Distr.: General
30 March 2016

English only

Economic Commission for Europe

Conference of European Statisticians

Joint Eurostat/UNECE Work Session on Demographic Projections
Geneva, 18-20 April 2016
Item 3 of the provisional agenda
Sub-national projections

Testing a simple averaged model for local and regional population forecasts

Tom Wilson, Australia[1]

Summary

Local and regional population forecasts inform a wide range of planning and budgeting activities, including those concerning educational provision, health facilities, electoral redistricting, and business location decisions. Unfortunately such forecasts often prove to be quite inaccurate. The aim of this paper is to evaluate a simple model for forecasting local and regional total populations in Australia which takes the average of two extrapolative methods. This is the Constant Share of Population – Variable Share of Growth (CSP-VSG) model, shown to have performed well at the local area scale in earlier research. This study extends that earlier work, making use of recently available historical local area population estimates. It reports on retrospective tests of the averaged model over several forecasting periods, and at three geographical scales. Forecasts are produced for three ten year forecast horizons and comparisons are made with simple linear extrapolation. It is shown that for all geographical scales and forecast horizons, the averaged model generally produces more accurate population forecasts than linear extrapolation. It is argued that the CSP-VSG averaged model is a useful addition to the population forecaster's toolkit as it produces forecasts of respectable accuracy with low input data requirements and production costs.

[1] [Prepared by Tom Wilson, Principal Research Fellow, Charles Darwin University].

I. Introduction

1. Subnational population forecasting is a challenging endeavour. Analyses of past forecasts confirm the difficulty of producing accurate population figures even for just one decade ahead, let alone further into the future (e.g. Statistics New Zealand 2008; Wilson & Rowe 2011). Age-specific population numbers usually turn out to be less accurate than total population forecasts, which is unfortunate because forecasts for particular age groups are widely used for service delivery and planning. Observed levels of accuracy in evaluation studies are also generally at odds with expectations. Users expect accurate forecasts with considerable demographic detail, such as population numbers by single years of age and sex for a fine-grained geography, for many decades into the future, and broken down by various socio-economic characteristics. This is difficult enough at the national scale. But demographers tasked with producing subnational population forecasts often have to deal with less detailed, lower quality, and noisier base period data than at the national scale – especially for the most important demographic component, migration – as well as more limited methodological and theoretical guidance from the literature. There are often hundreds or thousands of local areas and regions for which forecasts have to be produced. And in many organisations staff numbers and budgets are lower than they once were.

2. For local and regional population forecasting the literature offers a much greater variety of methods compared to the national scale (Smith et al. 2013; Wilson 2011). This is due to factors such as the diverse nature of available demographic data, staff resources, budgets, project timeframes, and stakeholder requirements, as well as the personal methodological preferences of those preparing the forecasts. Methods vary from simple extrapolation (e.g. Hachadoorian et al. 2011), various forms of cohort-component model (e.g. Swanson et al. 2010; Wilson 2011, 2015), dwelling-led approaches such as the housing-unit method (e.g. Foss 2002), regression models (e.g. Chi & Voss 2011), disaggregation approaches (e.g. Li & Corcoran 2011), land use development and dwelling allocation models (e.g. Bell et al. 2000; Pullar et al. 2015), and microsimulation (e.g. Marois & Bellanger 2015; Wu and Birkin 2013). Sometimes several methods are combined in a single forecasting system.

3. However, little research in demographic forecasting has investigated the combination of methods through averaging. An averaged forecast is simply the mean of two or more models' forecasts. The approach is more prominent in the general forecasting literature where its benefits have been appreciated for some time (e.g. Bates & Granger 1969; Clemen 1989; Goodwin 2009). The highly regarded textbook *Principles of Forecasting* has a whole chapter devoted to "Combining Forecasts" (Armstrong 2001). Why does averaging often yield greater accuracy? Goodwin (2009 p. 34) explains that "when several methods are combined, there is a likelihood that biases in different directions will counteract each other, thereby improving accuracy." Ahlburg and Lutz (2001 p. 5) advise that "combining forecasts is most likely to improve accuracy when the forecasts combined use different methods and approaches that capture different information sets or specifications".

4. The few demographic studies investigating averaging in population forecasting include Rayer & Smith (2010), Reinhold & Thomsen (2015), Smith and Shahidullah (1995), and Wilson (2014), who report that in many situations averaging yields improvements in forecast accuracy over individual methods, or at least provides forecasts about as

accurate as the best individual method. But so far the evidence base on averaging in population forecasting is small, and the methodology is far from fully developed. Not surprisingly, averaging approaches are rarely applied in practice. The Bureau of Economic and Business Research at the University of Florida is one of the few organisations that implements averaging as a standard technique in preparing sub-State population forecasts for Florida (Rayer & Wang 2015). But such applications are few and far between.

