A checklist for reviewing draft population projections

NORTHERN INSTITUTE
People. Policy. Place.

Tom Wilson

Northern Institute, Charles Darwin University, Ellengowan Drive, Darwin, NT 0909, Australia

Email: tom.wilson@cdu.edu.au

May 2017

Abstract

The literature on population projections says very little about one crucial step of the projections production process – checking draft projection outputs for errors and plausibility. Unfortunately, it is quite possible to obtain implausible projections even when using the latest projection models and carefully prepared input data. This paper presents a checklist to assist with checking draft population projection outputs for plausibility. Applying the checklist does not guarantee accurate projections, but it should help identify serious errors and embarrassingly implausible projections prior to publication.

Key words

Population projections; reviewing; plausibility; checklist

Comments?

Suggested additions or amendments to this checklist, or other comments, are most welcome. Please send them by email using the address above.

Introduction

Research on population projections has contributed many important methodological advances to the field in recent decades (discussed in Smith et al. 2013 and Wilson & Rees 2005 amongst others). The literature covers a wide range of projection methods, applications, computer programs, input data sources, and data estimation and adjustment techniques. But very few demographic projections textbooks and academic papers mention one very important part of the process of producing projections – checking draft projection outputs for errors and plausibility. Murdock et al. (1991) and Smith et al. (2013, chapter 12) are amongst a handful of publications that address the issue.

This paper presents a checklist to assist with checking draft population projection outputs for plausibility. It originated as a set of notes created by the author several years ago to systematically check his own projections (after failing to notice some input data errors which generated embarrassingly awful projections). What is published here is an expanded and revised version which hopefully others will find useful. It contains some limitations. The checklist focuses on projection outputs only and does not evaluate projection *methods* or *assumptions*, nor does it deal with forecast *accuracy*. It has been created largely for subnational, and especially local area, population projections, though many of the checks could also be applied to projections at the national scale. And the checklist assesses plausibility across four sets of projection outputs only:
(1) Projected total population trends
(2) Projected demographic components of change
(3) Projected age-sex structures
(4) Projected population change across areas.

What constitutes 'plausible' of course covers quite a wide range of possibilities, with large grey areas at the margins. There are no definite answers, especially at the smaller end of the population size spectrum because of the volatility of demographic trends at that scale. This checklist focuses on identifying serious problems, such as when draft projections are clearly temporally inconsistent (different from historical or likely future trends), spatially inconsistent (different from projections for similar areas), or internally inconsistent (e.g. the demographic components of change do not sum to total projected change).

It is the case, of course, that some projections are not designed to be plausible. They may illustrate the results of 'no change' in certain demographic rates, or extreme scenarios such as the future amount of migration needed to halt population ageing (UN Population Division 2000). They may show desired demographic futures or pathways towards particular targets. Or they may exclude one demographic component, such as migration, in order to highlight

the role of migration in a 'standard' projection which includes it. But even for these alternative projections, many of the checklist questions in this paper remain valid.

Why is it even necessary to review draft population projections? Surely the latest projection models produce projections free of problems? Unfortunately, even with state-of-the-art projection models and carefully prepared input data and assumptions, it is quite possible to obtain implausible projections. The reasons include:

- Poor quality, incorrect or poorly adjusted input data, such as
    - Data which have coverage or quality limitations
    - Data which have been perturbed or partially suppressed
    - Data which do not relate to the current boundaries of a region
    - Data which are highly disaggregated and very sparse so that even smoothed age profiles are unrepresentative of underlying rates/probabilities
    - Data based on a different definition to the one required in projections (e.g. census counts rather than the Estimated Resident Population)
- The use of unsmoothed and ragged age profiles which are unrepresentative of the underlying rates/probabilities.
- Sub-optimal choice of projection assumptions
- Inadvertent errors in preparing assumptions, such as
    - choosing the wrong value in a pull-down menu in the projection software;
    - pasting input data into the wrong place (e.g. mixing up male and female populations);
    - putting a decimal point in the wrong place;
    - inadvertent inclusion of a negative sign;
    - rates/probabilities containing calculation errors,
- Projection software errors or bugs
- The omission of mechanisms in projection models to maintain certain demographic relationships (such as correlations between males and females, and between different geographical areas) in even the most advanced projection models.

