The University of Manchester

MANCHESTER 1824

The Cathie Marsh Centre for Census and Survey Research

Estimation of local demographic variation in a flexible framework for population projections

CCSR Working Paper 2010-05
Ludi Simpson and Harvey Snowling
ludi.simpson@manchester.ac.uk

An experiment with small areas of Fife in Scotland shows that persuasive local population projections based on standard data for standard areas are feasible without the regular publication of migration flows. Three approaches are compared: (a) direct estimation of local area age-specific schedules of fertility, mortality and migration based on data available to the national statistical agency; (b) adaptation of national schedules using only local area population estimates by age, total numbers of births, and total number of deaths: age-specific migration is indirectly estimated from successive population estimates; (c) use of the same rates for each small area.

Estimation of local demographic variation in a flexible framework for population projections

Ludi Simpson (University of Manchester), Harvey Snowling (General Register Office for Scotland)

Background – and the problem

Demand for projections of the detailed age structure of the population in sub-national areas arises from demands on government, and by government on other organisations, to plan meticulously for the needs of all populations. As demographic microdata from administrative records are increasingly available to government through new technologies, the agencies responsible for local services expect demographic projections to be available for all the boundaries relevant to service planning and delivery.

In Scotland as elsewhere in the UK, local authorities are responsible for most local planning. Each service prefers projections for the geographical areas relevant to them: school catchment areas, the boundaries of each major housing scheme, and larger areas for fire service planning, for example. However, the data held by statistical agencies do not support robust estimates of fertility, mortality and migration for very small areas. Therefore it is considered not possible to produce a robust standard set of projections for every small area, that could be aggregated by users to their areas of interest. Release of the vital statistics and migration data with age detail to allow local authority and other independent researchers to produce their own demographic estimates and projections risks disclosure of sensitive material for individuals.

Migration poses a particular problem in this respect. While births and deaths can be released by a statistical agency for standard small areas and re-aggregated by users to areas appropriate to services, migration flows cannot be aggregated in the same way. Knowledge of the migration between each pair of small areas would be required, which would involve very small flows indeed.

The motivation for this paper is the possibility of independent projections supported by regular release of government data for standard areas, without loss of flexibility for users, without loss of confidentiality for the public, and without loss of plausibility in the projections. Plausibility here will entail a demonstration that distinct local demographic characteristics can be measured.

The cohort component model has no rival in providing projections of age structure, though simpler methods perform well for projections of the total population (Rowland, 2003; Siegel, 2002; Smith et. al. 2001). Its implementation within a single year of age and annual time periods provide the framework for this investigation, using the POPGROUP software (Andelin and Simpson, 2007; Simpson, 2006).

Various authors have specified and evaluated methods to graduate or calibrate reference schedules of mortality, fertility and migration to take into account local characteristics. Age schedules may be parameterised and their local parameters estimated as clearly reviewed and demonstrated by Congdon (1992). Multilevel modelling may distinguish local variability from sample variability and provide model-based confidence intervals (Johnson et al. 2010; Leyland and McLeod 1999; Westerling 1995). Migration attracts a variety of approaches for sub-national projection (reviewed by Wilson and Rees, 2005).

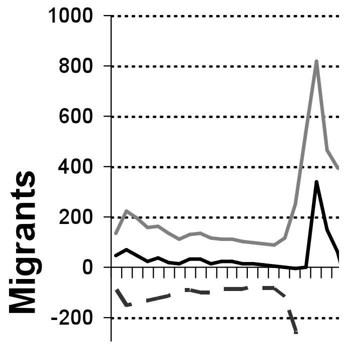
In this paper we test a practical approach to calibrate each component of change. National age schedules of fertility and mortality are scaled to reflect the local levels of fertility and mortality, maintaining a common age-sex structure. For migration, we calibrate a national age-sex schedule to match a time series of estimates of the local population age structure. This indirect estimation of local age-specific migration is the main methodological contribution of the paper. It is made possible by the publication in England, Wales and Scotland since 2001 of annual age-structured population estimates for small areas. The dependence of demographic methods on data availability will be discussed further later in the paper.

We first set the context of Fife in Scotland for which we present small area projections, and the data available. The methods are then described not only for the calibrated estimates of local components of change just referred to, but for two alternatives: on the one hand estimating demographic schedules directly from local data without recourse to calibration, and on the other ignoring local characteristics altogether. Comparison is made of the three sets of estimates, and of small area projections based on them. At the end of the paper we discuss extensions of the calibration approach to make the most of available data, and to project households and other characteristics related to age and sex.