5. This paper sets out to contribute to the empirical evidence and methodology of averaging in subnational population forecasting. It focuses on an averaged forecasting model for total populations, the Constant Share of Population – Variable Share of Growth (CSP-VSG) model, which was found to work well in Australia in an earlier investigation (Wilson 2014). Key strengths of this model include its ease and speed of calculation in a spreadsheet and minimal input data requirements. It is therefore a low cost option which appears to generate relatively accurate forecasts. However, the earlier study only covered a single period (2001-2011) at a single spatial scale (SA2 areas, which contain between 2,500 and 25,000 residents). This paper extends that earlier work by assessing the CSP-VSG averaged model over several forecast periods and at three spatial scales in Australia, namely, SA2, SA3 and SA4 areas of the Australian Statistical Geography Standard (termed here local areas, minor regions, and major regions, respectively). It does so with retrospective applications of the model over the periods 1991-2001, 1996-2006 and 2001-2011, and uses simple linear extrapolation as a comparative basic forecast. Specifically the paper addresses the question 'Does the CSP-VSG averaged model generate short-term population forecasts for a range of subnational geographies which are of acceptable accuracy and more accurate than those from linear extrapolation?' Acceptable accuracy is taken to be under 10% error (Tye 1994). The paper also undertakes analysis to answer the question 'For which areas were CSP-VSG averaged model forecasts successful and where were they less successful?'. Such analysis will be helpful in assessing the circumstances under which the averaged model should generally work well, and where further research is required.

II. Data and methods

a) Averaged forecasting model

6. The averaged forecasting model is the Constant Share of Population – Variable Share of Growth (CSP-VSG) model (Wilson 2014). Population forecasts are calculated as the mean of the two constituent models' outputs. In the Constant Share of Population model the population of each local area or region is forecast as a fixed proportion of an independent State (or national) population forecast, i.e.

$$P^i\ (t+h) = P^{State}(t+h)\ p^i$$

where P denotes population, i the local area, t+h a point in time in the future, and p the local area's share of the State population. In this study the shares were calculated from the jump-off year populations. The Variable Share of Growth model shares out forecast State population growth for each projection interval amongst local areas, i.e.

$$P^i\ (t+h) = P^i\ (t) + \left[G^{State}(t,t+h)\ \ g^i(t,t+h)\right]$$

where G refers to State population growth over the t,t+h forecast interval and g each local area's share of State growth. Provisional local area population growth is obtained from a linear model if the area's base period growth is positive and an exponential model if negative. This is a conservative approach which avoids excessive population increases or decreases (including 'negative' populations). Provisional growth is then adjusted across local areas to match State growth using the plus-minus method (Smith et al. 2013). The resulting averaged forecasts pull local area growth rates towards the average, reducing rapid decline or growth.

7. Population forecasts from the CSP-VSG model were created for the periods 1991-2001 (minor regions and major regions only due to a lack of data for local areas), 1996-2006 and 2001-2011. For each set of forecasts a ten year base period was used to estimate the local areas' shares of growth. A small number of SA2 local areas, those with populations under 100, were excluded from the analysis. These were mostly non-residential areas, such as ports, airports and industrial estates.

b) Linear extrapolation

8. Comparative forecasts for the same forecast periods were created from linear extrapolations of recent population growth, i.e.

$$P^i\ (t+h) = P^i\ (t) + h\ G^i$$

where G represents annual growth and h the number of years from the jump-off year t. Annual growth was that of the preceding decade:

$$G^i = \frac{P^i\ (t) - P^i\ (t-10)}{10}.$$

9. To be consistent with the CSP-VSG model, the linear forecasts were adjusted to sum to State and Territory total populations. The extent to which the averaged forecasts out-perform linear extrapolation indicates the 'value added' by the averaged model over this naïve forecasting approach.

c) State and Territory population forecasts

10. The CSP-VSG model requires independent State or Territory population forecasts. Medium series population projections produced by the Australian Bureau of Statistics (ABS) published shortly after the jump-off date of each set of forecasts were used. These are the likely State/Territory forecasts which would have been applied had the local and regional forecasts been produced on the release of 1991, 1996 and 2001 Estimated Resident Populations (ERPs). In addition, a second set of forecasts were produced using actual ERPs as the State forecasts in order to remove the effect of any inaccuracy in the ABS projections.

