

## **THE FORECAST ACCURACY OF LOCAL GOVERNMENT AREA POPULATION PROJECTIONS: A CASE STUDY OF QUEENSLAND**

**Tom Wilson**

Queensland Centre for Population Research, School of Geography, Planning and Environmental Management, Chamberlain Building, The University of Queensland, Brisbane, Qld, 4072, Australia. Email: [tom.wilson@uq.edu.au](mailto:tom.wilson@uq.edu.au)

**Francisco Rowe**

Queensland Centre for Population Research, School of Geography, Planning and Environmental Management, Chamberlain Building, The University of Queensland, Brisbane, Qld, 4072, Australia

**ABSTRACT:** Users of population projections tend to assume that they provide accurate predictions of future demographic trends. However, there is limited evidence to either support or refute this assumption because almost no research has evaluated the accuracy of past projections in Australia. This paper assesses the forecast accuracy of seven previous rounds of local government area population projections for Queensland. The analysis reveals errors to be quite large in absolute terms and in relation to State forecast errors, but respectable compared to those reported in other studies. Relative to simple extrapolative forecasts, the official Queensland projections are shown to have performed quite well. Fractional response models are employed to determine the extent to which forecast errors can be predicted on the basis of local area characteristics. The concluding section suggests ways in which forecast error might be reduced and how users can be informed about the possible magnitude of error in current projections.

Key words:

population projections; forecast accuracy; local government areas; Queensland; fractional response model

### **1. INTRODUCTION**

Local government area (LGA) population projections are produced regularly by State and local governments. They inform planning for education, health, transport, power, water and sewerage services, housing provision, road construction and retail facilities to name few examples.

Major decisions and investments are made on their basis. They also generate considerable controversy and criticism from councils when projected population change is at odds with a council's development stance. Decisions and reactions are often made on the assumption that the projections are highly likely to eventuate (Swanson and Tayman 1995). In fact, as this paper will demonstrate, that assumption is not supported by evidence.

There are at least two major reasons for studying the errors of past population projections. First, they may highlight problems with certain aspects of the forecasting methodology or data. For example, are there certain types of LGA which are always over- or under-predicted? Which LGAs' populations have been subject to the largest forecast errors, and why? Such an analysis may reveal persistent problems which can be remedied in future sets of projections. Second, past forecast errors can provide users with estimates of uncertainty in the current set of forecasts. As Keyfitz (1981 p 579) argued, "Demographers can no more be held responsible for inaccuracy in population forecasting 20 years ahead than geologists, meteorologists or economists when they fail to announce earthquakes, cold winters, or depressions 20 years ahead. What we can be held responsible for is warning one another and our public what the error of our estimates is likely to be." Whilst future errors are not guaranteed to resemble those of the past, past errors can at least provide a ballpark guide to uncertainty where there is reasonable stability over time in error distributions (Smith and Sincich 1988). There are various ways in which past errors may be used to represent uncertainty. They range from simple tabular presentations of past forecast errors, through more complex approaches where models are fitted to past forecast error distributions (Tayman, Schafer and Carter 1998), to sophisticated probabilistic population forecasts in which past errors are used calibrate statistical models to represent the variability of demographic processes (Keilman, Pham and Hetland 2002; Wilson and Bell 2007).

There is a small but growing literature on the accuracy of past population forecasts. Much of it focuses on national and global populations; for example Keilman (2001, 2008), Keilman and Pham (2004), Keyfitz (1981), Khan and Lutz (2008), Mulder (2002), National Research Council (2000), Shaw (2007) and Wilson (2007). The literature on subnational population forecast error is much more limited. Examples include Bell and Skinner (1992), Campbell (2002), Murdock et al. (1984), Office for National Statistics (2008), Smith and Tayman (2003), Statistics New Zealand (2008), Tayman (1996), and Tayman, Schafer and

Carter (1998). As a general rule, studies at both national and subnational scales have found that population forecasts tend to be more accurate for shorter rather than longer forecast horizons and more populous rather than less populous regions. Few analyses of subnational forecast errors have been undertaken for Australia. Although some studies have assessed past State population projections (Adam 1992, Bell and Skinner 1992, ABS 2000) there is only one published analysis of sub-State forecast errors we are aware of, the Western Australian Planning Commission's (2004) study of the short-term (five year) accuracy of its State, regional and LGA projections.

A few forecast accuracy studies have additionally compared the errors of past projections with those of retrospective projections produced using simple, data-light, extrapolative methods (sometimes referred to as naive methods). Examples of these comparative studies include Campbell (2002), Long (1995), Rayer (2008), Smith and Sincich (1992), Smith and Tayman (2003) and Tayman, Smith and Rayer (2010). Often a great deal of resources (and therefore expense) is devoted to the preparation of a set of local area projections. Whilst many demographers may feel uncomfortable with the atheoretical nature of extrapolative methods and their ignorance of age structure effects, if they consistently perform better than cohort-component models then a case can be made for their use. Some studies comparing simple methods with cohort-component models have found that for projecting total populations there is no great difference in overall accuracy between the two types of model (e.g. Smith and Sincich 1992, Smith and Tayman 2003). Of course, simple methods only produce projections of total populations, so where age and sex detail is required then cohort-component models must be used. But even in these situations it may still be worth using simple methods to produce population totals to act as constraints on age-disaggregated projections.

There is a second reason for producing retrospective extrapolative projections. They provide a benchmark which can be used to assess how good a 'real' set of projections was. Simply comparing two sets of past projections and declaring the one with the lowest mean error the better projection overlooks the varying difficulty of producing them. Projections with the lowest mean error may have been produced at a time of little demographic change, with demographic trends fortuitously turning out to follow the projections quite closely. The value of a set of past projections may be measured by comparing past projections to projections calculated using a naïve extrapolative method (Swanson and Tayman 1995). The gain in accuracy from using the actual projection rather than the naïve one represents the value added, or 'proportionate reduction in error'. For

example, a comparison of two sets of projections could result in the one with the higher mean error values being declared the 'better' projection because of its greater proportionate reduction in error in relation to a naive forecast.

A great deal remains to be learned about the nature and extent of local area population forecast accuracy internationally, but especially in Australia. Do the findings of mostly North American forecast accuracy case studies also apply elsewhere? This paper makes an initial contribution to learning more, aiming to both improve understanding of Australia-specific projections issues and add to the international body of evidence on local area population forecast accuracy. It makes a novel contribution to the literature through the use of a fractional response regression model, a more statistically appropriate model than the commonly employed linear least squares method. It also promotes the little-used error measure Mean Percentage Absolute Deviation (MPAD). The paper asks:

- (1) How accurate have past projections of total populations for Queensland LGAs been? Specifically, how do errors vary by projection horizon, population size and growth rate?
- (2) For which LGAs have the projections proved to be very poor population forecasts? Why have these projections been poor?
- (3) How do forecast errors compare to those of a simple linear extrapolation model?
- (4) How do forecast errors compare to those reported by other studies?
- (5) To what extent can forecast errors be predicted on the basis of local area characteristics?

The following terms are used in the paper. *Base period* refers to the period of past data which a projection model uses as its inputs. The *launch year* is the final year of the base period and the population estimate from which the forecast is 'launched'. *Projection period* refers to the whole time period between the launch year and the *final year* of the projections, whilst *projection horizon* describes the interval between the launch year and any particular projection year in question. This paper

analyses projections for projection horizons of 5, 10 and 15 years. Distinctions are often made between a *projection* and a *forecast*, with the former referring to a numerical statement about future population based on certain assumptions, and the latter describing a particular projection which is believed to be the most likely demographic future. Although many demographers stress that they produce projections and not forecasts, their middle series projections are nearly always interpreted as forecasts by users. We follow Smith (1987) by labelling the data under analysis as ‘projections’, but evaluating them as if they were forecasts and thus refer to ‘forecast accuracy’.

The plan of the paper is as follows. The next section describes the projections data and the population estimates which they were compared against. Section 3 lists the various error measures used in the study, whilst the following results section presents answers to the five questions posed above. The final section includes a number of recommendations on how forecast accuracy might be improved, and how forecast uncertainty could be communicated to users.

## 2. DATA

### *Population Projections Data*

Seven rounds of official Queensland LGA population projections were evaluated in this study, beginning with the earliest set published in 1989 (Skinner, Bell and Gillam 1989) extending to the 2003 edition projections (Queensland Department of Local Government and Planning 2003). Table 1 presents summary information about the seven sets of projections. Data on total populations were obtained either from hard copy publications or from spreadsheets supplied by the Demography and Planning section of Queensland Treasury. We analysed total populations only because age-specific projections data were not available for all projection rounds.

The LGAs under consideration were those in existence in 2006, with boundaries as defined in the 2006 Australian Standard Geographical Classification. Following the 2008 local government reorganisation in Queensland many of these LGAs experienced boundary changes or ceased to exist, but for the purposes of assessing the characteristics of past projections and the lessons they may hold for future projections, this is unimportant.