In the event of finding implausible projections it may be useful to check that:

- the jump-off populations are all correct (in terms of year, region, sex, age groups, etc.)
- all projection assumptions have been correctly entered (e.g. males and female inputs have not been inadvertently swapped; or five year values have not been entered instead of annual values)
- the projection program has been run with the intended options, constraints and other settings
- any constraining mechanisms (such as those constraining local area projections to sum to State projections) are not altering projections too much from the input assumptions

- projected demographic components of change are in the right ballpark. They can often provide clues about the nature of the problem
- the selected projection method/software really is appropriate to the task in hand.

If the source of the problem is not obvious it might be useful to produce projections using a simple a cohort-component model (e.g. Wilson 2014a), or at least create total population projections using a basic extrapolative approach (e.g. Wilson 2014b).

Checklist

Draft population projections are best assessed with population data covering at least the past 10 years, and preferably longer periods. Checks under headings 1-3 are for individual areas, whilst those under heading 4 are comparative checks across all areas in a set of projections to assess spatial consistency. Not all checks listed here will be relevant to every set of projections, and at the same time, it is important to stress that the checklist is not exhaustive and will not include some checks required for particular types of projections.

1. Total population

(a) *Does the projected total population seem plausible in the light of past trends?*

- Create a graph which shows past and projected total population.
- Do the projections join the estimates well? Normally there should be no discontinuity between the population estimates and the projections, or at least no greater difference between the jump-off population and the first year of the projections than the recent past range of year-to-year population changes.
- How does the projected amount of growth over 5 or 10 year intervals compare to past trends? The direction and rate of projected population growth should not be too dissimilar from the past unless there are credible reasons for it.
- If there is a noticeable change from past trends, one reason may be that considerable dwelling growth is anticipated (often projected by a dwelling-led model). Or the region might experience fluctuating population numbers due to the type of industry located there. Or a large non-private dwelling (e.g. a prison) might be scheduled to open in the future.
- As a general rule, the larger the population, the greater the likely continuity in trend over time (because larger populations tend to experience smaller net migration fluctuations relative to population size).

(b) *If projected growth is very high, is it credible?*

- Where projected population growth in an area is very high, calculate and graph the projected population density.
- Particularly for urban areas, check that the projected population density is (i) likely to be within planning controls, and (ii) not considerably higher than other built-out high-density areas.
- If the high growth is due to the development of a new suburb, check that the projected population trajectory is similar to that of other recently constructed suburbs (often a flat s-shaped curve).

(c) *Is the projected total population similar to previous projections for this area, and those produced by other reputable forecasters?*

- Graph the current and previous projections, and projections produced by others if they are available.
- Is the current projection close to previous projections and others' projections?
- If not, is there a good reason for the difference, such as?
    - ABS made significant revisions to the base period data;
    - migration trends have experienced a major change in direction;
    - dwelling forecasts for the area have changed dramatically;
    - a major non-private dwelling is scheduled (or no longer scheduled) to be built;
    - a significant piece of proposed infrastructure is no longer likely.

(d) *Does the total projected population decline by a substantial amount?*

- If so, be absolutely sure it is a robust projection – it may prove very unpopular! Investigate whether recent population decline (on which the projections are based) is likely to be a one-off event or a long-term phenomenon.
- It may be wise to err a little on the generous side for such areas for political reasons. Because such areas tend to have small populations to start with, this can usually be achieved with very minor relative changes to the assumptions for other areas.

2. Demographic components of change

(a) *Do the demographic components of change match total population change for each projection interval?*

- Calculate population change over each projection interval as (i) the result of applying a demographic accounting equation using the demographic components of change, and (ii) the difference between the total populations at the start and end of the interval. They should be the same.

(b) *Do the projected demographic components of change seem plausible in the light of past trends?*

- Graph past and projected components of change (total births, deaths and migration flows).
- Do the projected components join recent observed values well? Normally there should be no discontinuity between past trends and projections of births, deaths, and directional migration, or at least no difference greater than recent year-to-year changes. (Net migration trends may well fluctuate considerably, especially for small areas).
- The projected trajectory of each component should not be too dissimilar from the past unless there is a credible reason for it.
- Projected births and deaths should exhibit fairly smooth trends over time. Migration may be more variable, especially for smaller areas.

(c) *Do the projected births and deaths seem plausible in the light of projected age structure changes?*

- Does the trend in projected births seem about right given projections of the Total Fertility Rate and size of the female childbearing age population?
- Does the trend in projected deaths seem about right given mortality assumptions and the projected size of the elderly population? If an area is projected to experience substantial growth of the elderly population, then even with declining mortality there should be a significant increase in the number of deaths.

