent of LGAs experienced errors of 20 per cent or more.

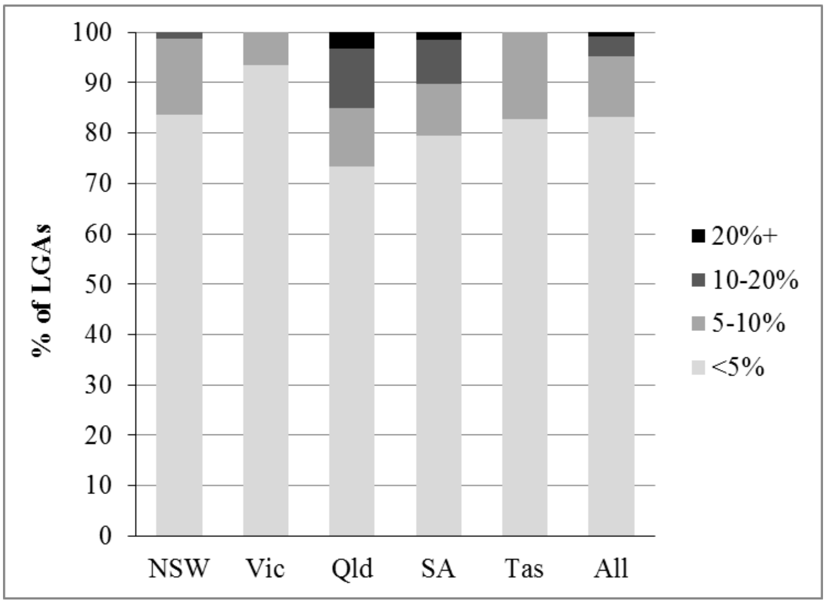

Figure 1. The Distribution of Absolute Corrected Percentage Errors across Local Government Areas After 5 Years. Source: the Author.

In terms of bias, LGA projections in NSW and Queensland proved a little low overall with Mean Percentage Errors after 5 years of -1.2 per cent and -2.2 per cent respectively. In Victoria, South Australia and Tasmania the projections were a little high overall with MPE values of 0.2 per cent, 2.8 per cent and 0.6 per cent respectively. Across all five states MPE was -0.2 per cent, indicating little bias overall. All these values are fairly low and demonstrate that bias is not a significant issue.

Why were 5 per cent of LGAs' populations forecast with large errors (exceeding 10 per cent)? An examination of projection assumptions and/or local area characteristics for individual LGAs can be useful in diagnosing the causes of error. In NSW two LGAs were forecast with more than 10 per cent error. One was Camden, an area of south-west Sydney undergoing residential development. The population projection for Camden was driven by optimistic dwelling forecasts, and although the area's population grew by 15 per cent over the 2006-11 period, it was less than anticipated. The other LGA was Murrumbidgee, home to about 2 600 people in 2006. The population projection was for very slight growth, but in fact population declined. Census data show significant job losses in the ‘agriculture, forestry and fishing’ and ‘manufacturing’ industries over the 2006-11 period (ABS, 2012).

In Queensland nine LGAs experienced more than 10 per cent error, eight of which had populations under 5 000. The one LGA with a sizeable population was the mining town of Mount Isa. Substantial population growth in this LGA was projected for 2006-11 due to resource development (Queensland Department of Infrastructure and Planning, 2008 p.29), but more moderate growth eventuated. Employment in mining and associated construction is often subject to considerable volatility due to fluctuations in global commodity prices, and is therefore very hard to predict. The nine small LGAs with large errors were Aurukun, Blackall Tambo, Boulia, Bulloo, Carpentaria, Cook, Croydon, and McKinlay. All of these LGAs were under-projected, and all experienced upward changes in the direction of their population trends from 2006 or 2007. Such results confirm the long-established fact that projections, created from extrapolative models in this case, tend to be accurate when population trends are on a ‘business as usual’ setting. The challenge remains to predict discontinuities and turning-points. However, it is possible that some of the problem may be due to ERPs rather than projections. It is well known that Indigenous census counts have a tendency to fluctuate and be inconsistent from one census to the next, resulting in ERP reliability issues for areas with large Indigenous populations. For example, Aurukun (over 90 per cent Indigenous) recorded very large increases in its Indigenous ERP between 2006 and 2011 (ABS, 2008; 2013c).