**Table 1.** Summary Characteristics of Evaluated Queensland LGA Population Projections. Source: the Authors

Label	Launch year	Final year	Series	Method for projecting LGA total populations	Reference
1989	1986	2001	Medium	Cohort-component model using net migration rates (SAASSPS)	Skinner, Bell and Gillam (1989)
1992	1986§	2006	Medium	As above	Queensland Department of Housing and Local Government (1992)
1994	1991	2011	Medium	As above	Queensland Department of Housing, Local Government and Planning (1994)
1996	1995	2011	Medium	Population totals projected as average of a share of population model and a share of growth model. Two sets of adjustments were then made (see text)	Queensland Department of Local Government and Planning incorporating Rural Communities (1996)
1998	1996	2016	Medium	As above	Queensland Department of Communication and Information, Local Government and Planning (1998)
2001	2000	2021	Medium	As above	Queensland Department of Local Government and Planning (2001)
2003	2001	2026	Medium	Population totals projected as average of a share of population model and a share of growth model; combined with a dwelling-led model for LGAs in south east Queensland (see text)	Queensland Department of Local Government and Planning (2003)

§ Although 1991 ERPs based on the 1991 Census were not available at the time the projections were prepared, “for the period 1986-90 the assumptions have been heavily influenced by ABS intercensal estimates of overseas and interstate migration to Queensland and by ABS population estimates for each statistical division” (Queensland Department of Housing and Local Government 1992 p.9).

The following approach was taken where boundary changes occurred between the launch year of each set of projections and 2006. If the impact of boundary changes on population numbers was small, defined as the launch year population in the projections being within 1% of the 2006 ASGC population estimate for that year, then the LGA was retained in

the analysis. If not, an attempt was made to merge it with a neighbour to meet the 1% population difference criterion. Failing this, the LGA was discarded from the analysis. Merging was in fact only applied to three small Indigenous councils created in 2002 which were merged back with the local government areas from which they were carved. The projections dataset used in the evaluation thus comprised 90 LGAs (or areas consisting of merged LGAs) totalling 83% of Queensland's population in 2006. These LGAs are listed in the Appendix.

Projection rounds 1989, 1992 and 1994 were prepared using the projection package SAASPPS (Small Area Age-Sex Population Projection System) developed by the University of Queensland (APRU 1992). SAASPPS is a cohort-component model of the population disaggregated by sex and five year age groups up to 75 years and over. Migration is handled with age-sex-specific net migration rates. Total populations were obtained by summing projections by age and sex after constraining to separately-produced regional projections. Preliminary projections were often adjusted in an iterative process on the basis of feedback from councils and assessments of the face validity of the preliminary projected demographic components of change.

From the 1996 projections onwards a new approach was implemented in which LGA population totals were calculated using the average of two extrapolative methods. One of these extrapolates each LGA's share of its encompassing regional population (which is projected independently); the other method extrapolates each LGA's share of the region's recent population *growth*. Adjustments were then made "in cases where the model produced future population growth considered unable to be absorbed by individual LGAs or the model produced accelerating population declines which were also considered unlikely in practice" (Queensland Department of Local Government and Planning, 2001 p 155). At this stage the projections were circulated to councils for comment and further adjustments made on the basis of council feedback where necessary. For the 2003 projections the extrapolative approach was applied as before, but additionally, population totals for LGAs within the densely populated region of south east Queensland were projected using a housing-unit model. Projections from the housing unit model were combined with those from the extrapolative approach to obtain the final projections (though details of exactly how the two sets of figures were combined are not given in the projections publication).

*Population Estimates Data*

Projections were assessed against the best estimate of the usually resident population of each LGA. Estimated Resident Population (ERP) data for 1981, 1986, 1991, and every year from 1996 to 2006 on the 2006 ASGC were supplied by the Demography and Planning section of Queensland Treasury (based on original data from the Australian Bureau of Statistics (ABS)). ERPs for 1976, required as part of the base period data for the simple extrapolations, were obtained from a population estimates release published by ABS (1984). Forecast accuracy has not been assessed beyond 2006 because post-2006 ERPs are not finalised and will be revised in the light of 2011 Census results.

*Extrapolative Projections Data*

Retrospective projections of LGA population totals for the same launch years as the official projections were produced by linear extrapolation. This model was chosen from the many available extrapolative methods because it is very simple to apply and has often been found to be amongst the most accurate of extrapolative methods (e.g. Openshaw and van der Knapp 1983, Smith and Shahidullah 1995, Smith and Sincich 1992). The model was applied by fitting a linear regression to the three data points which formed the base period dataset (ERPs for the launch year, five years earlier and ten years earlier). Constraining to State projections was not applied. A ten year base period represents the usual length of time series supplied by ABS on a current set of LGA boundaries. Interestingly, both Smith and Sincich (1990) and Rayer and Smith (2010) have found that, for simple methods, increasing the base period beyond ten years generally yields little improvement in forecast accuracy.

*Regression Model Data*

To answer the question ‘To what extent can forecast errors be predicted on the basis of local area characteristics?’ we estimated a regression model with error regressed against several characteristics of each LGA (details of which are presented in section 4.5). The choice of predictor variables was informed by the existing literature (Tayman et al. 2010) as well as our own experimentation. The first variable, *POP*, is the natural logarithm of the launch year Estimated Resident Population, obtained



from ABS. The proportion of the population identifying as Indigenous at the preceding census, *IND*, was calculated from ABS census data, as was the proportion of the workforce employed in the mining industry at the preceding census, *MIN*. The dependent variable was also regressed against errors of a previous projection after a 5 year horizon, *PRV*. This previous projection is the most recent projection which could have been evaluated after 5 years at the time of publication of each projection round.

### 3. MEASURES OF FORECAST ERROR

#### *Measures of Average Error*

Three measures of average forecast error were used in this study: the Mean Absolute Percentage Error (MAPE), the Median Absolute Percentage Error (MedAPE) and the Mean Percentage Absolute Deviation (MPAD). The first two are based on the Percentage Error (*PE*) of individual projections, defined for any area *i* as

$$PE^i = \frac{F^i - A^i}{A^i} 100$$

where *F* denotes the forecast population and *A* the actual value (the subsequently published Estimated Resident Population). A positive value indicates a projection which is too high; a negative value signifies an under-prediction.

MAPE was selected because it is widely used and easily interpretable. It is calculated by summing the absolute values of Percentage Errors for all *n* observations and then dividing by *n*:

$$MAPE = \frac{\sum_i \left| \frac{F^i - A^i}{A^i} 100 \right|}{n}$$

Despite its widespread use, some authors have criticised the use of MAPE in population forecast accuracy studies because the distribution of absolute error values is frequently right-skewed, thus 'overstating' error (Tayman, Swanson & Barr 1999; Tayman & Swanson 1998). MedAPE, the middle value of the ranked set of Absolute Percentage Errors, was

therefore selected as a complementary measure because it is unaffected by outliers.

The third measure, the Mean Percentage Absolute Deviation (MPAD), provides a rather different perspective (Murdock et al. 1984). It is defined as the mean absolute error as a percentage of the mean actual population. It can be calculated as:

$$MPAD = \frac{\sum_i |F^i - A^i| / n}{\sum_i A^i / n} 100 \quad \text{or simply} \quad \frac{\sum_i |F^i - A^i|}{\sum_i A^i} 100.$$

MPAD is also known as the MAD/mean ratio (Mean Absolute Deviation/mean) (Kolassa and Schutz 2007) and the population-weighted MAPE (WMAPE) (Siegel 2002), i.e.

$$\sum_i \left( \left| \frac{F^i - A^i}{A^i} \right| 100 \right) \frac{A^i}{\sum_i A^i}$$

where the population weight is denoted  $\frac{A^i}{\sum_i A^i}$ . Whereas MAPE gives each observation equal weighting in the calculation of average error, MPAD weights the APE for each observation by its population size. For users who regard a 10% error in projecting a population of 1,000,000 as far more costly than the same percentage error for a population of 1,000, MPAD is a useful measure. It is particularly suitable for this case study where population sizes vary enormously between LGAs.

Note that in this paper we do not focus much on bias – whether projections were too high or too low overall – because for the purposes of using past errors to warn users about the likely error of current forecasts, it is the absolute value of error that is of greater value. Previous research has demonstrated absolute errors to be more predictable than bias (Tayman et al. 2010). In addition, an analysis of bias would have resulted in a very long paper.

#### *Measures of Error Distribution*

Forecast error distributions were measured simply by classifying LGA projection errors as:

- (i) small – where Absolute Percentage Errors (APEs) are within 10%,
- (ii) moderate – where errors extend from 10 to 20% APE, or
- (iii) large – where APEs exceed 20%.

The 10 and 20% category boundaries were chosen on the basis of workshop discussions involving UK local area population estimates and projection users reported by Tye (1994). Errors of 5% were viewed as unproblematic, those up to 10% were generally considered acceptable, whilst errors of 20% or more were universally regarded as very serious. These error categories were determined from a user's perspective in which a very large error – for example, 30% – is a serious error irrespective of whether it occurs 5 or 15 years into the forecast.<sup>1</sup>

#### *Comparative Error Measures*

Two comparative error measures were used: Proportionate Reduction in Error (PRE) and Percentage Better. PRE measures the percentage reduction in error, or gain in accuracy, by using a 'real' projection over a simple method (Swanson and Tayman 1995; Tayman and Swanson 1996). It was applied here to MAPE, MedAPE and MPAD in the form:

$$PRE = \frac{\text{Average error of linear extrapolation} - \text{Average error of real projection}}{\text{Average error of real projection}} 100$$

where *PRE* denotes the Proportionate Reduction in Error. Positive values signify the percentage reduction in error from using the real projection method, whilst negative values indicate superior results from the linear extrapolation.