11. Two further sets of forecasts were then produced, one constrained to forecasts for the metropolitan section of each state/territory, and the other to forecasts for its remaining non-metropolitan part. The reason for doing so is because rapid population growth is occurring in parts of many metropolitan regions due to large-scale housing (re)development which, at least in theory, is better suited to dwelling-led, rather than extrapolative, models of forecasting. The two-way division of each state/territory into metropolitan and non-metropolitan was based on the ABS's definition of Greater Capital City Statistical Areas. Because no real population forecasts were produced for the current geographical definition of these regions, the forecasts were constrained to the relevant metropolitan and non-metropolitan ERPs for each State and Territory.

d) Estimated Resident Populations

12. Forecasts were evaluated by comparing them with Estimated Resident Populations, the official usually resident population estimates produced by the Australian Bureau of Statistics. Following the introduction of the current statistical geography in 2011, the ABS produced total population ERPs on this geography back to 1991 for SA2 local areas and back to 1981 for SA3 minor regions and SA4 major regions. SA2 population numbers for 1986 were estimated by Wilson et al. (2015), but that study was unable to produce reliable SA2 estimates for 1981 due to data limitations. The lack of a 1981-91 base period prevented SA2 local area forecasts being produced for 1991-2001.

13. The statistical geography of 2011 was designed to keep the range of populations at each geographical scale fairly small. In 2011 SA2 local areas had a median population of 9,039 (with an interquartile range of 5,458 to 14,913), SA3 minor regions had a median population of 57,411 (38,011 to 87,677), while the equivalent figures for SA4 major regions were 224,465 (157,124 to 317,428).

e) Error measures

14. Forecasts for each area were assessed with Absolute Percentage Error (APE):

$$APE = \frac{|F - ERP|}{ERP} \; 100\%$$

where F denotes the population forecast. APE then formed the basis of several summary measures of error, including Median Absolute Percentage Error (MedAPE), and the percentage of areas forecast with less than 10% APE (%<10APE). MedAPE gives a good indication of typical error when the error distribution is skewed, as it is in the case of APEs for population forecasts. However, whilst low average errors should be celebrated, it is also preferable to forecast as many areas as possible within acceptable limits and avoid very high errors. The %<10APE measure is useful in recording the proportion of forecasts that are within acceptable (10%) limits.

III. Results

a) Overall results

15. Table 1 presents average errors from all four sets of forecasts 10 years out. Those from the averaged model are on the left hand side of the table, whilst forecasts from the comparative linear extrapolations are on the right. It can be seen that errors decrease as population size increases, matching findings from previous research. MedAPEs from the averaged model were mostly 6-8% for SA2 local areas, 3-6% for SA3 minor regions and 2-4% for SA4 major regions. Differences in accuracy between the forecasts constrained to State projections and those constrained to State ERPs are mostly small, suggesting that constraining to State ERP ‘forecasts’ provides a reasonable indication of the magnitude and patterns of local and regional errors which would be obtained from real forecasts.

16. Overall, the averaged model performed better across geographies and time periods than linear extrapolation by a reasonable, but not huge, margin (though not for every individual set of forecasts). The percentage point gain in accuracy in using the averaged model rather than linear extrapolation proved greatest at the SA2 local area scale and smallest for the SA4 major regions. For example, the reduction in MedAPE for all State

projection-constrained forecasts was 2.3% at the local area scale (7.0% rather than 9.3%), 1.3% for minor regions (5.0% rather than 6.3%), and 0.7% for major regions (3.4% rather than 4.1%). The proportional reduction in MedAPE did not show a clear pattern across geographical scales and, aside from two sets of SA4 major region forecasts where the averaged model proved inferior to the linear model, these reductions varied from 7% to 38%.

Table 1: Median Absolute Percentage Errors of population forecasts after 10 years

Geography	Constraint	CSP-VSG averaged model				Linear extrapolation			
		1991-2001	1996-2006	2001-2011	All periods	1991-2001	1996-2006	2001-2011	All periods
SA2 local areas	State projections	n/a	7.7	6.5	7.0	n/a	10.1	8.6	9.3
	State ERPs	n/a	7.9	7.3	7.6	n/a	9.8	7.9	8.8
	Metro region ERPs	n/a	8.2	7.6	8.0	n/a	10.4	8.2	9.3
	Non-metro ERPs	n/a	7.4	6.0	6.7	n/a	8.4	7.1	7.8
SA3 minor regions	State projections	5.5	4.6	4.8	5.0	6.8	6.0	6.2	6.3
	State ERPs	5.7	4.8	4.6	4.9	6.6	5.8	5.5	5.9
	Metro region ERPs	5.2	5.1	4.6	4.9	7.9	6.8	5.7	6.9
	Non-metro ERPs	5.7	3.3	3.5	4.1	4.4	4.7	4.7	4.7
SA4 major regions	State projections	3.4	3.0	3.9	3.4	4.4	3.7	4.3	4.1
	State ERPs	3.2	2.7	3.2	3.1	3.7	3.3	3.4	3.6
	Metro region ERPs	2.8	2.4	3.0	2.6	3.4	3.9	4.1	3.9
	Non-metro ERPs	3.7	1.9	2.1	2.4	3.1	2.5	2.5	2.7

Note: Shaded cells indicate where linear extrapolation out-performed the averaged model.