In South Australia the seven LGAs of Anangu Pitjantjatjara, Cleve, Coober Pedy, Elliston, Karoonda East Murray, Maralinga Tjarutja, and Peterborough experienced errors of more than 10 per cent. A modest increase in population was projected for Anangu Pitjantjatjara but subsequently published ERPs reveal that the population actually increased substantially between 2006 and 2011. This was the only LGA under-projected. For the other six LGAs negligible total population change was projected, but in reality they all lost population.

In summary, the most erroneous projections were largely amongst non-metropolitan LGAs that were either small, had significant mining employment or significant Indigenous populations. The findings confirm those of Wilson and Rowe (2011).

Error in Projecting Total Population by Population Size

The upper panel of Table 2 shows how errors of the states' projections varied by population size (measured as size at the start of the projections). Previous research has found error is usually greater for smaller populations. To some extent this finding is reflected in the table, though there is little difference between the two larger population categories, and only modest differences between the middle two categories. A possible contributor to the fractionally larger error in the 50 000+ population category is overseas migration: these larger LGAs attract more overseas migration and are likely to experience greater annual fluctuations in population growth due to the volatility of net overseas migration trends. The main finding here is that population projections for LGAs with just a few thousand (or even a few hundred) people are very difficult to get right, even over the short-term.

Table 2. WMAPEs of Local Government Area Population Projections after Five Years, by Population Size Category.

State	**Population size**			
	0 – 4,999	**5,000 – 19,999**	**20,000 – 49,999**	**50,000+**
	per cent			
	States' projections			
NSW	4.2	3.0	2.6	2.7
Vic	*	2.7	2.3	2.3
Qld	6.7	*	2.7	2.4
SA	5.1	2.4	0.7	*
Tas	*	2.4	*	*
All	4.9	2.7	2.2	2.4
	Naïve projections			
NSW	6.7	4.5	5.7	5.5
Vic	*	4.2	4.0	4.1
Qld	7.8	*	4.8	7.9
SA	7.7	3.8	4.3	*
Tas	*	3.3	*	*
All	7.2	4.1	4.8	5.5

Note: * Values not shown for categories with fewer than 10 observations. Source: the Author.

The lower panel of Table 2 shows show the naïve linear extrapolations fared by population size category. For every size category and every state the naïve projections were less accurate. Interestingly, the states' projections increased their accuracy relative to the naïve projections as population size increased. For the smallest 0-4 999 category the average error of the states' projections was 69 per cent of the naïve projections (4.9 per cent versus 7.2 per cent WMAPE), while for the 50 000+ category it was 44 per cent (2.4 per cent versus 5.5 per cent).

Error in Projecting Total Population by Growth Rate

There was some variation in error according to population growth rates over the preceding 2001-06 period. Table 3 presents average errors for LGA projections by state, indicating some agreement with previous research which shows that areas with the highest positive or negative growth rates in the recent past tend to experience the largest forecast errors. This may be related to the fact that past growth rates do not always provide a good indication of the future. There does seem to be a tendency for 'regression to the mean' in local area population trends in which areas experiencing the largest growth or decline in one period often grow at a rate closer to the average in the next (Wilson, 2014; Smith, 1987). However, many of the areas which underwent the largest declines over the 2001-06 period also had very small populations so growth rate versus population size effects are difficult to disentangle in this study. On the matter of bias, very little bias was evident in the projections for the middle three categories of growth rate. Interestingly, the MPE for areas declining the most (<-2 per cent) was -2.6 per cent indicating slight under-projection overall for these areas, while in the highest growth category (2 per cent+) there was slight over-projection (MPE of 1.7 per cent). It suggests that, overall, the projections would have benefitted from slightly less projected decline in areas which had recently declined the most and slightly less growth in areas which had recently grown the fastest – i.e. just a little more regression to the mean.