Percentage Better simply expresses the percentage of observations forecast more accurately by method A than method B (Armstrong 2001). It was applied in this study to describe the percentage of LGAs whose populations were forecast more accurately by the real projections rather than by linear extrapolation.

---

<sup>1</sup> From a forecast producer's perspective an alternative categorisation scheme could have classified forecasts as good, satisfactory and poor with the Absolute Percentage Errors for each category increasing over time to reflect the greater difficulty of forecasting further into the future.

#### 4. RESULTS

##### 4.1. HOW ACCURATE HAVE PAST PROJECTIONS OF TOTAL POPULATIONS FOR QUEENSLAND LGAs BEEN?

###### *Error by Projection Horizon*

Table 2 contains average measures of error for combinations of projection rounds at projection horizons of 5, 10 and 15 years. With the most recently available final ERP data being for 2006, only three projection rounds could be assessed after 15 years (1989, 1992 and 1994). The 1996 and 1998 rounds could be evaluated for up to 10 years and the 2001 and 2003 rounds could be assessed at only a five year horizon. The table shows how average forecast error is positively associated with forecast horizon, a finding which is consistent across the three average error measures, and in line with many other studies. Average State errors are smaller than average LGA errors, which also corresponds with existing evidence. The values of MAPE exceed MedAPE due to right-skewness in the distribution of Absolute Percentage Errors. MPAD displays the smallest numbers because many LGAs in Queensland have very small populations. Recalling that MPAD is effectively a population-weighted MAPE, the Absolute Percentage Errors of these small LGAs receive little weight in the calculation of MPAD.

**Table 2:** Average Error in Queensland LGA and State Projections by Projection Horizon. Source: the Authors

Projection horizon (years)	LGAs			State
	MAPE	MedAPE	MPAD	MAPE
<i>2001 &amp; 2003 projection rounds</i>				
5	6.3	4.3	2.9	2.2
<i>1996 &amp; 1998 projection rounds</i>				
5	7.1	5.4	4.5	1.4
10	8.8	7.0	7.5	2.0
<i>1989, 1992 &amp; 1994 projection rounds</i>				
5	6.5	4.4	3.7	0.7
10	10.9	7.9	8.2	2.4
15	14.6	11.3	12.2	1.9

Whilst producers of projections are keen to minimise average errors, they usually also wish to maximise the number of areas with small errors and avoid, as much as possible, embarrassingly large errors. To what extent was this achieved with past Queensland LGA projections? Table 3 provides answers by presenting the percentage of LGAs whose populations were projected with small, moderate and large errors, as defined in section 3. Because error tends to cumulate over time error distributions become wider with longer projection horizons. After 5 years approximately 80% of LGAs were within 10% error, falling to 60-70% after 10 years and a little under half after 15 years. The percentage of LGAs with large errors was only about 4-5% five years into a projection, but unfortunately rose substantially with projection horizon length.

**Table 3.** Percentage of Queensland LGAs by Error Category and Projection Horizon. Source: the Authors.

Projection horizon (years)	Small errors <10% APE	Moderate errors 10-20% APE	Large errors >20% APE
<i>2001 &amp; 2003 projection rounds</i>			
5	79.4	16.7	3.9
<i>1996 &amp; 1998 projection rounds</i>			
5	78.9	16.1	5.0
10	70.0	20.6	9.4
<i>1989, 1992 &amp; 1994 projection rounds</i>			
5	80.7	15.2	4.1
10	58.1	25.6	16.3
15	45.2	29.6	25.2

#### *Error by Population Size*

How has error varied by launch year population size? Some previous studies have revealed a negative relationship between error and population size. Does this also apply to Queensland? Five categories of LGA are distinguished: those with populations under 2,000 (23 LGAs in 2006), 2,000-4,999 (20 LGAs), 5,000-14,999 (22 LGAs), 15,000-49,999 (10 LGAs) and 50,000 and above (15 LGAs). MAPE and MPAD are the chosen measures for this question because MedAPE is less stable with

small sample sizes (Rayer 2007). Table 4 presents the results. As can be seen, both error measures reveal a general trend of average error declining with increasing population size, albeit at a diminishing rate. Some of the more populous LGAs proved exceptions to this rule, however<sup>2</sup>.

**Table 4.** Average Error of Queensland LGA Projections by Launch Year Population Size and Projection Horizon. Source: the Authors.

Projection horizon (years)	Launch year population size				
	0-1,999	2,000-4,999	5,000-14,999	15,000-49,999	50,000+
<i>MAPE</i>					
<i>2001 &amp; 2003 projection rounds</i>					
5	10.0	6.9	5.4	3.3	2.8
<i>1996 &amp; 1998 projection rounds</i>					
5	9.2	7.3	7.1	5.2	4.0
10	12.9	8.7	7.2	5.7	6.9
<i>1989, 1992 &amp; 1994 projection rounds</i>					
5	11.2	5.9	5.2	4.5	3.4
10	13.4	11.4	10.1	10.0	7.1
15	19.9	13.7	15.2	9.7	10.9
<i>MPAD</i>					
<i>2001 &amp; 2003 projection rounds</i>					
5	9.2	6.8	5.3	3.2	2.5
<i>1996 &amp; 1998 projection rounds</i>					
5	7.9	7.3	6.6	5.0	4.0
10	11.3	8.2	6.7	5.1	8.1
<i>1989, 1992 &amp; 1994 projection rounds</i>					
5	9.9	6.0	5.0	4.3	3.3
10	14.4	11.3	9.5	9.6	7.6
15	20.2	13.2	13.9	8.2	12.7

<sup>2</sup> A referee suggested we check whether a power function could be employed to predict Absolute Percentage Error on the basis of launch year population size. Unfortunately the association between the two variables is very weak and the R<sup>2</sup> values for five, 10 and 15 year projection horizons were all very low.

*Error by Growth Rate*

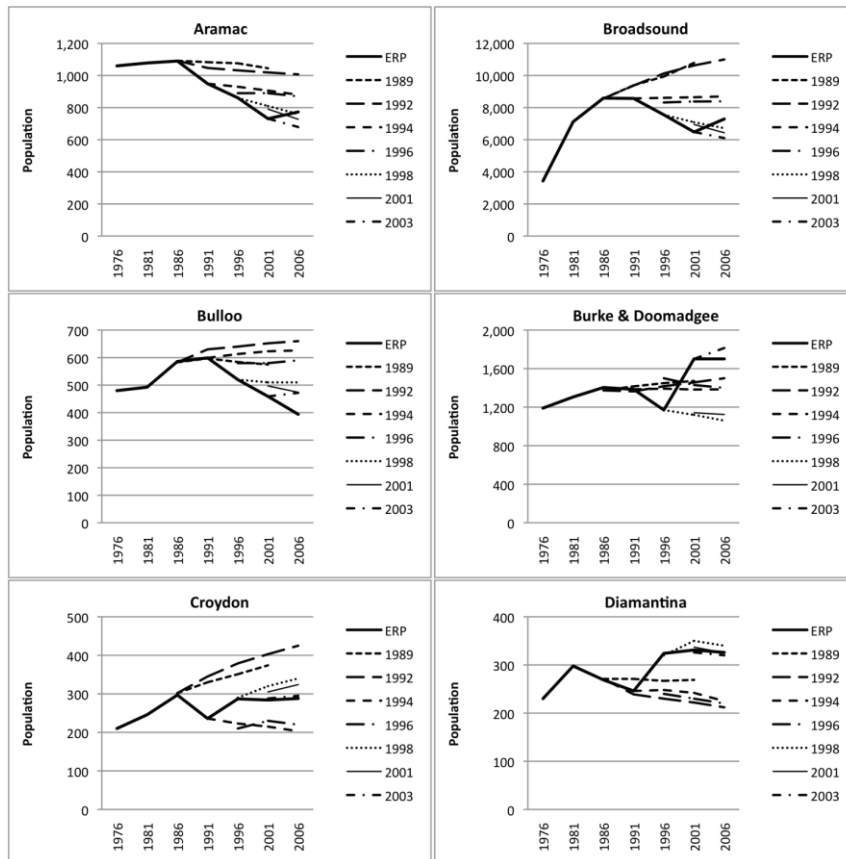
Some studies have suggested that an approximately u-shaped relationship exists between population growth rates in the base period and subsequent forecast error, with moderately growing areas tending to be forecast more accurately than those growing or declining rapidly (e.g. Smith 1987; Rayer and Smith 2010). Table 5 indicates that the u-shaped relationship is not evident in the case of Queensland LGA projections. In fact there appears to be little pattern to the errors. One possibility is that the u-shaped relationship is more likely to result from simple extrapolative models because areas growing or declining rapidly in the base period often experience subsequent moderation in their rates of growth. Extrapolative models, by their nature, maintain base period trends and will always give highly erroneous forecasts when the growth rate changes, whereas real projections – which incorporate human judgement – may well include growth rate alterations. Alternatively, it is possible that the relatively small sample size of 90 LGAs may be obscuring the underlying pattern, and had a greater number of LGAs been analysed a clearer pattern might have emerged.

**Table 5.** Average Error of Queensland LGA Projections by Base Period Growth Rate and Projection Horizon. Source: the Authors.