(d) *Does the total net migration assumption (internal + overseas migration) in broad terms align with what is known about the local economic prospects of the area?*

- Is projected total net migration (internal + overseas migration) plausible in light of the area's general economic prospects?
- For example, if local information suggests no or little prospect of employment increases in an area which does not have any retirement/lifestyle migration, then you would not expect high positive levels of net migration.

3. Age-sex structure

(a) *Do the projections by age and sex sum to the output total population projections? (This applies only if projected population totals and age-sex populations are output separately)*

- Sum the projected population by sex and age group.
- Ensure that they match projected total populations.

(b) *Are all projections by age and sex free of negative values?*

- Check all age-sex projections for negatives. This may seem like an unnecessary check, but unfortunately some projection models are capable of generating negatives (e.g. cohort-component models using net migration numbers).

(c) *Does the projected population age structure change in line with the projected net migration age pattern and cohort flow from earlier age groups?*

- Check that the projected age structure is consistent with migration age profile assumptions. For example, if an area is projected to have net migration losses in the young adult ages and net migration gains at older adult ages then the projected age profile would be expected to maintain an indentation in the young adult ages and experience significant growth in the older age groups.
- For smaller populations/those with high migration rates the age structure will be shaped largely by migration. For larger populations/those with low migration rates cohort flow from younger ages will be a greater influence in the development of the age structure over time.

(d) *Does the projected age structure change in line with past population age structures?*

- Create population pyramids or similar graphs illustrating past and future population age structures (e.g. 20 years before jump-off, 10 years before, the jump-off year, and projected populations in 10 year intervals)
- Often there is only gradual change, or even continuity, in projected local and regional population age structures over time. This is especially case in the childhood and younger adult ages at the local scale where migration is substantial and age profiles are often fairly stable. In the upper half of the population pyramid age profile changes are common because of cohort flow and declining mortality rates.
- However, if age-specific net migration patterns have changed markedly, some changes will occur. This is more likely for small areas undergoing major socio-economic changes (e.g. new immigrant groups moving in; new student accommodation blocks, etc.)

(e) *Does the relative size of the infant age group align with fertility assumptions?*

- Check that the projected number of 0 or 0-4 year olds makes sense given the Total Fertility Rate and numbers of women of childbearing age.
- If the TFR is significantly above replacement then you would expect the infant age group to be fairly large relative to the female childbearing age population. If the TFR is low then the number of infants should be small relative to the female childbearing age population. Some allowance will have to be made for net migration.

(f) *Are peaks and troughs in the age structure related to non-private dwelling populations maintained in the projections?*

- Most non-private dwelling populations (e.g. boarding schools, prisons, nurses' accommodation, and aged care homes) have an unchanging or only slowly changing age structure, so their shapes can be expected to be mostly maintained in the overall population age structure over time.
- Check that these institutional populations do not shift up the population pyramid over time.

(g) *If an area is projected to grow rapidly, is most of the growth in the peak in-migration ages?*

- Rapid growth is usually the result of large net in-migration, and so you would expect the bulk of growth to be in the peak in-migration age groups.

(h) *Do sex ratios by age change gradually over time?*

- Graph past and projected sex ratios by age.
- Births, and mortality and migration numbers by age, are highly correlated by sex, so there should normally be only gradual change in projected population sex ratios by age into the future (unless projection assumptions have been deliberately selected to achieve this).

- In the context of male-female mortality convergence, sex ratios should gradually rise over time at the oldest ages.
- In some areas and at specific ages, migration has a long-established sex ratio which deviates noticeably from 1.0. Unless a change in the causes of these patterns is projected, the peaks and troughs in the sex ratio age profile should be maintained.

(i) *Are sex ratios for the infant age group close to the sex ratio at birth?*
- Unless there is huge differential in migration and mortality by sex, the sex ratio for the youngest age group should be close to the assumed sex ratio at birth.
- As a consequence the projected numbers of male and female young children should be fairly similar and reflective of this sex ratio.

4. Population change across areas

Comparing population projections across areas is useful in spotting inconsistencies and unusual projection growth patterns in individual areas that may otherwise be missed. Such spatial inconsistencies may also be difficult to explain to users, thus reducing confidence in the projections.

(a) *Do projected populations across all areas sum to projections for higher geographies? (Applicable only where more than one geographical scale of projections is produced)*
- Sum projections for lower levels of geography to the larger regions at the higher level of geography for both total and age-sex populations.
- Check to ensure the two sets of projections are consistent.

(b) *Do projected demographic components of change across all areas sum to projections for higher geographies?*
- Sum projected births, deaths, immigration, emigration, and net internal migration for lower levels of geography to the larger regions at the higher level of geography.
- Check that these components are consistent.