Table 3. WMAPEs of Local Government Area Population Projections after Five Years, by Growth Rate Category.

State	**Annual average growth rate, 2001-06**				
	<-2.0%	**-2.0 – -0.5%**	**-0.5 – 0.5%**	**0.5 – 2.0%**	**2.0%+**
	per cent				
NSW	*	3.7	2.5	2.5	*
Vic	*	3.8	1.9	2.8	*
Qld	7.5	1.8	*	2.4	2.4
SA	*	*	1.5	1.1	*
Tas	*	*	2.2	1.7	*
All	4.9	3.2	2.3	2.3	2.6

Note: * Values not shown for categories with fewer than 10 observations. Source: the Author

Error in Projecting Total Population by Density

LGAs were classified into three population density categories: high density (>200 persons per km^2), medium density (10-200 persons per km^2), and low density (up to 10 persons per km^2). Average errors in these three categories are shown in Table 4. Across all five states low population density LGAs were projected slightly less accurately than those in the medium and high density categories. However, many of these low density LGAs also had very small populations. The key finding is that projections in all three density categories were quite good overall and that there is no obvious urban or rural factor substantially affecting error.

Table 4. WMAPEs of Local Government Area Population Projections after Five Years, by Population Density Category.

	High density	**Medium density**	**Low density**
	per cent		
NSW	2.9	2.0	3.1
Vic	2.6	1.7	2.4
Qld	*	2.9	3.1
SA	1.0	1.9	3.0
Tas	*	1.9	2.5
All	2.4	2.2	2.9

Note: * Values not shown for categories with fewer than 10 observations. Source: the Author.

Error in Projecting Specific Age Groups

Many users of projections have only a passing interest in the errors of projected total populations, focusing primarily on specific age groups related to the services they provide (e.g. education, prisons, aged care). How well were age-specific populations projected? Figure 2 illustrates WMAPEs by age for New South Wales, Victoria, South Australia and Tasmania, plus all four of these states combined. Projections by age group were not published for Queensland. Age group 0-4 is not shown in the graph because Corrected Percentage Errors require the start-of-period cohort population in their calculations, in this case births, which would be conceptually inconsistent with the other ‘corrected’ errors.

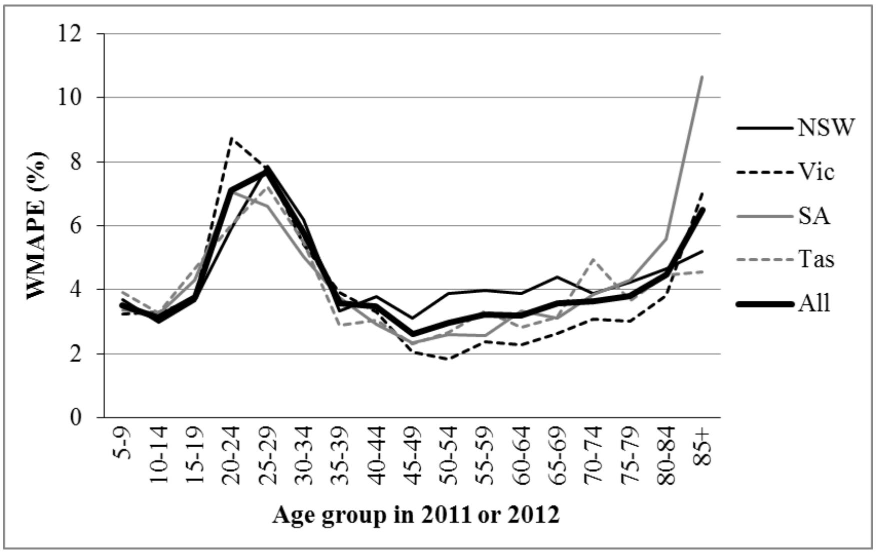

Figure 2. WMAPEs of Local Government Area Population Projections After Five Years, by Age Group. Source: the Author.