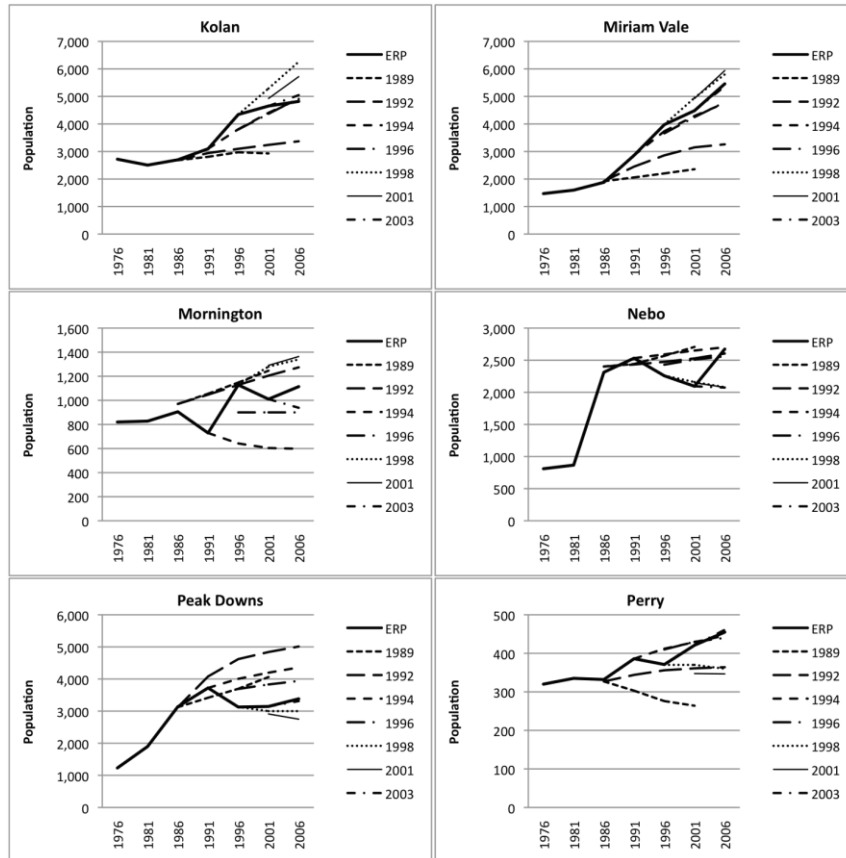
Projection horizon (years)	Annual average growth rate (%) over the 10 year base period						
	<-1	-1 to 0	0-1	1-2	2-3	3-4	4+
<i>MAPE</i>							
<i>2001 &amp; 2003 projection rounds</i>							
5	7.9	6.7	5.8	6.0	8.5	2.9	4.4
<i>1996 &amp; 1998 projection rounds</i>							
5	8.6	7.7	4.8	6.9	10.2	5.6	3.4
10	10.8	10.7	6.4	7.5	7.5	6.6	7.2
<i>1989, 1992 &amp; 1994 projection rounds</i>							
5	7.7	5.8	6.7	5.8	7.0	10.1	5.6
10	9.6	10.1	8.8	12.5	12.1	11.8	12.7
15	15.1	13.3	12.1	13.3	17.6	18.2	17.1
<i>MPAD</i>							
<i>2001 &amp; 2003 projection rounds</i>							
5	7.3	4.6	3.5	2.6	7.5	2.5	2.6
<i>1996 &amp; 1998 projection rounds</i>							
5	8.1	8.4	4.6	3.5	6.0	4.9	4.4
10	8.9	9.0	8.8	7.1	5.4	7.5	6.9
<i>1989, 1992 &amp; 1994 projection rounds</i>							
5	3.5	3.6	3.4	3.7	4.4	3.4	4.0
10	10.0	5.8	8.9	7.1	6.9	6.2	8.4
15	15.7	7.3	15.1	6.7	13.5	10.4	11.3

**4.2. FOR WHICH LGAs HAVE THE PROJECTIONS PROVED TO BE CONSISTENTLY POOR POPULATION FORECASTS?**

‘Consistently poor’ forecasts are defined here as those projections which after a projection horizon of five years have APes exceeding an average of 10% over all projection rounds. The affected LGAs are Aramac, Broadsound, Bulloo, Burke & Doomadgee, Croydon, Diamantina, Duaringa & Woorabinda, Ilfracombe, Kolan, Miriam Vale, Mornington, Nebo, Peak Downs, and Perry. These areas share a common feature of several changes in population trend, as Figure 1 demonstrates.







**Figure 1:** Projections and ERPs of LGAs with Consistently Poor Population Forecasts. Source: the Authors

Fluctuating mining employment is a key factor in several areas. As discussed by Tonts (2010) mining activity is often characterised by volatility, being dependent on the variability of global commodity prices, availability of the resource base, mining technologies, and labour supply amongst other factors. Periodic expansion and contraction of the mining-related labour force is common. The LGAs of Broadsound, Duaringa & Woorabinda, Nebo and Peak Downs are located in the Bowen Basin mining region and have a significant proportion of their employment provided by coal mining and associated construction. Coal mining employment in Queensland declined in the late 1990s but recovered from about 2000, increasing substantially from the mid-2000s due to rising

global demand, and prices, for coal (Rolfe et al. 2007). Census data show all four areas mirroring this state-wide trend, with substantial increases in mining and construction employment occurring between 2001 and 2006 (ABS 2007).

To what extent can forecast errors for these mining areas be blamed on statistical division projections for Fitzroy and Mackay to which the LGA projections were constrained? Table 6 shows that the answer is 'only some'. Whilst positive (negative) statistical division Percentage Errors are often accompanied by positive (negative) LGA Percentage Errors, the correlation between the two sets of errors is weak.

Other LGAs have also been affected by fluctuating mining employment. Burke & Doomadgee in the far north-west corner of the State also has a significant proportion of its workforce employed in mining as a result of a zinc mine opening in the late 1990s (CSRM 2004, 2008). In the local government area of Perry, growth of the small population has been largely due to the commencement of gold mining in 2001 (LGL 2009). Whilst many workers in the mining industry are employed on a fly-in-fly-out basis and live far away in major regional centres and capital cities, some employees do reside locally. For LGAs with small populations these locally-based employees can make a relatively large impact on ERPs. Given the difficulties of predicting mining activity, and crucially where employees are likely to be based, population forecasting for such areas is likely to remain problematic. Similar difficulties in forecasting local populations in mining areas have been experienced in Western Australia (WAPC 2004).

**Table 6.** Percentage Errors of Selected LGAs and Statistical Divisions.  
Source: the Authors.

	Projection round						
	1989	1992	1994	1996	1998	2001	2003
<i>Fitzroy statistical division</i>							
5 year projection horizon							
Duaringa & Woorabinda	7.2	0.4	14.4	30.1	17.0	-0.2	-17.4
Peak Downs	-8.2	9.6	28.1	22.7	-4.6	-16.3	-2.1
Fitzroy statistical division	2.8	-0.2	2.1	8.8	3.6	-2.8	-4.4
10 year projection horizon							
Duaringa & Woorabinda	29.1	19.2	38.2	28.9	6.1		
Peak Downs	18.2	47.7	33.4	17.8	-11.3		
Fitzroy statistical division	5.6	1.9	6.4	4.7	-1.2		
<i>Mackay statistical division</i>							
5 year projection horizon							
Broadsound	9.6	9.4	14.0	26.0	9.4	-8.1	-16.3
Nebo	-3.6	-3.9	14.8	18.8	3.2	-17.2	-22.4
Mackay statistical division	5.4	1.8	-1.5	2.9	3.1	-5.4	-7.9
10 year projection horizon							
Broadsound	31.6	34.1	33.4	18.0	-7.8		
Nebo	13.7	9.6	26.6	0.8	-22.2		
Mackay statistical division	7.9	2.6	0.6	-7.2	-6.4		

The problems with Mornington are quite different. Mornington shire consists of several islands just off the north west Queensland coast. Its resident population in 2006 was estimated to be around 1,100, about 90% of whom identified as Indigenous. Obtaining reliable estimates of past population trends is a major challenge for such areas. First, there are difficulties in enumerating Indigenous peoples in the census, especially in remote areas such as Mornington (Martin et al. 2002). Second, there are difficulties in trying to estimate annual population change between censuses. Indeed, *post-censal* ERPs for Mornington (rolled forward from the most recent census only) regularly show considerable differences from subsequently revised *intercensal* ERPs (based on two censuses). For example, the 1995 ERP rolled forward from the 1991 Census-based ERP was 741 (ABS 1997); the subsequently published intercensal ERP, 'anchored' to both 1991 and 1996 Census-based ERPs, was 1,042. The producers of the Queensland 1996 edition projections, which used 1995 as the launch year, felt uncomfortable with the 1995 postcensal ERP, noting that voter numbers at recent council elections raised questions about the ERP (Queensland Department of Local Government and

Planning incorporating Rural Communities 1996 p. 30). In the absence of robust estimates of past demographic trends there is little hope of producing reliable projections, and limited confidence in ERPs used to evaluate projections.

Reasons for Miriam Vale's forecast inaccuracy are harder to determine. The area has experienced substantial population increases over the last twenty years, with an annual average growth rate over its highest growth period, 1986-96, of 7.5%. Net migration has been the principal driver of population increase (Queensland Department of Infrastructure and Planning 2007). Possible reasons for consistently large forecast errors are not obvious except to say that the pace of migration, and therefore population growth, has fluctuated over time. It is not the case that the producers of the projections thought such a high growth rate unsustainable and under-forecast the population because two of the seven projections proved to be too high.