(c) *Does projected net internal migration across the whole country sum to zero?*
- Sum net internal migration across all subnational areas.
- A basic rule of population accounting holds that net internal migration must sum across internal areas to zero.

(d) *Do similar types of area (e.g. in terms of size, population density, remoteness classification, etc.) have similar projected growth?*
- Create scatter plots of projected growth rates against jump-off year population size, population density, etc.

- Areas that fall well outside the general pattern will need closer inspection to determine whether their projections are realistic. Outliers might be reasonable, but there will need to be credible reasons for why their projections differ from other similar areas. Some stakeholders (e.g. the mayor of an area with lower projected growth than a council with similar socio-economic characteristics) may ask the question.
- Create graphs of total population and age-specific populations for areas sharing similar growth trajectories in recent decades. Do they have similar projected growth rates? If not, are there good reasons for the differences?

(e) *Do projected populations aggregated up to states/capital/balance of state areas give results fairly similar to ABS or other projections?*

- Graph projections against those produced by ABS or other respected projection producers.
- Assess the degree of difference between the projections and those of ABS. If there are noticeable differences, are there good reasons for them? This is not to imply that ABS projections are superior (often they are not) but users will probably ask the question.

Summary

The checklist questions are reproduced in Table 1 below.

Table 1: A checklist for reviewing draft population projections

Projection outputs	Questions
1. Total population	(a) Does the projected total population seem plausible in the light of past trends? (b) If projected growth is very high, is it credible? (c) Is the projected total population similar to previous projections for this area, and those produced by other reputable forecasters? (d) Does the total projected population decline by a substantial amount?
2. Demographic components of change	(a) Do the demographic components of change match total population change for each projection interval? (b) Do the projected demographic components of change seem plausible in the light of past trends? (c) Do the projected births and deaths seem plausible in the light of projected age structure changes? (d) Does the total net migration assumption (internal + overseas migration) in broad terms align with what is known about the local economic prospects of the area?
3. Age-sex structure	(a) Do the projections by age and sex sum to the output total population projections? (b) Are all projections by age and sex free of negative values? (c) Does the projected population age structure change in line with the projected net migration age pattern and cohort flow from earlier age groups? (d) Does the projected age structure change in line with past population age structures? (e) Does the relative size of the infant age group align with fertility assumptions? (f) Are peaks and troughs in the age structure related to non-private dwelling populations maintained in the projections? (g) If an area is projected to grow rapidly, is most of the growth in the peak in-migration ages? (h) Do sex ratios by age change gradually over time? (i) Are sex ratios for the infant age group close to the sex ratio at birth?
4. Population change across areas	(a) Do projected populations across all areas sum to projections for higher geographies? (b) Do projected demographic components of change across all areas sum to projections for higher geographies? (c) Does projected net internal migration across the whole country sum to zero? (d) Do similar types of area (e.g. in terms of size, population density, and remoteness classification) have similar projected growth rates? (e) Do projected populations aggregated up to states/capital/balance of state areas give results fairly similar to ABS or other projections?

It should be possible to automate some of the checks listed in this paper in a spreadsheet or in a dedicated subroutine in a population projection program. This should prove particularly useful if there are hundreds, or even thousands, of local areas in a set of projections. However, some types of check will still require examination and the judgement of the forecaster.

What constitutes a plausible projection will sometimes be difficult to judge, especially if major demographic changes are anticipated and/or the areas are small and subject to volatile demographic trends. In such cases it may prove useful to generate alternative projections using different types of projection model, or use the current projection program and undertake a short-term test projection from 5 or 10 years prior to the jump-off date to assess how well the model reproduces the jump-off population.

Reviewing draft projections may take up time in the projections production process. But it is far better to discover errors or implausible projection outcomes during this process rather than be alerted to problems by users after publication!

References

Murdock S H, Hamm R R, Voss P R, Fannin D, & Pecotte B (1991) Evaluating small area population projections. *Journal of the American Planning Association* 57.4: 432-443.

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

United Nations Population Division (2000) *Replacement Migration: Is it A Solution to Declining and Ageing Populations?*. New York: UN.

Wilson T & Rees P (2005) Recent developments in population projection methodology: a review. *Population, Space and Place* 11.5: 337-360.

Wilson T (2014a) Simplifying local area population and household projections with POPART. In Hoque N and Potter L (eds) *Emerging Techniques in Applied Demography*. Dordrecht: Springer; 25-38.

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