Age-specific projections which had average errors under 5 per cent, and which may be classified as ‘good’, were the childhood ages, and most middle and older adult ages. Not surprisingly, average errors were greatest in the 20s and early 30s, which are the most migratory age groups. Migration is the most volatile of the demographic components of change, especially at the local area scale, and is relatively hard to predict. The very oldest age groups also experienced greater error than most. The reasons for this are not obvious, but are likely to be related to the increase in migration rates with age that often occurs at the highest ages at the local area scale, and errors in mortality rate projections at high ages.

Projections for most age groups were on average more erroneous than the total population errors shown in Table 1. This occurs because of a mix of over-projection and under-projection by age group. For example, the average absolute error for the total population projection for South Australia is lower than those at every age group in Figure 2 due to the mix of positive and negative age-specific errors which partially offset one another when summed over all ages.

4. DISCUSSION

Making Use of Past Error Data

The results section of this paper described average errors and the distribution of errors in the form of proportions of LGAs in different error categories. An alternative way of presenting error distributions is to report errors at certain points on the distributions. Table 5 displays selected percentiles of absolute Corrected Percentage Errors of local government area projections across all five states after five years by population size category. For example, the table reports that 80 per cent of LGAs with under 5 000 people had population forecast errors under 9.6 per cent for a five year projection horizon, and that 67 per cent of LGAs with 50 000 or more people had errors under 3.0 per cent.

Table 5. Absolute Corrected Percentage Errors at Selected Percentiles of the Error Distribution by Population Size Category.

	Population size			
Percentile	**0 – 4,999**	**5,000 – 19,999**	**20,000 – 49,999**	**50,000+**
	per cent			
67th	6.7	3.4	2.8	3.0
80th	9.6	4.1	4.4	3.6
95th	19.3	5.9	5.5	5.4

Source: the Author.

These sorts of errors have the potential to be used as ballpark indications of likely future error with the latest sets of local government area, or other similar small area, projections. It makes the significant assumption that the distributions of errors observed in the past will approximate those of the future, although there are studies which lend support to this assumption (e.g. Smith and Sincich, 1988). The emphasis is very much on ballpark indications, however, because factors other than population size clearly affect error and unique local factors may result in especially large errors in some areas. For more sophisticated and comprehensive indications of likely forecast error, probabilistic projections are required (e.g. Bell *et al*., 2011; Wilson, 2013). However, probabilistic models are very data-hungry and have yet to be modified from the national and large region scales for application at the local level.

Although indicative and approximate only, the data in Table 5 are still useful. Imagine a local government area with a population of 15 000 which is projected to grow to 16 000 five years' later. Assuming that the error distributions shown in Table 5 are applicable, we could say it is probable that absolute Percentage Error will not exceed 4.1 per cent and very likely (though not impossible) lie within 5.9 per cent. Therefore it is probable that the population in five years' time will be 16,000 ± 656 (i.e. 0.041 × 16 000) and unlikely it will exceed 16 000 ± 944 (i.e. 0.059 × 16 000). This scale of uncertainty is probably greater than most users realise (or would like), but it reflects the reality of local area population forecasting.

Effectively these calculations comprise a simple method of estimating empirical prediction intervals around a population projection (Tayman, 2011; Wilson, 2012). Further research is needed to create a more refined regression-based approach which uses a larger sample of past projections and accounts for factors such as recent growth rates, mining employment, majority Indigenous populations, the varying volatility of migration trends, and urban localities slated for rapid development or redevelopment. Nonetheless, even basic indications of uncertainty are better than none, and certainly preferable to traditional high and low projection variants which have been shown to be poor at representing likely error ranges (e.g. Keilman *et al*., 2001; Wilson and Bell, 2007). Information on the likely range of future population should prove useful to decision-makers facing significant investment decisions. For example, will the population of town X have grown sufficiently in five years' time to justify the building of a new supermarket? Different decisions might result from projections which have a relatively small prediction interval compared to those with a very large range of uncertainty which indicate either growth or decline is possible.

The Scope for Improving Accuracy

Providing information on the uncertainty of projections is important, but it would also be beneficial to undertake research aimed at improving the accuracy of projections. The local area projections assessed in this paper have been shown to be quite accurate overall. Is there really much scope for reducing errors in the future? This author's view is that there is potential to reduce errors a little. A number of suggestions are made.