Aramac, Bulloo, Croydon, Diamantina, and Ilfracombe are LGAs with very small populations which are notoriously difficult to forecast. The loss or gain of a dozen jobs can result in a relatively large impact on populations which number just a few hundred people. Agriculture, Forestry and Fishing employment is significant for all these areas, though resident population change has not matched employment change in these industries very closely (ABS 2007). Population trends for such small areas tend to be inherently erratic and volatile, and realistically it is difficult to see how projections can improve in accuracy for such small populations.

#### **4.3. HOW DO FORECAST ERRORS COMPARE WITH THOSE FROM LINEAR EXTRAPOLATION?**

Table 7 compares projection errors of the official Queensland projections with those of linear extrapolation. It shows the percentage of the 90 LGAs whose populations were forecast more accurately by the official projections. As can be seen, in most projection rounds, a majority of LGAs were forecast more accurately by the official projections. However, this is not universal, and in every round a substantial percentage of LGAs' populations were forecast more accurately by extrapolation.

**Table 7.** Percentage of LGAs Whose Population was Forecast More Accurately by the Official Projections than by Linear Extrapolation. Source: the Authors.

Projection Period (years)	Projection round						
	1989	1992	1994	1996	1998	2001	2003
5	61.1	63.3	53.3	45.6	48.9	67.8	66.7
10	58.9	66.7	54.4	37.8	56.7		
15	52.2	64.4	53.3				

Table 8 provides a little more detail. It compares the two sets of projections using Proportionate Reduction in Error (PRE) of the three average error measures shown. The mostly positive values of PRE indicate the percentage reduction in average error when switching from linear extrapolation to the official projections. For example, the MPAD of the official 2003 round of projections after five years was 2.15% whilst for the linear extrapolation it was 3.84%. The PRE was thus  $(3.84 - 2.15) / 2.15 \times 100 = 78.6\%$ , a large reduction in error. For some projection rounds and horizons, however, the official projections performed less well. The 1996 round was not more accurate than linear extrapolation, and the 1989 round also performed relatively poorly when assessed by MedAPE and MPAD. In terms of MPAD, the best projections have proved to be the 1992, 1994 and 2003 rounds: these projections achieved the greatest reduction in error compared to naive linear extrapolation.

**Table 8:** Proportionate Reduction in Error (PRE) of Three Average Error Measures for Queensland LGA Projections Compared to Linear Extrapolations. Source: the Authors.

Projection Period (years)	Projection round						
	1989	1992	1994	1996	1998	2001	2003
	PRE of MedAPE (%)						
5	31.8	74.5	8.3	-8.6	11.3	17.8	33.0
10	-12.7	18.2	23.4	-28.6	17.1		
15	-5.4	14.7	0.6				
	PRE of MAPE (%)						
5	28.0	57.4	24.8	2.3	8.0	30.3	21.7
10	21.3	43.8	37.2	-4.7	12.3		
15	23.8	43.1	33.2				
	PRE of MPAD (%)						
5	13.1	67.2	72.7	-4.6	24.8	16.5	78.6
10	3.0	39.0	35.6	-1.6	4.6		
15	-0.3	23.7	33.5				

Note: Positive PRE values indicate the percentage reduction in average forecast error in switching to the official Queensland projections from the linear extrapolations.

#### 4.4. How do forecast errors compare with those reported by other studies?

How do Queensland LGA projections compare in accuracy to local area projections produced for other states and regions? To make valid comparisons only studies which assessed actual projections, as opposed to retrospective comparative projections, were considered. For illustrative purposes comparisons were made with forecast errors for:

- Western Australian LGAs (WAPC 2004);
- census tracts in San Diego county, California (Tayman 1996), and
- counties and places in Texas and North Dakota (Murdock et al. 1984).

These three studies were chosen because they reported their results in sufficient detail to enable reasonable comparisons. Although they cannot provide a representative sample of international forecast accuracy for local areas, they do at least permit some degree of context for the Queensland LGA projections. To permit some standardisation for

different population size distributions, errors are compared by population size category.

*Western Australian 1996-based LGA projections*

Forecast errors for Western Australian LGA projections were calculated from data presented in the Western Australian Planning Commission's detailed report on the accuracy of its 1996-based LGA projections after a projection horizon of five years (WAPC 2004). They are compared against the Queensland 1996-based projections (the round published in 1998) for 2001. The Western Australian projections have a MAPE of 9.1% whilst the equivalent error for Queensland is 5.9%. On the basis of MAPE the Queensland projections would appear to be significantly more accurate. However, a greater proportion of Western Australia's LGAs fall into the 0-1,999 population category where errors tend to be higher (45 out of 142 LGAs in Western Australia (32%) compared to 23 out of 90 in Queensland (26%). A more complex picture emerges if MAPE is presented by population size category, and if MPAD, which weights error by population size, is also calculated. Table 9 presents the figures. For the very smallest LGAs both MAPE and MPAD are lower for Queensland. For LGAs of 15,000 people or more both measures indicate smaller errors for Western Australia. For the 2,000-4,999 size category, MAPE and MPAD are in disagreement. Taking into account the states' different LGA population distributions, the two sets of projections are approximately comparable in accuracy, with MPAD being 3.3% in Queensland and 3.0% in Western Australia.

**Table 9.** Average Error of Western Australian and Queensland LGA 1996-based Projections After a Five Year Projection Horizon. Source: the Authors

Error	Launch year population size					All LGAs
	0-1,999	2,000-4,999	5,000-14,999	15,000-49,999	50,000+	
	Western Australia					
MAPE	13.7	7.9	6.1	2.6	2.0	9.1
MPAD	14.3	3.7	5.3	2.7	1.9	3.0
	Queensland					
MAPE	9.1	5.9	5.5	3.7	2.5	5.9
MPAD	8.5	6.6	5.3	3.3	3.0	3.3

*San Diego census tract projections*

Tayman (1996) reports on the errors of population projections for census tracts in San Diego county, California, produced by a spatial interaction land-use model. These projections were launched from 1980 and evaluated against 1990 census counts; errors were reported as MAPEs and MedAPEs but not MPADs. Queensland LGA projections were not produced as far back as 1980, so they are compared against Queensland projections over rounds 1989 to 1998 (all those which can be assessed after a 10 year horizon). Table 10 compares the two sets of errors. The timing differences render the comparison only partially valid, but nonetheless instructive. In this case the Queensland projections come out as less inaccurate.

**Table 10.** Average Error of San Diego Census Tract and Queensland LGA Projections After a Ten Year Projection Horizon. Source of San Diego projections: Tayman, 1996, Table 2.

	0-2,999	Launch year population size			All
		3,000-4,999	5,000-7,499	7,500+	
San Diego census tracts					
MAPE	34.6	18.6	17.6	17.7	20.9
MedAPE	18.1	13.0	14.4	13.0	14.6
Queensland LGAs					
MAPE	15.6	11.4	7.9	10.3	12.3
MedAPE	10.2	7.0	7.4	6.4	7.6

*Projections for local areas in Texas and North Dakota*

Murdock et al. (1984) assessed the accuracy of a regional economic-demographic model used to produce population projections for counties and incorporated places in North Dakota and Texas. The projections were launched from 1970 and assessed against 1980 census counts. Table 11 presents selected results from the Murdock et al. study and compares them with errors from Queensland LGA projections over rounds 1989 to 1998. For areas with the smallest populations, the Queensland projections are less inaccurate than the Texas and North Dakota projections according to both MAPE and MPAD, but as population size increases the differences in error are less marked. Overall, the Queensland projections



are more accurate than the place projections and roughly comparable with the county projections.

**Table 11:** Average Error of North Dakota and Texas County and Place Projections, and Queensland LGA Projections, After a Ten Year Projection Horizon. Source of North Dakota and Texas County and Place projections: Murdock et al. (1984), Tables 1 and 3

Error	0-500	501-1,000	Launch year population size				100,001+	All
			1,001-2,500	2,501-10,000	10,001-25,000	25,001-100,000		
Texas and North Dakota								
Counties								
MAPE	-	-	-	11.2	10.3	11.7	-	11.0
MPAD	-	-	-	10.9	9.3	11.4	-	10.6
Places								
MAPE	36.5	26.5	18.1	17.2	12.1	13.9	-	27.9
MPAD	32.2	29.1	17.7	16.8	13.3	14.3	-	14.5
Queensland LGAs								
MAPE	19.4	17.8	14.4	11.3	13.0	7.0	9.6	12.3
MPAD	13.1	13.4	15.3	13.5	12.9	10.4	7.7	9.0

Note: Results for categories with fewer than ten observations are not shown.

#### 4.5. TO WHAT EXTENT CAN FORECAST ERRORS BE ANTICIPATED?

Can the magnitude of population forecast errors for LGAs be anticipated? If this were the case then such information could provide the warning to users about likely future error that Keyfitz (1981) advocated. To answer this question, multiple regression was used with Absolute Proportional Error (APrE) as the dependent variable, i.e. APE/100. Proportional rather than percentage error was used because the model we chose required the distribution of the dependent variable to be between 0 and 1. Social science research regularly employs Ordinary Least Squares (OLS) regression models for such analyses. However, APrE has properties that prevent the valid application of OLS models. These are:

- The lack of constancy in its conditional variance and, in particular, the tendency of variance to decrease when the mean approaches one of the boundaries (heteroscedasticity);
- The non-normal distribution of errors;

- The logical range of APrE values from 0 to 1 which OLS models can potentially exceed (thus giving nonsensical predictions);
  - The tendency of explanatory variables to be non-linear in their effects.
- Hence, an alternative to OLS was sought. For variables like APrE, Kieschnick and McCullough (2003) advise researchers to use either parameterised models based upon a dependent variable with a beta distribution, or a fractional response model proposed by Papke and Wooldridge (1996). We adopt the latter approach because the former does not include values of 0 and 1 in the distribution.