Context and data

Fife lies to the north-east of Edinburgh in Scotland. Its 360 thousand population is divided between the major towns of Dunfermline, Glenrothes and Kirkcaldy which expanded through housing developments in the past decade, several smaller towns, and rural areas. Three quarters of its boundary is water, the firths (rivers) of Tay and Forth and a coastline which hosts tourism, sport and the University of St Andrews. Public service planning divides Fife into sub-areas in a variety of ways. Here we use the 23 ‘multi-member ward’ electoral areas, henceforth called ‘wards’ (Figure 1) but our larger aim is to find a method that will apply to any set of local boundaries.

Fife’s variety provides a good test bed for methods of local projection. Figure 2 shows the estimated migrants to and from Fife as a whole, taken from GROS (2009, which also describes the sources of data used to measure both within-Scotland and international migration). The twin peaks of out-migration illustrate the different demographic experiences within Fife. One younger out-flow reflects mainly indigenous young adults leaving the parental homes throughout Fife for work and study outside Fife, while the slightly older out-flow reflects mainly graduating university students located specifically in St Andrews. Estimating an equivalent of Figure 2 for each ward of Fife is one of the challenges for local area projections.



The data used in this paper are those available to the national statistics agency, the General Register Office for Scotland (GROS), and thus are potentially more widely available through release to local and other planning organisations. GROS is responsible for the registration of births and deaths, and the estimation of migration and population for Scotland and its Council areas of which Fife is one. Since 2001, GROS has used the boundaries of 'data zones' for standard data releases relating to births, deaths and annual estimates of the age-sex structure of local population. Data zones have an average population in Scotland of 795. Fife's 453 data zones have an average population of 799 (estimates for mid-2008).

GROS derives data zone small area population estimates by annually updating Census-based population age-sex structures, 'ageing on' the population and applying information on births, deaths and migration. Births and deaths are straightforwardly aggregated from individual records which include a detailed geographical reference. Internal migration into and out of each data zone is estimated by GROS using health service re-registrations, interpreted as moves between the registered residential addresses of the patients, based on the universal National Health Service in the UK. International migration to and from data zones is less straightforwardly estimated (GROS, 2010). However, these estimates of migration flows which are available for each data zone cannot be aggregated to give flows for other areas of interest to local planning. They would require deduction of inter-zone flows, most of which are so small that their release would risk disclosure of personal data.

The standard data that are available for each data zone are:

- Counts of births and deaths for each mid-year annual period from mid-2001/2002,
- Annual population estimates by single year of age and sex from mid-2001.

These data were aggregated to the 23 wards of Fife using a conversion table from data zones, containing proportional allocations where a data zone crossed the boundaries of two or more wards.

Estimation methods

Of interest is the impact on projections of alternative estimates of the recent level and age-sex structure of fertility, mortality and migration in each area. In all the projections presented below the future trend of fertility and mortality rates is taken from the projections for Scotland as a whole prepared by the Office for National Statistics (Bray, 2008). The *changes* in these rates specific to each age and sex are assumed to be the same for each area. Future changes in migration are also assumed to be similar for each area in the ways described below. Interest focuses on the local *level and age-structure* of fertility, mortality and migration, which vary between each projection and between areas.

The projections are each a set of single-area cohort component projections with single year of age and annual cycles, carried out using the POPGROUP software. The software demands entry of single year of age schedules of cohort rates of fertility, mortality and transition migration in- and out- flows with one or two external areas. Rates are expressed as events per thousand population specific to males or females with age at the start of the projection year.

The software allows optional constraints as (a) numbers of births, deaths and migrants in any of the flows, and (b) population (Simpson, 2006). These constraints may be totals or age and sex specific (only by sex for births) and may be specified in any year of the projection. A constraint specifying a number of births, deaths or migrants is implemented by scaling the initial projections of that component, and recalculating demographic indicators after scaling. A population constraint is implemented by adjusting the migration flows initially estimated through the rates and migration constraints; the adjustment is shared between flows according to user-provided weights. The POPGROUP software is based on Microsoft Excel worksheets aided by VBA routines; its options provide a flexible framework for a variety of projection strategies, documentation of its algorithms is available and may be implemented on other platforms.

The Fife ward projections each begin in 2001 with six annual periods from mid-2001 to mid-2007 for which start and end population, births and deaths are recorded or estimated by GROS. Migration at each age and sex during this initial period is unknown and is indirectly estimated by the software as the residual after births and deaths are deducted from population change. Thus the model estimates the net migration to each age-sex cohort, but it adjusts gross flows that are initially based on standard age-profiles in order to estimate in- and out- flows of migration consistent with the changing population. The projection into the future begins after mid-2007. Each projection continues until 2026, a target year of importance to local planning procedures, providing a forward projection of nineteen years.

Three estimates of local area demographic rates are implemented and compared. This section provides a summary of each estimation method. More detail is provided in later sections where relevant.