First, more robust ERPs are required. The recent ERP recasting has shed light on the issue of uncertainty in official population estimates. Good population projections require a solid foundation of past population trends. In addition, reliable and consistent local area births, deaths, internal migration and overseas migration data are required to understand how populations are changing. Population accounts should be free of any 'unexplained growth': the ERP at 30th June in year t plus the subsequent 12 months of births, minus deaths, plus net internal migration, plus net overseas migration should equal the ERP one year later.

Second, regular assessments of past projections are worthwhile. Diagnosing problems is half-way to providing solutions. Where possible it is useful to assess the accuracy of projected demographic components of change (births, deaths and migration) as well as projected populations. Are there particular demographic components or types of areas that are often poorly forecast? In many cases there may be no obvious answers as to what went wrong, but in others there could be. It may be wise to experiment with different types of projection models for such areas, or averages of several different models which bring different strengths to the overall projection.

Better projections of migration are also crucial to reducing error, as Figure 2 suggests. Improving accuracy will always be hard in a western liberal democracy with freedom of internal movement and with much of the country's overseas migration not subject to migration controls (e.g. the immigration of Australian and New Zealand citizens, and all emigration flows). Experimenting with different methods of projecting migration, and investigating alternative data sources, may prove helpful. For local areas within metropolitan regions, high-quality data on the region-wide distribution of anticipated residential development or re-development are helpful. For non-metropolitan areas, recent analysis by Argent (2014) found a strong correlation between building approvals and net migration volumes in non-metropolitan areas in NSW, suggesting that dwelling data may also be valuable for projections in these areas.

Finally, it has been demonstrated that LGAs with the smallest populations tend to have the most inaccurate population projections. The merging of LGAs to avoid any with populations under 5 000 would reduce the likelihood of obtaining errors exceeding 10 per cent after five years, or worse, greater than 20 per cent (classified here as poor or bad forecasts). This would, of course, be a statistical artefact rather than a genuine improvement in projection accuracy, but it would nonetheless make the populations of LGAs at the lower end of the population size spectrum slightly less small and a little more forecastable.

5. SUMMARY AND CONCLUSION

This paper has examined the short-term forecast error of the 2006 round local government area population projections produced by five states. The key findings can be summarised as follows:

- Average LGA total population forecast errors were low – low relative to naïve linear extrapolation and the 10 per cent error cut-off deemed acceptable in Tye (1994). Across the five states studied the average five year forecast error as measured by WMAPE was just 2.4 per cent.

- Across all LGAs in the five states, 95 per cent were forecast within 10 per cent absolute Percentage Error (with 83 per cent forecast within 5 per cent).

- A small proportion of LGAs (5 per cent) experienced large errors (>10 per cent after 5 years) in their population projections. These were mostly very small and/or Indigenous, or declining rapidly in the recent past, characteristics known to adversely affect forecast accuracy. Projections for LGAs with fewer than about 5 000 people were the most inaccurate.

- Age groups with the highest average errors were the young adult ages and the very elderly. The childhood age groups had low errors, on average.

In presenting these findings the paper has contributed evidence to a very small literature in Australia and internationally on subnational population forecast accuracy. Future research by the author will concentrate on generating a more comprehensive model of empirical prediction intervals, and experimenting with ways of reducing error further. A particular focus of this research should be on those LGAs prone to the largest errors, especially with regards to the accuracy of dwelling forecasts for areas expected to grow rapidly, and the economic prospects of small areas heavily dependent on individual industry sectors.

In the meantime, users of the states' local area population projections can generally be fairly confident about the short-term reliability of the figures. However, users should exercise particular caution with projections for very small populations (under 5 000 people) and those with rapidly declining populations in the recent past, and accept that projections for the young adult ages (20-34) are not as accurate as those for other age groups. In addition, users wishing to obtain a rough indication of the uncertainty of current sets of projections can make use of the error distributions of recent projections (Table 5).