Formally, the general model for the conditional expectation of the fractional response variable can be written as follows:

$$E(y_i|x_i) = G(x_i \beta)$$

where  $0 \leq y_i \leq 1$  denotes the dependent variable of observation  $i$ ,  $x$  a vector of explanatory variables and  $G(\cdot)$  a known cumulative distribution function that ensures the predicted values of  $y_i$  lie in the interval (0,1). Typically, that distribution function is the logistic function

$$G(x_i \beta) = \frac{\exp(x_i \beta)}{[1 + \exp(x_i \beta)]}.$$

The model is estimated using robust estimators of the asymptotic variance of  $\beta$  based on the well-known sandwich formula (Cameron and Trivedi 2005) through the maximisation of the Bernoulli log likelihood (McCullagh and Nelder 1989):

$$l_i(\beta) = y_i \log[G(x_i \beta)] + (1 - y_i) \log[1 - G(x_i \beta)].$$

A non-linear specification of the model was estimated using a pseudolikelihood estimator as described in Cameron and Trivedi (2005). The choice of predictor variables was informed by the existing literature (e.g. Tayman et al. 2010) and our own experimentation<sup>3</sup>, and resulted in APrE being predicted on the basis of:

---

<sup>3</sup> We excluded population density, east coast location, base period growth rate, and percentage of people living at a different address one year ago because either

- the natural logarithm of the LGA launch year population, *POP*,
- the proportion of the population identifying as Indigenous in the census preceding the launch year, *IND*,
- the proportion of the workforce employed in mining in the previous census, *MIN*,
- the most recently available five year APE of a previous projection, *PRV*, and
- dummy variables for each specific projection round.

The inclusion of the predictor variable PRV meant that the earliest round of projections which could be used to estimate the equation was that of 1994. The model for predicting error after a 5 year horizon was thus:

$$E(APrE|x) = G(\beta_1 + \beta_2 \ln(POP) + \beta_3 IND + \beta_4 MIN + \beta_5 PRV + \beta_6 1994 + \beta_7 1996 + \beta_8 1998 + \beta_9 2001)$$

The model for predicting APrE after 10 years could only use errors from rounds 1994 to 1998 (1998 being the most recent round assessable after a 10 year horizon). Thus:

$$E(APrE|x) = G(\beta_1 + \beta_2 \ln(POP) + \beta_3 IND + \beta_4 MIN + \beta_5 PRV + \beta_6 1994 + \beta_7 1996)$$

We estimated both models with and without dummy variables for each projection round (termed the ‘full model’ and the ‘no time effects model’, respectively).

Table 12 presents the estimated model coefficients and diagnostic statistics. To judge the goodness of fit we considered the Wald test to evaluate the joint significance of the estimated coefficients, and two information criteria for model comparison: the Akaike’s and the Schwarz’s information criterion (AIC and BIC, respectively). The lower the values of the AIC and BIC, the better the fit of the model (Long 1997). Hypothesis testing as for the traditional OLS model is widely used to evaluate the significance of individual coefficients. We therefore calculated z statistics and p values. Since our estimations are the result of a non-linear model, marginal effects at median values of the explanatory variables are presented to interpret the effect of estimated coefficients on APrE (Cameron and Trivedi 2005). The Wald test indicates that all coefficients are simultaneously different from zero at 1% significance. Comparison of the coefficient estimations in the full and no time effects

---

no identifiable relationship with APrE was apparent, or they were highly correlated with other variables in the model creating collinearity problems.

models suggests that projections in 1994 and 1996 were relatively more imprecise than those for the corresponding reference year. There is thus no convincing evidence that the change of projection method implemented from the 1996 projections onwards (Table 1) improved forecast accuracy.

**Table 12.** Coefficients and Diagnostic Statistics of the Fractional Response Models

Variable	5 year horizon				10 year horizon			
	Full model		No time effects		Full model		No time effects	
	Coefficien ts	Marginal effects	Coefficien ts	Marginal effects	Coefficien ts	Marginal effects	Coefficien ts	Marginal effects
POP	-0.182*** (0.026)	1.473** *	-0.186*** (0.026)	-1.493*** (0.210)	-0.139*** (0.037)	-1.101*** (0.291)	-0.144*** (0.037)	-1.126*** (0.291)
IND	1.167*** (0.336)	0.027** *	1.214*** (0.334)	0.027*** (0.008)	1.351*** (0.350)	0.026*** (0.007)	1.431*** (0.349)	0.028*** (0.007)
MIN	1.965*** (0.362)	0.011** *	1.978*** (0.395)	0.011*** (0.002)	2.188*** (0.500)	0.011*** (0.003)	2.207*** (0.534)	0.011*** (0.003)
PRV	1.253** (0.534)	0.054** (0.023)	0.992* (0.584)	0.042* (0.025)	0.661 (0.561)	0.027 (0.023)	0.277 (0.658)	0.011 (0.026)
1994 <sup>§</sup>	0.297** (0.125)	0.000** (0.000)			0.369*** (0.126)	0.000*** (0.000)		
1996	0.370*** (0.136)	0.000** (0.000)			0.332** (0.129)	0.000** (0.000)		
1998	-0.041 (0.133)	0.000 (0.000)						
2001	0.086 (0.139)	0.000 (0.000)						
CONSTANT	-1.504*** (0.273)		-1.304*** (0.247)		-1.600*** (0.346)		-1.290*** (0.351)	
N	450		450		270		270	
R-squared	0.33		0.30		0.27		0.24	
Wald test	203.41		164.57		100.08		81.67	
P-value	0.00		0.00		0.00		0.00	
AIC	0.41		0.39		0.50		0.49	
BIC	-2677.85		-2701.53		-1457.90		-1468.53	
Log pseudolikelihood	-82.20		-82.57		-60.51		-60.79	

\* 10% significance, \*\* 5% significance, \*\*\* 1% significance.

Note: Standard errors in brackets.  $R^2$  corresponds to the correlation between observed and fitted values of APrE. Marginal effects are calculated as elasticities. It measures the proportional change in APrE associated with a given proportionate change in one of the explanatory variables. For further details, see Cameron and Trivedi (2005).

§ The reference projection round for the dummy variables for the 5-year horizon models was 2003, while for the 10-year horizon model it was 1998.

The models provide further statistical evidence to support our previous findings. Population size is negatively associated with APrE, while the proportion of the population identifying as Indigenous, the proportion of the workforce employed in mining, and the errors of previous projections are positively associated with APrE. Most of these estimated coefficients are statistically significant, except for those associated with the APrE corresponding to earlier rounds of projections for the 10 year projection horizon models. The results from marginal effects associated with 5 year horizon projections suggest that APrE decreases by approximately 1.5% (1.1% in the case of 10 year horizon projections) when the natural logarithm of population increases by 1%, whereas it rises between 0.04% and 0.05% when the APrE for a previous projection increases by 1%. Likewise, 1% increases in the proportion of the population identifying as Indigenous and the proportion of the workforce employed in mining increase APrE by around 0.27% and 0.11% respectively. These effects are similar at both 5 and 10 year horizons. Marginal effects associated with individual projection rounds are virtually nil.

To aid interpretation of the models' ability to predict error we additionally calculated the squared-correlation between the predicted and actual values of APrE, commonly called  $R^2$ . The modest values shown in Table 12 confirm the response to the question 'To what extent can the magnitude of error be anticipated?', is 'only to some extent'.

## 5. SUMMARY AND RECCOMENDATIONS

This paper has evaluated the accuracy of all official Queensland government LGA population projections up to and including the 2003 round. The key findings of the study can be summarised as follows.

- Errors in Queensland LGA population forecasts were large compared to those of the State and large relative to user tolerability (section 3), with only 60-70% of LGAs being forecast within 10% APE after a decade (Table 3).
- Error increased substantially with projection horizon length (Table 2).
- Error was negatively associated with launch year population size in a nonlinear relationship (Table 4).
- There was no relationship between the growth rate of an LGA's population in the base period and subsequent forecast error (Table 5).

- LGAs whose populations were especially poorly forecast were those in mining regions, with majority Indigenous populations, and with very small populations (section 4.2).
- Queensland LGA projections proved more accurate than simple linear extrapolations (though not for every projection round) (section 4.3).
- They were also respectably accurate relative to some other local area population projections (section 4.4). In addition, the value of different perspectives on error, as measured by MAPE and MPAD for example, was demonstrated in this section.
- An attempt to predict population forecast error on the basis of local characteristics met with only some success (section 4.5).

What lessons can be learned from this assessment of LGA forecast errors? We make some suggestions under two broad headings: first, forecast accuracy, and second, information on the likely error of current projections.

#### *Forecast Accuracy*

On the issue of forecast accuracy, it would be worth evaluating different methods of producing local area population projections to ascertain whether certain methods proved consistently better than others. Alternative methods could be assessed both conceptually and empirically, the latter by producing retrospective projections and comparing them to current population figures. For example, how do projections differ between net migration rate cohort-component models, directional migration probability cohort-component models, the Hamilton-Perry model (Hamilton and Perry 1962), various extrapolative models, and so on? Much of the previous work comparing projection models has focused primarily on which type of extrapolative model produces the best projections, with different forms of cohort-component model largely absent from the analyses. These tests can never *completely* replicate the conditions under which past projections were produced, but if they are produced according to an explicit set of rules they can come close.