17. Forecasts constrained to the ERPs of metropolitan and non-metropolitan sections of States resulted in slightly different levels of error. Constraining to metropolitan region ERPs using the averaged model produced marginally higher forecast errors than State-level constraining for SA2 local areas, about the same for SA3 minor regions, and slightly lower errors for SA4 major regions. Importantly, constraining to non-metropolitan region ERPs gave lower average errors across all geographies than State-constrained forecasts. For example, MedAPEs for all averaged model forecasts were 6.7% at the SA2 local area scale compared to 7.6% when constrained to State-level ERPs, 4.1% at the SA3 minor region scale (compared to 4.9%), and 2.4% at the SA4 major region scale (compared to 3.1%). However, at all three geographical scales, the percentage point reduction in error relative to linear extrapolation was smaller for non-metropolitan-constrained forecasts than metropolitan-constrained forecasts.

18. With regards to variations in the averaged model's performance across the three forecast periods, the 1991-2001 period proved the most difficult to forecast. It had slightly larger average errors than the other two periods, and a lower percentage point and proportional reduction in error from the linear model's forecasts. The 1996-2006 period was the most successfully forecast, having the lowest average errors and the greatest reduction in error from the linear model.

19. The forecast accuracy picture is similar – though not identical – from the perspective of the percentage of areas forecast with less than 10% error (%<10APE). Table 2 presents the results. The proportion of areas forecast with under 10% error from the averaged model was 58-67% for SA2 local areas, 66-86% for SA3 minor regions and 88-100%

for SA4 major regions. Mostly, the averaged model out-performed linear extrapolation, and the percentage point gain in %<10APE was greatest at the SA2 local area scale, and greater for the metropolitan-constrained forecasts than the non-metropolitan-constrained forecasts.

Table 2: Percentage of areas forecast with less than 10% error after 10 years

Geography	Constraint	CSP-VSG averaged model				Linear extrapolation			
		1991-2001	1996-2006	2001-2011	All periods	1991-2001	1996-2006	2001-2011	All periods
SA2 local areas	State projections	n/a	61	67	64	n/a	50	56	53
	State ERPs	n/a	59	62	60	n/a	51	59	55
	Metro region ERPs	n/a	58	61	59	n/a	48	57	53
	Non-metro ERPs	n/a	64	67	66	n/a	57	62	60
SA3 minor regions	State projections	71	77	80	76	70	71	74	72
	State ERPs	69	79	79	76	71	73	74	73
	Metro region ERPs	66	77	76	73	66	68	71	68
	Non-metro ERPs	75	84	86	82	83	85	86	85
SA4 major regions	State projections	88	92	90	90	93	91	86	90
	State ERPs	91	95	97	94	91	92	98	93
	Metro region ERPs	95	95	98	96	84	84	95	88
	Non-metro ERPs	90	98	100	96	100	100	98	99

Note: Shaded cells indicate where linear extrapolation out-performed the averaged model.

20. So in response to the first research question posed in the introduction to this paper, ‘Does the CSP-VSG averaged model generate short-term population forecasts for a range of subnational geographies which are of acceptable accuracy and more accurate than those from linear extrapolation?’, the answer to the linear extrapolation part is ‘yes, for the majority of sets of forecasts’. The answer to the acceptable accuracy (under 10% error) part is ‘to some extent’ at the SA2 local area scale, ‘mostly’ at the SA3 minor region scale, and ‘to a large extent’ at the SA4 major region scale.

b) Errors by type of area

21. To explore the sorts of areas that were forecast well or badly by the averaged model, average errors were calculated by jump-off population size, metropolitan/non-metropolitan region, base period growth rate, and base period growth rate variability. This last variable was defined as the absolute difference between the annual average growth rate of the first five years of the base period and the annual average growth rate of the second five years.

22. Table 3 presents MedAPEs for the State projection-constrained set of forecasts after 10 years by population size category. At the SA2 scale errors were negatively associated with population size, though not linearly. There is a sudden drop in the magnitude of error from the very smallest population category of 100-2,499 to the next smallest, and then more modest falls with increasing population size. For all SA2 local area size categories the CSP-VSG model out-performed linear extrapolation by a noticeable margin, though by a larger percentage point and proportional amount for the smallest population categories. At the SA3 and SA4 scales the averaged model mostly out-

performed linear extrapolation. Again, there is a generally negative association between average error and population size at the SA3 scale and an unclear relationship at the SA4 scale.