REFERENCES

ABS (2008). *Experimental Estimates of Aboriginal and Torres Strait Islander Australians, Jun 2006*. [Excel file] Cat. No. 3238.0.55.001. ABS, Canberra.

ABS (2012). *2011 Census of Population and Housing. Murrumbidgee Time Series Profile*. [Excel file] Cat. No. 2003.0. Canberra: ABS.

ABS (2013a). *Regional Population Growth, Australia, 2012*. Cat. No. 3218.0. ABS, Canberra.

ABS (2013b). *Australian Demographic Statistics*. 2011 Census Edition – Final. December quarter 2012. Cat. No. 3101.0. ABS, Canberra.

ABS (2013c). *Estimates of Aboriginal and Torres Strait Islander Australians, June 2011*. [Excel file] Cat. No. 3238.0.55.001. ABS, Canberra.

ACT Government (2013). *ACT Public School Enrolment Projections 2013 to 2017.* ACT Government, Canberra.

AEC (Australian Electoral Commission) (2015). Enrolment projections overview – NSW federal redistribution. Online version accessed 14 July 2015, http://www.aec.gov.au/electorates/Redistributions/2014/nsw/enrolment-projections-overview.htm.

Argent, N. (2014). *The Movement of People in Regional New South Wales, 2006-2011*. Report to the NSW Department of Planning and Environment. University of New England, Armidale.

Bell, M., Wilson, T. and Charles-Edwards, E. (2011). Australia's Population Future: Probabilistic Forecasts Incorporating Expert Judgement. *Geographical Research*, 49(3), pp. 261-275.

Carey, K. D. (2011). *School District Master Planning*. Rowman and Littlefield, Lanham.

Diamond, I., Tesfaghiorghis, H. and Joshi, H. (1990). The Uses and Users of Population Projections in Australia. *Journal of the Australian Population Association*, 7(2), pp. 151-170.

Dovey, C. (2015). Schoolyard Crush. *The Monthly*, February 2015.

Johnstone, K. (forthcoming, 2015). Communicating Population Projections to Stakeholders: a Case Study from New South Wales. In Wilson, T., Bell. M., and Charles-Edwards, E. (Eds) *Demography for Planning and Policy: Australian Case Studies*. Springer.

Keilman, N. (1999). How Accurate are the United Nations World Population Projections? In W. Lutz, J.W. Vaupel and D. A. Ahlburg (Eds) *Frontiers of Population Forecasting*, pp. 15-41, Supplement to Population and Development Review vol. 24, Population Council, New York.

Keilman, N., Pham, D.Q. and Hetland, A. (2001). *Norway's Uncertain Demographic Future*. Oslo: Statistics Norway.

Murdock, S.H., Leistritz, F.L., Hamm, R.R., Hwang, S. and Parpia, B. (1984). An Assessment of the Accuracy of a Regional Economic-Demographic Projection Model. *Demography*, 21(3), pp. 383-404.

NSW Department of Planning and Environment (2014). 2014 NSW Population Projections data [Excel file]. Sydney: Department of Planning and Environment. Online version accessed 21 January 2015, http://www.planning.nsw.gov.au/en-au/deliveringhomes/populationandhouseholdprojections/data.aspx.

NSW Department of Planning (2009). *New South Wales Statistical Local Area Population Projections, 2006-2036*. Department of Planning, Sydney.

Queensland Department of Infrastructure and Planning (2008). *Queensland's Future Population: 2008 edition*. Department of Infrastructure and Planning, Brisbane.

Queensland Government Statistician's Office (2013). Projected Population (Medium Series), by Five-Year Age Group (Males, Females, and Persons), by Local Government Area, Queensland, 2011 to 2036 [Excel file]. Queensland Treasury and Trade, Brisbane. Online version accessed 21 January, http://www.qgso.qld.gov.au/products/tables/proj-pop-medium-series-age-group-sex-lga-qld/index.php.

Rayer, S. (2008). Population Forecast Errors: a Primer for Planners. *Journal of Planning Education and Research*, 27, pp. 417-430.