Related work is needed to better understand local demographic and economic processes and their links with migration. Migration is often the largest and most volatile of the demographic components of change in LGAs but our understanding of its relationship with employment, housing, commuting, and other social and cultural factors is piecemeal.

Various local migration projection methods could be evaluated as part of the retrospective projection tests.

Can much be done about the large percentage errors for LGAs with the smallest populations? Some short-term gain in accuracy may be obtained from detailed local studies just mentioned, but it is likely the smallest areas will always experience large errors. A more radical alternative would be to only produce projections for areas above a certain population threshold, grouping neighbouring small areas with below-threshold populations into larger regions. But perhaps the issue of large *percentage* errors in the forecasts of these small populations is not such a serious problem after all because their *absolute* errors remain small. For resource allocations and decisions dependent on *numbers* of people, absolute rather than percentage errors are more relevant.

Although stating the obvious, a more reliable dataset on past population change is also likely to reduce the likelihood of really bad projections. The example of Mornington shire in section 4.2 clearly shows the problems of basing projections on unreliable Estimated Resident Populations. Given that post-censal ERPs are provisional and revised following the subsequent census, it would be best to base all local area projections on census-year ERPs which have been finalised. Although these ERPs are unlikely to be perfect (they are, after all, estimates), they are probably more reliable than non-census year ERPs because they are ‘anchored’ to a census count of the population.

#### *The Likely Error of Projections*

Some gains in accuracy may be obtained as a result of the above suggestions, but forecast error will always exist, and at the local scale it is likely to remain quite substantial. User expectations about accuracy need to be carefully managed. There is a difficult balancing act here, of course, between honesty about likely error on the one hand and the appearance of competency and professionalism on the other. It is important to stress to users that there are many factors affecting local demographic change which are essentially unpredictable, and that similar evaluations of forecasts from economics, marketing, transport and other disciplines also reveal large errors (e.g. Loungani 2000; Parthasarathi and Levinson 2010). In fact, rather than focusing too much on the word ‘error’ with its connotations of wrongdoing or incompetence, it may be more useful to emphasise ‘forecast uncertainty’ instead.

One possible approach to providing users with information on uncertainty was attempted in our regression analysis. The modest success

of this was disappointing, and it means that reliable prediction intervals for projections probably cannot be created in this way. Certainly our analysis could be extended by searching for additional predictor variables, but it would seem unlikely to result in the huge increases in  $R^2$  which would be required to give robust error predictions. Other ways of warning users about uncertainty are needed. The conventional approach has been to produce high and low projection variants. But high and low variants have been widely criticised by demographers: they are very imprecise in their meaning (how likely are they?) and tend not to be based on much evidence (Lee 1999; Keilman et al. 2002). A better alternative would be to publish forecast errors from previous projection rounds for each LGA together with average errors for LGAs in that population size category. Doing so would be far from precise, but it would at least offer users a ballpark indication of uncertainty. Interestingly, the publication of past errors has very occasionally appeared in official projection reports, such as the US Census Bureau's 1998-based national population projections (US Bureau of the Census, 1989 pp 14-15).

An alternative would be to produce probabilistic projections for individual LGAs, a conceptually more refined approach but one which would be complex and time-consuming. Although a few examples of subnational probabilistic projections do exist in the literature (Wilson and Bell 2007; Cameron and Poot 2011) considerable methodological development would be required. In addition, researchers would have to tackle the challenge that probabilistic methods are more data hungry than conventional deterministic methods, but data tend to be less detailed, less accurate and less temporally extensive at the local scale.

Complementing technical work on quantifying forecast uncertainty, research on the ways in which uncertainty is communicated to, and perceived by, non-technical users would be valuable. There has been almost no exploration of this in demography, but fortunately it has been addressed in meteorology and its literature offers some useful guidance (e.g. WMO 2008; Morss et al. 2008). For example, certain prediction intervals could be expressed in qualitative terms, ranging from "virtually certain" (greater than 99% probability) and "very likely" (90+% probability) to "unlikely" (less than 10% probability) and "exceptionally unlikely" (less than 1% probability). Various alternative methods of communicating uncertainty should be trialled. By providing users with indications of forecast uncertainty demographers are not only being scientifically more honest, but are providing users with additional



information which will assist them in making better informed decisions. Paradoxically, by being frank about population forecast uncertainty, user confidence in projections may increase.

**ACKNOWLEDGEMENTS:** The assistance of C-J Rohlin of Queensland Treasury in providing some of the data for this study is gratefully acknowledged. Comments on an earlier draft by Martin Bell and Jim Cooper, and the two anonymous referees, helped to improve the paper. All errors, of course, remain our own.

## REFERENCES

- ABS (1984). Estimated Resident Population in Local Authority Areas, Queensland, 30 June 1976 to 1981. Catalogue No. 3212.3. ABS: Brisbane.
- ABS (1997). *Estimated Resident Populations and Area, Preliminary, Queensland*. Catalogue No. 3201.03. ABS: Brisbane.
- ABS (2000). 'Past ABS projections: how well have they matched reality?' Appendix in ABS (2000) *Population Projections, Australia: 1999 to 2101*. Catalogue No. 3222.0. ABS: Canberra, pp 147-152.
- ABS (2007). 2006 Census Community Profile Series: Time Series Profile [Excel workbooks]. ABS: Canberra.
- Adam, A. Y. (1992). The ABS population projections: overview and evaluation. *Journal of the Australian Population Association*, 9(2) pp. 109–130.
- APRU (Applied Population Research Unit) (1992). *SAASPPS: Multi Regional Population Projection System*. Department of Geographical Sciences and Planning, The University of Queensland: Brisbane.
- Armstrong, J. S. (2001). Evaluating forecasting methods. In J.S. Armstrong (Ed) *Principles of Forecasting: A Handbook for Researchers and Practitioners*, Kluwer Academic Publishers: Norwell, MA, pp. 443-472.
- Bell, M. and Skinner, J. (1992). Forecast accuracy of Australian subnational population projections. *Journal of the Australian Population Association*, 9(2) pp. 207-235.
- Cameron, M. and Poot, J. (2011). Lessons from stochastic small area population forecasts: the case of Waikato sub-regions in New Zealand. Forthcoming in the *Journal of Population Research*.
- Cameron, C. and Trivedi, P. (2005). *Microeconometrics: Methods and Applications*, Cambridge University Press: Cambridge.
- Campbell, P. R. (2002). Evaluating Forecast Error in State Population Projections Using Census 2000 Counts. Population Division Working Paper No. 57. US Census Bureau: Washington DC.
- CSRM (Centre for Social Responsibility in Mining) (2004). Aboriginal employment at Century Mine, Centre for Social Responsibility in Mining Research Paper No. 3, Centre for Social Responsibility in Mining, The University of Queensland: Brisbane.

- CSRSM (Centre for Social Responsibility in Mining) (2008). Completion of Mining at Oz Minerals Century Mine: Implications for Gulf Communities, Centre for Social Responsibility in Mining, The University of Queensland: Brisbane.
- Khan, H. T. A. and Lutz, W. (2008). How well did past UN population projections anticipate demographic trends in sex south-east Asian countries? *Asian Population Studies* 4(1) pp. 77-95.
- Hamilton, C. and Perry, J. (1962). A short method for projecting population by age from one decennial census to another. *Social Forces*, 41(December), pp. 163–170.
- Keilman, N. (2001). Data Quality and Accuracy of United Nations Population Projections. 1950-95, *Population Studies*, 55(2) pp. 149-164.
- Keilman, N. (2008). European Demographic Forecasts Have Not Become More Accurate Over the Past 25 Years. *Population and Development Review*, 34(1) pp. 137-153.
- Keilman, N. and Pham, D. Q. (2004). Empirical errors and predicted errors in fertility, mortality and migration forecasts in the European Economic Area. Discussion Paper No. 386, Statistics Norway: Oslo.
- Keilman, N., Pham, D.Q., and Hetland, A. (2002). Why population forecasts should be probabilistic - illustrated by the case of Norway. *Demographic Research*, 6, pp. 409-454.
- Keyfitz, N. (1981). The limits of population forecasting. *Population and Development Review*, 7(4), pp. 579-593.
- Kieschnick, R. and McCullough, B. (2003). Regression analysis of variates observes on (0,1): percentages, proportions and fractions. *Statistical Modelling*, 3, pp. 193-213.
- Kolassa, S. and Schutz, W. (2007). Advantages of the MAD/mean ratio over the MAPE. *Foresight*, 6, pp. 40-43.
- Lee, R. D. (1999) Probabilistic approaches to population forecasting. In W. Lutz, J. W. Vaupel and D. A. Ahlburg (Eds) *Frontiers of Population Forecasting*, Supplement to *Population and Development Review* 24, Population Council: New York, pp. 156–190.
- LGL (Lihir Gold Limited) (2009). *Mt Rawdon*. LGL: Brisbane.
- Long, J. (1997). *Regression models for categorical and limited dependent variables*. SAGE Publications Inc: California.
- Long, J. (1995). Complexity, accuracy, and utility of official population projections. *Mathematical Population Studies*, 5, pp. 203-216.