Table 3: Median Absolute Percentage Errors of population forecasts after 10 years by jump-off population size category; projection-constrained

Geography	Jump-off population	CSP-VSG averaged model	Linear extrapolation
SA2 local areas	100-2,499	20.8	30.6
	2,500-4,999	8.2	11.5
	5,000-9,999	7.5	9.3
	10,000-14,999	5.7	7.9
	15,000-24,999	4.9	6.6
	25,000+	4.6	5.2
SA3 minor regions	100-24,999	7.9	7.8
	25,000-34,999	5.5	6.9
	35,000-44,999	4.7	5.9
	45,000-59,999	5.1	6.0
	60,000-99,999	4.4	6.3
	100,000+	3.2	5.4
SA4 major regions	100-149,999	4.1	4.0
	150,000-249,999	3.1	4.1
	250,000+	3.3	4.2

Note: Shaded cells indicate where linear extrapolation out-performed the averaged model.

23. Table 4 presents the equivalent MedAPEs for areas categorised by how fast they grew in population over the ten year base period. For almost all categories the averaged model gave lower average errors than linear extrapolation, with the most notable differences in error observed for areas declining in population during the base period. Errors for SA2 local areas and SA3 minor regions exhibit an approximately U-shaped pattern of error, confirming previous findings, though the pattern is more pronounced for linear extrapolation. Forecasts for areas which experienced very high growth in the base period were the most erroneous and are not reliable.

24. Table 5 presents MedAPEs for areas classified according to the volatility of their base period population growth, as measured by the absolute difference between the annual average growth rates of the first and second halves of the ten year base period. Again, the averaged model gave lower average errors than linear extrapolation for almost all categories for both models. Differences in average errors between categories up to about 2% are not huge. High errors occurred for those areas experiencing large swings in growth rates over the base period, indicating that areas undergoing the greatest variations (3% or more) were not forecast reliably.

Table 4: Median Absolute Percentage Errors of population forecasts after 10 years by annual average base period growth rate; projection-constrained

Geography	Base period growth rate (%)	CSP-VSG averaged model	Linear extrapolation
SA2 local areas	< -1%	6.3	16.2
	-1 to 0%	5.3	8.0
	0 to 1%	5.1	5.4
	1 to 2%	7.1	7.3
	2 to 3%	9.1	12.0
	3 to 4%	10.0	12.2
	4 to 5%	10.1	13.8
	5% +	17.8	24.9
SA3 minor regions	< -1%	7.5	10.9
	-1 to 0%	4.1	7.9
	0 to 1%	4.0	4.3
	1 to 2%	5.8	6.3
	2 to 3%	6.0	7.5
	3 to 4%	4.6	6.8
	4 to 5%	8.7	7.7
	5% +	10.1	10.4
SA4 major regions	< -0%	3.2	7.2
	0 to 2%	3.4	3.9
	2% +	4.0	4.6

Note: Shaded cells indicate where linear extrapolation out-performed the averaged model.

Table 5: Median Absolute Percentage Errors of population forecasts after 10 years by base period growth rate variability; projection-constrained

Geography	Growth rate variability (%)	CSP-VSG averaged model	Linear extrapolation
SA2 local areas	0-0.1	5.7	8.5
	0.1-0.5	5.3	7.1
	0.5-1	5.0	6.9
	1-2	6.3	8.4
	2-3	9.3	9.7
	3+	13.9	19.0
SA3 minor regions	0-0.1	4.1	6.4
	0.1-0.5	4.2	5.2
	0.5-1	4.4	6.3
	1-2	5.1	6.4
	2-3	8.6	7.8
	3+	12.9	11.6
SA4 major regions	0-0.5	2.9	3.9
	0.5-2	4.3	4.3
	2+	5.6	6.8

Note: Shaded cells indicate where linear extrapolation out-performed the averaged model.

25. Table 6 shows MedAPEs for areas classified by metropolitan or non-metropolitan location. The averaged model performed better than linear extrapolation at all three geographical scales and for both region types. Although areas within non-metropolitan regions were forecast with lower average errors, the value added by employing the averaged model over linear extrapolation proved greater in metropolitan regions.

Table 6: Median Absolute Percentage Errors of population forecasts after 10 years by metropolitan/non-metropolitan location; projection-constrained

Geography	Location	CSP-VSG averaged model	Linear extrapolation
SA2 local areas	Metropolitan	7.1	10.5
	Non-metropolitan	6.9	7.9
SA3 minor regions	Metropolitan	5.3	7.7
	Non-metropolitan	4.6	5.1
SA4 major regions	Metropolitan	3.5	4.7
	Non-metropolitan	3.3	3.7

c) Comparison with other forecast evaluation studies

26. How do the averaged model's forecasts compare with those from other studies? Given the use of a variety of summary error measures and different ways of categorising populations by size, growth rate, and other characteristics, the ability to make comparisons is limited. As a consequence, just one case study is used for comparison here: that of Wilson and Rowe (2011) which examined the forecast accuracy of several sets of official Queensland local government area population projections. These were produced with several methods, including cohort-component models, a housing-unit model in metropolitan areas, and extrapolative techniques, although all sets of projections incorporated manual adjustments on the basis of feedback from local government areas.