Rayer, S. and Smith, S.K. (2014). Population Projections by Age for Florida and its Counties: Assessing Accuracy and the Impact of Adjustments. *Population Research and Policy Review*, 33(5), pp. 747-770.

Salt, B. (2014). Population Forecasts Key to Good Planning. *The Australian*, 11 September 2014, Property page 25.

Siegel, J.S. (2002). *Applied Demography: Applications to Business, Government, Law and Public Policy*. Academic Press, San Diego.

Smith, S. and Sincich, T. (1988). Stability Over Time in the Distribution of Population Forecast Errors. *Demography*, 25, pp. 461-474.

Smith, S.K. (1987). Tests of Forecast Accuracy and Bias for County Population Projections. *Journal of the American Statistical Association*, 82, pp. 991-1003.

Smith, S.K. and Rayer, S. (2011). *An Evaluation of Population Forecast Errors for Florida and Its Counties, 1980–2010*. Bureau of Economic and Business Research, University of Florida.

South Australian Department of Planning and Local Government (2011). Age-Sex Population Projections by Statistical Local Area, 2006-2026. [Excel file]. Department of Planning and Local Government, Adelaide. Online version accessed 7 February 2012, www.sa.gov.au/planning/populationprojections.

Statistics New Zealand (2008). *How Accurate are Population Projections? An evaluation of Statistics New Zealand population projections, 1991–2006*. Statistics New Zealand, Wellington.

Tasmanian Demographic Change Advisory Council (2008). Population Projections. [Excel file].Demographic Change Advisory Council, Hobart. Online version accessed 11 October 2009, www.dcac.tas.go.au/.

Taylor, A. (forthcoming, 2015). Not Just Ageing: Policy and Service Delivery Implications from Changing Population Compositions in the Northern Territory of Australia. In T. Wilson, M. Bell and E. Charles-Edwards (eds) *Demography for Planning and Policy: Australian Case Studies*. Springer.

Tayman, J. (2011). Assessing Uncertainty in Small Area Forecasts: State of the Practice and Implementation Strategy. *Population Research and Policy Review*, 30, pp. 781-800.

Tayman, J. and Swanson, D.A. (1999). On the Validity of MAPE as a Measure of Population Forecast Accuracy. *Population Research and Policy Review*, 18, pp. 299-322.

Tayman, J., Smith, S.K. and Rayer, S. (2011). Evaluating Population Forecast Accuracy: a Regression Approach Using County Data. *Population Research and Policy Review*, 30(2), pp. 235-262.

Tye, R. (1994). Local Authority Estimates and Projections: How are They Used and How Accurate Do They Need to Be? Working Paper 9, Estimating with Confidence project. Department of Social Statistics, University of Southampton, UK.

Victoria Department of Transport, Planning and Local Infrastructure (2014). *Victoria in Future 2014*. Department of Transport, Planning and Local Infrastructure, Melbourne.

Victorian Department of Planning and Community Development (2009). Victoria in Future 2008 - second release. [Excel file]. Department of Planning and Community Development, Melbourne. Online version accessed on 11 October 2009, www.dpcd.vic.gov.au/victoriainfuture.

Wilson, T. (2012). Forecast Accuracy and Uncertainty of Australian Bureau of Statistics state and territory population projections. *International Journal of Population Research*, Volume 2012, DOI: 10.1155/2012/419824, http://dx.doi.org/10.1155/2012/419824.

Wilson, T. (2013). Quantifying the Uncertainty of Regional Demographic Forecasts. *Applied Geography*, 42, pp. 108-115.

Wilson, T. (2014). New Evaluations of Simple Models for Small Area Population Forecasts. *Population, Space and Place,* 21, pp. 335-353. DOI: 10.1002/psp.1847. http://dx.doi.org/10.1002/psp.1847.

Wilson, T. and Bell, M. (2007). Probabilistic Regional Population Forecasts: the Example of Queensland, Australia. *Geographical Analysis*, 39(1), pp. 1-25.

Wilson, T. and Rowe, F. (2011). The Forecast Accuracy of Local Government Area Population Projections: a Case Study of Queensland. *Australasian Journal of Regional Studies*, 17(2), pp. 204-243.