- Loungani, P. (2000). How accurate are private sector forecasts? Cross-country evidence from consensus forecasts of output growth, IMF Working paper 00/77, International Monetary Fund: Washington DC. <http://www.imf.org/external/pubs/ft/wp/2000/wp0077.pdf>
- Martin, D. F., Morphy, F., Sanders, W. G. And Taylor, J. (2002). Making Sense of the Census: Observations of the 2001 Enumeration in Remote Aboriginal Australia. Research Monograph No. 22, Centre for Aboriginal Economic Policy Research, the Australian National University, Canberra.
- McCullagh, P. and Nelder, J. (1989). *Generalized Linear Models*, 2nd edition, Chapman and Hall: London.
- Morss, R. E., Demuth, J. L. and Lazo, J. K. (2008). Communicating Uncertainty in Weather Forecasts: A Survey of the U.S. Public. *Weather and Forecasting*, 23, pp. 974-991.
- Mulder, T. J. (2002). Accuracy of the US Census Bureau national population projections and their respective components of change, *Population Division Working Paper Series No. 50*, US Census Bureau: Washington DC.
- Murdock, S. H., Leistriz, 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.
- National Research Council (2000). *Beyond Six Billion: Forecasting the World's Population*. National Academy Press: Washington DC.
- ONS (2008). Subnational Population Projections Accuracy Report. Office for National Statistics: London.  
[http://www.statistics.gov.uk/downloads/theme\\_population/SNPP\\_Accuracy\\_Report.pdf](http://www.statistics.gov.uk/downloads/theme_population/SNPP_Accuracy_Report.pdf)
- Openshaw, S. and van der Knapp, G. A. (1983). Small area population forecasting: some experience with British models. *Tijdschrift voor economische en sociale geografie*, 74(4), pp. 291-304.
- Papke, L. and Wooldridge, J. (1996). Econometric methods for fractional response variables with an application to 401 (k) plan participation rates. *Journal of Applied Econometrics*, 11(6), pp. 619-632.
- Parthasarathi, P. and Levinson, D. (2010). Post-construction evaluation of traffic forecast accuracy. *Transport Policy*, 17, pp. 428-443.
- Queensland Department of Communication & Information, Local Government and Planning (1998). *Population Projections for Queensland: 1998 Edition*. Department of Communication & Information, Local Government and Planning: Brisbane.

- Queensland Department of Housing and Local Government (1992). *Population Projections for Queensland and Statistical Divisions 1986-2021 and Local Government Areas and Statistical Districts 1986-2006: 1991 Revision*. Department of Housing and Local Government: Brisbane.
- Queensland Department of Housing, Local Government and Planning (1994). *Population Projections for Queensland and Statistical Divisions 1991-2031 and Local Government Areas and Statistical Districts 1991-2011*. Department of Housing, Local Government and Planning: Brisbane.
- Queensland Department of Infrastructure and Planning (2007). *Miriam Vale Shire: Population and Housing Factsheet*. Department of Infrastructure and Planning: Brisbane.
- Queensland Department of Local Government and Planning (2001). *Population Trends and Prospects for Queensland: 2001 Edition*. Department of Local Government and Planning: Brisbane.
- Queensland Department of Local Government and Planning (2003). *Queensland's Future Population: 2003 Edition*. Department of Local Government and Planning: Brisbane.
- Queensland Department of Local Government and Planning (incorporating Rural Communities) (1996). *Population Projections for Queensland*. Queensland Department of Local Government and Planning (incorporating Rural Communities): Brisbane.
- Rayer, S. (2007). Population forecast accuracy: does the choice of summary measure of error matter? *Population Research and Policy Review*, 26, pp. 163–184.
- 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. (2010). Factors Affecting the Accuracy of Subcounty Population Forecasts. *Journal of Planning Education and Research*, 30(2), pp. 147-161.
- Rolfe, J., Miles, B., Lockie, S. and Ivanova, G. (2007). Lessons from the social and economic impacts of the mining boom in the Bowen Basin 2004-2006. *Australasian Journal of Regional Studies*, 13(2), pp. 134-153.
- Shaw, C. (2007) Fifty years of United Kingdom national population projections: How accurate have they been? *Population Trends*, 128, pp. 8–23.
- Siegel, J. S. (2002). *Applied Demography: Applications to Business, Government, Law and Public Policy*. Academic Press: San Diego.

- Skinner, J. L., Bell, M. and Gillam, M. E. (1989). *Population Projections for the Local Government Areas of Queensland 1986-2001*. Applied Population Research Unit, Department of Geographical Sciences, The University of Queensland: Brisbane.
- Smith, S. K. and Tayman, J. (2003). An evaluation of population projections by age. *Demography*, 40(4), pp. 741-757.
- Smith, S. (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 Sincich, T. (1988). Stability Over Time in the Distribution of Population Forecast Errors. *Demography*, 25, pp. 461-74.
- Smith, S., and Tayman J. (2003). An evaluation of population projections by age. *Demography*, 40, pp. 741-757.
- Smith, S., and Shahidullah, M. (1995). An evaluation of population projection errors for census tracts. *Journal of the American Statistical Association*, 90: pp. 64-71.
- Smith, S. and Sincich, T. (1990). The relationship between the length of the base period and population forecast errors. *Journal of the American Statistical Association*, 85, pp. 367-375.
- Smith, S. and Sincich, T. (1992). Evaluating the forecast accuracy and bias of alternative population projections for states. *International Journal of Forecasting*, 8, pp. 495-508.
- Statistics New Zealand (2008). *How Accurate are Population Projections? An Evaluation of Statistics New Zealand Population Projections, 1991-2006*. Statistics New Zealand: Wellington.
- Swanson, D.A. and Tayman, J. (1995). Between a rock and a hard place: the evaluation of demographic forecasts. *Population Research and Policy Review*, 14(2): pp. 233-249.
- Tayman, J. (1996). The accuracy of small-area population forecasts based on a spatial interaction land-use modelling system. *Journal of the American Planning Association*, 62(1), pp. 85-98.
- Tayman, J. and Swanson, D. A. (1996). On the utility of population forecasts. *Demography*, 33(4), pp. 523-528.
- Tayman, J. and Swanson, D. A. (1998). 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. (2010). Evaluating population forecast accuracy: a regression approach using county data. *Population Research and Policy Review*, Online First.

- Tayman, J., Swanson, D. A. and Barr, C. F. (1999). In search of the ideal measure of accuracy for subnational demographic forecasts. *Population Research and Policy Review*, 18, pp. 387–409.
- Tayman, J., Schafer, E. and Carter, L. (1998). The Role of Population Size in the Determination and Prediction of Population Forecast Errors: An Evaluation Using Confidence Intervals for Subcounty Areas. *Population Research and Policy Review*, 17, pp. 1–20.
- Tonts, M. (2010). Labour market dynamics in resource dependent regions: an examination of the Western Australian goldfields. *Geographical Research*, 48(2), pp. 148-165.
- 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.
- US Bureau of the Census (1989). *Projections of the Population of the United States by Age, Sex and Race 1988 to 2080*. Current Population Reports Series P 25 No. 1018. US Government Printing Office: Washington DC.
- WAPC (Western Australian Planning Commission) (2004). *Are our Population Projections on Target?* Western Australian Planning Commission: Perth.
- Wilson, T. (2007). The forecast accuracy of Australian Bureau of Statistics national population projections. *Journal of Population Research*, 24, pp. 1-27.
- Wilson, T. and Bell, M. (2007). Probabilistic Regional Population Forecasts: The Example of Queensland, Australia. *Geographical Analysis*, 39, pp. 1-25.
- WMO (World Meteorological Organization) (2008) *Guidelines on Communicating Forecast Uncertainty*. WMO: Geneva, Switzerland. [www.wmo.int/pages/prog/amp/pwsp/documents/TD-1422.pdf](http://www.wmo.int/pages/prog/amp/pwsp/documents/TD-1422.pdf)

**APPENDIX: LGAs included in the study**

The 90 LGAs (or LGA combinations) included in the study were: Aramac, Atherton, Balonne, Banana, Barcaldine, Barcoo, Bauhinia, Beaudesert, Belyando, Biggenden, Blackall, Booringa, Boulia, Bowen, Brisbane, Broadsound, Bulloo, Bungil, Burke & Doomadgee, Caboolture, Calliope, Caloundra, Cardwell, Cherbourg & Murgon & Wondai, Chinchilla, Croydon, Dalby, Diamantina, Duaringa & Woorabinda, Eacham, Eidsvold, Emerald, Etheridge, Fitzroy, Flinders, Gayndah, Gladstone, Gold Coast, Goondiwindi, Herberton, Hervey Bay, Ilfracombe, Inglewood, Isis, Isisford, Jericho, Johnstone, Jondaryan, Kilkivan, Kilcoy, Kolan, Livingstone, Logan, Longreach, Mackay, Mareeba, Maroochy, McKinlay, Millmerran, Mirani, Miriam Vale, Monto, Mornington, Mount Isa, Mount Morgan, Mundubbera, Murilla, Murweh, Nebo, Noosa, Paroo, Peak Downs, Perry, Pine Rivers, Quilpe, Redcliffe, Redland, Richmond, Rockhampton, Roma, Sarina, Stanthorpe, Taroom, Tiaro, Toowoomba, Townsville, Waggamba, Warwick, Whitsunday, Winton.