27. The database of the Queensland forecast error study included five sets of official projections with errors available for 10 year forecast horizons. With 90 local government areas in the study, this gave a sample of 450 projections. The Queensland areas were categorised into the same population size categories used for the present paper, and because they varied in population size from a few hundred to hundreds of thousands they were compared against the averaged model and linear extrapolation across all three levels of geography. Table 7 shows the MedAPEs from the Queensland local government area projections alongside those for the projection-constrained forecasts from the CSP-VSG model and linear extrapolation.

28. As can be seen from the table, the performance of the CSP-VSG model relative to the Queensland local area projections was mixed. Overall, the averaged model did better, with the MedAPE across all areas being 6.3% rather than 7.6%. But when the results are examined by population size category, the Queensland projections were more accurate for the smallest populations, especially the smallest category of 100-2,499 people. Part of the explanation for this lower average error might be related to the nature of the very smallest areas in Queensland and the SA2 local area geography. Boundary changes in Queensland meant some very small Aboriginal communities (which are notoriously hard to forecast) had to be merged with neighbouring areas. SA2 local areas were

designed to have minimum populations of 2,500 wherever possible, and those with fewer than this number are often newly developing urban fringe areas, remote Aboriginal communities, and small mining-dominated settlements. The better performance of the Queensland projections for areas with 2,500 to 4,999 people is harder to explain, but might be related to better projections of rapidly developing urban fringe locations with the housing-unit model. Despite the mixed results overall, the averaged model was the best performer for populations of 10,000 or more.

Table 7: Median Absolute Percentage Errors of population forecasts after 10 years by jump-off population size category compared across three types of projections

Geography	Jump-off population	CSP-VSG averaged model	Linear extrapolation	Queensland projections
All geographies	100-2,499	20.1	30.2	10.7
	2,500-4,999	8.2	11.5	7.3
	5,000-9,999	7.5	9.3	7.4
	10,000-14,999	5.9	7.9	6.5
	15,000-24,999	5.1	6.6	6.4
	25,000+	4.3	5.4	6.2
	All sizes	6.3	8.1	7.6

Source: Queensland projection errors obtained from the database constructed for Wilson & Rowe (2011).
Note: Shaded cells indicate where the Queensland projections out-performed the averaged model.

d) Modelling errors

29. The types of areas which experienced small and large forecast errors were then assessed through multiple regression. Only the results from the State projection-constrained forecasts are shown here because the ERP-constrained forecasts are very similar. The dependent variable was Absolute Percentage Error. Only a limited number of predictor variables were available on the SA2 local area geography unfortunately, and these mostly describe population size, recent growth and location. Consideration was given to including some 2011 Census variables such as the proportion of the labour force in mining employment and the proportion of the population identifying as Indigenous. An obvious weakness of these variables is their incorrect reference date, and in any case it was found that they added very little explanatory power to the regression model, and were therefore excluded.

30. The final model included four common predictor variables: the natural log of population size at the jump-off year, the absolute base period growth rate, metropolitan or non-metropolitan location, and volatility of the base period growth rate. The variables were chosen on the basis of previous research indicating their importance (e.g. Tayman et al. 2011; Wilson and Rowe 2011), and their impact in terms of contributing to R^2. Table 8 presents a summary of the regression results for both 1996-2006 and 2001-2011 forecast periods.

31. As expected, errors were negatively associated with population size and positively associated with the absolute base period growth rate and the volatility of base period growth. Being located in a metropolitan region increased the error relative to a non-metropolitan location. On the basis of these regressions, it would be expected that SA2 local areas subject to the highest forecast errors would generally have several of the following characteristics: location within a metropolitan region, small in population,

high rates of base period growth or decline, and considerable volatility in this base period growth. Unfortunately the explanatory power of the models is weak so the ability to predict the error for individual areas is limited.

Table 8: Regression results for CSP-VSG model forecast errors for SA2 local areas after 10 years; projection-constrained

	Forecast period	
Effect	1996-2006	2001-2011
Intercept	73.218***	61.658***
ln(population)	-7.376***	-6.299***
Base period growth rate	0.378***	0.312***
Base period growth rate volatility	0.269***	0.960***
Metropolitan / non-metropolitan	3.873***	3.992***
Adjusted R^2	0.220	0.235

*** Significant at 0.001

32. However, the ability to predict whether forecasts will fall within three broad Absolute Percentage Error categories of 0-10%, 10-20% and 20%+ is a little better. The regression model for the 1996-2006 period correctly placed 51% of SA2 local areas in the right error category while the model for 2001-2011 achieved 58%. Of course, regression coefficients will change over time and the regression models estimated here will not be exactly right for forecasts being prepared today. But the use of models such as these might provide forecasters with some clues about which areas are likely to be difficult-to-forecast with the averaged model. The initial forecasts might then be amended, or an alternative forecasting model applied for such areas.

33. Similar regressions were run for SA2 local area forecasts constrained to metropolitan and non-metropolitan ERPs. Again, the adjusted R^2 values were disappointing and are not reported here. Regression models were also created for SA3 minor region and SA4 major region 4 forecast errors but most of the predictor variables were statistically insignificant; these results are not reported here.

IV. Discussion and conclusions

34. The results of this paper have demonstrated that the CSP-VSG averaged model produces lower average forecast errors, and a greater proportion of areas forecast within 10% error, than simple linear extrapolation. Relative to the Queensland local area population projections the average errors also appear reasonably low for areas of 10,000 people or more, though further investigation is required to achieve lower errors for areas with the smallest populations. The percentage point reduction in error from linear extrapolation is greater at the SA2 local area scale than at the higher SA3 and SA4 geographies, though the results at these higher geographies are still quite good. The non-metropolitan-constrained forecasts achieve noticeably more accurate results than those constrained to State-level forecasts, while the metropolitan-constrained forecasts were slightly better or worse than the State-level forecasts depending on the spatial scale. Importantly, the forecasts are simple and easy to produce, requiring minimal data inputs and staff time and resources.

35. However, it is important to be aware of the limitations of the CSP-VSG model. Although it requires minimal data inputs, it also has limited data outputs in the form of total populations only. It is largely atheoretical in that neither the proximate drivers of population change (fertility, mortality, migration, and age structure), nor the underlying social and economic drivers, are considered. It will not work for areas with zero population at the start of the base period, such as greenfield areas at the urban fringe undergoing rapid residential development. Although average forecast errors are fairly low, a proportion of areas will be forecast poorly and have high errors. And there is limited flexibility to create alternative assumptions, except through changes to the independent projection for the State or major region that the CSP-VSG model requires as an input. The strengths and weakness of the averaged model are summarised in Table 9 below.

Table 9: Strengths and weaknesses of the CSP-VSG averaged model

Strengths	**Weaknesses**
Simple model: easy to understand	Atheoretical: doesn't incorporate demographic drivers of population change, including the effects of age structure on population change
Low input data requirements: only total populations for two past time points required plus independent State/large region population forecast totals	Outputs total population only (no components of change or age detail)
Largely automated: ready-made Excel spreadsheet template available	As an extrapolative model cannot be applied to areas with zero population at the start of the base period
Forecasts can be produced very quickly and cheaply for hundreds or thousands of areas	Performs poorly for some areas undergoing large-scale residential (re)development
Relatively low average errors demonstrated for Australia over 10 year forecast horizons	Some areas will have large errors (see Tables 8 and 9)
Links to an independent forecast for a State/large region or other 'parent region'; local area forecasts automatically sum to that of parent region	Difficult to incorporate local area-specific assumptions and alternative scenarios
Reduces decline of declining populations and slows growth of rapidly growing populations	

36. So what do the findings presented in this paper mean for population forecasting practitioners? I suggest that the averaged model could form a useful part of a subnational population forecasting system. Three ways in which it could be used are proposed here.

- First, it could form an integral part of a subnational forecasting system for all, or just the non-metropolitan part, of a State. The averaged model would produce the total population forecasts, and a cohort-component model would be added to produce age-sex forecasts. Inward and outward migration flows in the cohort-component model would be adjusted to match the averaged model's population total (as in Wilson 2015). For areas with the metropolitan section of a State, a housing-unit (or similar dwelling-led) model could be applied where reliable dwelling forecast data were available.

- Second, it could play a role in validating the forecasts from another model by providing an independent set of forecasts. Where large differences in the two sets of forecasts were evident, it would be useful to ask if there was a good reason for the large difference. If there was indeed a compelling reason for the difference, then there would be no problem; but in the absence of clear reasons, further checks would need to be made.

- Third, the averaged model could be used to produce a benchmark set of forecasts when undertaking retrospective tests of potential forecasting models. The averaged model forecasts would be the ones to beat.

37. Would the CSP-VSG averaged model work well in other countries? Work by Wilson (2014) which evaluated averaged models for three countries indicated that averaged models added more value in Australia than in England & Wales and New Zealand. Part of the explanation might lie in the greater variability and volatility of growth rates experienced by Australian small areas compared to those in the other case study countries. The only way to be sure is to conduct a similar evaluation to the one reported in this paper in which alternatives are assessed empirically.

References

Ahlburg, D., & Lutz, W. (1999). Introduction: The need to rethink approaches to population forecasts. In W. Lutz, J. Vaupel, & D. Ahlburg (eds.), Frontiers of Population Forecasting (pp 1-14). New York: Population Council.

Armstrong, J. S. (2001). Combining forecasts. In J. S. Armstrong (ed.), Principles of Forecasting: A Handbook for Researchers and Practitioners (pp 417-439). New York: Springer.

Bates, J. M., & Granger, C. W. J. (1969). The combination of forecasts. Operational Research Quarterly, 20.4, 451-468.

Bell, M., Dean, C. & Blake, M. (2000). Forecasting the pattern of urban growth with PUP: a web-based model interfaced with GIS and 3D animation. Computers, Environment and Urban Systems, 24, 559-581.

Chi, G., & Voss, P. (2011). Small-area population forecasting: borrowing strength across space and time. Population, Space and Place, 17.5, 505-520.

Clemen, R. T. (1989). Combining forecasts: a review and annotated bibliography. International Journal of Forecasting, 5: 559-583.

Foss, W. (2002). Small area population forecasting. The Appraisal Journal, 70.2, 163-172.

Goodwin, P. (2009). New evidence on the value of combining forecasts. Foresight, 12, 33-35.

Hachadoorian, L., Gaffin, S. R., & Engelman, R. (2011). Projecting a gridded population of the world using ratio methods of trend extrapolation. In R.P. Cincotta & L. J. Gorenflo (eds.), Human Population: Its Influences on Biological Diversity. (pp 13-15). Berlin: Springer-Verlag.

Li, T. & Corcoran, J. (2011). Testing dasymetric techniques to spatially disaggregate the regional population forecasts for South East Queensland. Journal of Spatial Science, 56:2, 203-221

Marois, G. & Belanger, A. (2015), Analyzing the impact of urban planning on population distribution in the Montreal metropolitan area using a small-area microsimulation projection model. Population and Environment, 37.2, 131-156.

Pullar, D., Bell, M., Cooper, J., Stimson, R. & Corcoran, J. (2015). Forecasting Patterns of Metropolitan Growth Using an Optimised Allocation Procedure. In T. Wilson, E. Charles-Edwards, & M. Bell (eds.), Demography for Planning and Policy: Australian Case Studies (pp 109-129). Cham: Springer.

Rayer, S., & Smith, S. (2010). Factors affecting the accuracy of subcounty population forecasts. Journal of Planning Education and Research, 30.2, 147-161.

Rayer, S., & Wang, Y. (2015). Projections of Florida Population by County, 2015-2040, with Estimates for 2014. Bureau of Economic and Business Research, University of Florida.

Reinhold, M., & Thomsen, S. L. (2015). Subnational Population Projections by Age: An Evaluation of Combined Forecast Techniques. Population Research and Policy Review, 34.4, 593-613.

Smith, S. K., & Shahidullah, M. (1995), An evaluation of population projection errors for census tracts. Journal of the American Statistical Association, 90.429, 64-71.

Smith, S. K., Tayman, J., & Swanson, D. A. (2013). A Practitioner's Guide to State and Local Population Projections. Dordrecht: Springer.

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

Swanson, D., Schlottmann, A., & Schmidt, B. (2010). Forecasting the population of census tracts by age and sex: an example of the Hamilton-Perry method in action. Population Research and Policy Review, 29.1, 47-63.

Tayman, J., Smith, S.K., & Rayer, S. (2011). Evaluating population forecast accuracy: a regression approach using county data. Population Research and Policy Review, 30.2, 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.

Wilson, T. (2011). A Review of Sub-Regional Population Projection Methods. Report to the Office of Economic and Statistical Research. Queensland Centre for Population Research, School of Geography, Planning and Environmental Management, The University of Queensland, Brisbane.

Wilson, T. (2014). New evaluations of simple models for small area population forecasts. Population, Space and Place, 21.4, 335-353.

Wilson, T. (2015). Evaluation of alternative cohort-component models for local area population forecasts. Population Research & Policy Review http://dx.doi.org/10.1007/s11113-015-9380-y

Wilson, T., & Rowe, F. (2011). The forecast accuracy of local government area population projections: a case study of Queensland. Australasian Journal of Regional Studies, 17, 204-243.

Wilson, T., Ueffing, P., & Bell, M. (2015). Spatially-consistent local area population estimates for Australia, 1986-2011. Journal of Population Research, 32.3, 285-296.

Wu, B., & Birkin, M. (2013). Moses: A Dynamic Spatial Microsimulation Model for Demographic Planning. In R. Tanton & K.L. Edwards (eds.) Spatial Microsimulation: A Reference Guide for Users (pp 171-192). Doredecht: Springer.