(a) The *No Local Variation* projection assumes the same fertility, mortality and migration rates in each ward. Fertility and mortality age-sex schedules are based on the experience of Fife as a whole, as used in GROS' 2006-based projections (GROS 2008). Migration schedules are the same for each ward. Fife's total of births, deaths and age-sex structured population as projected by GROS are set as constraints. Thus the annual net impact of migration at each age on Fife's population is allocated to each ward proportional to population size. The resulting age-sex schedules of in-migration and out-migration rates are the same for each ward, and vary little over the projection.

(b) The *Local Calibration* projection estimates local age-sex schedules of fertility, mortality and migration indirectly.

A 'training projection' is run from 2001 to 2007 in which the *no-variation* assumptions are augmented by the standard data recorded as described above (births, deaths and age-sex-structured population). Included in the output from this projection are estimates of each demographic component for each annual period mid-2001/2002 (Andelin and Simpson 2007):

- Total Fertility Rate: local TFR = Fife TFR x local births recorded / (births expected from Fife fertility rates and local starting population structure).
- Standardised Mortality Ratio (SMR) = Fife SMR x local deaths recorded / (deaths expected from Fife mortality rates and local starting population structure).
- Migration flow at age-sex group *a*: Migration expected from Fife migration rates and local starting population after local births and deaths + (end population estimate recorded for age-sex group *a* – end population estimate expected for age-sex group *a*). The correction referred to within brackets is allocated to in- and out-migration flows according to user-provided weights, in this case 50% for each of in-flow and out-flow.

In these ways, the training projection provides estimates of the TFR, SMR and single year of age-sex migration in- and out-flows, for each ward and each annual period. For each of these components, the subsequent *local calibration projection* used an average of the estimates for the six years mid-2001 to mid-2007 for each future year. Thus for fertility and mortality, reference rates were adjusted up and down by a single local factor. A local set of single year migration rates replaced the standard one, for each of in- and out-flows to one external area.

(c) The *Local Direct Estimates* projection uses local age-sex schedules of fertility, mortality and migration estimated from the detailed data available to GROS. These are based on mid-2001 to mid-2007, for comparability with the local calibration projection. For migration, age-sex schedules of in- and out-flows were estimated separately for movement within Fife and for all other movement. In each case the direct estimates were smoothed by averaging three adjacent age groups, except that migration rates for single ages were smoothed only when the adjacent age groups have fewer than 200 events, to avoid dampening extreme but realistic values of young adult migration.

Evaluation

The three projections specified are successively more complex to prepare. The *no-variation* projection assumes implausibly equal demographic characteristics in each locality. The *calibration* projection attempts a practical approach which respects local variation in demographic characteristics. It enables projections for any local area using standard data for standard data zones, because it requires only data that can be straightforwardly aggregated to non-standard areas, but it is also an approximation. The *direct* projection is methodologically most sound, using full detail of local data to estimate each component. However it is less practical to provide such detailed data for all types of locality and risks both disclosure of information about individuals, and poor quality estimates due to sparse data.

Our evaluation aims to measure the extent to which the calibration approach as implemented does capture local variations in fertility, mortality and migration. For this reason each projection shown here is constrained to the same total number of future births, deaths and age-sex-structured population projected by GROS for Fife as a whole. The differences between projections are thus only those relating to wards. Separate comparison of the three projections without constraint to the Fife totals showed slightly more discrepancies but the same patterns of difference that are reported here, since the impact of the constraints was not large.

We report the local characteristics as estimated by each approach, and judge the plausibility of the results for the calibration approach from our prior understanding of the areas involved and similarity with the direct estimates. We then compare how close each of the no-variation and calibration approaches is to the direct approach, when the population is projected 19 years ahead. We examine the mean absolute deviation between the no-variation and the direct projections across the 23 wards, for each of births, deaths, and population. When calculated again between the calibration and direct projections, success would involve much lower mean deviations, to indicate that the calibration projection captures variation between wards.

In order to isolate the separate impact of estimates of fertility, mortality and migration, we compare projections with different estimates of one component when all other aspects of a projection remain constant.

Estimates of population

All projections incorporate the 2001 and 2007 population estimates, illustrated for the entire Fife area and three of the wards in Figure 3 to further illustrate the variety of local circumstances. Fife's pyramid shows a growing area with ageing, the postwar baby boom reaching age 60 in 2007. There are reductions among teenagers and those in their thirties showing an overall loss of families through migration. Dunfermline South is one of the larger, mainly urban, wards that has been growing steadily at most ages following substantial housing developments. Lochgelly & Cardenden contains two small towns and is a less populous and more rural area; it shows the variability between neighbouring ages which is characteristic of small populations. Finally, St Andrews is dominated by students in higher education who mostly arrive in their late teens from other parts of Scotland and further afield, and leave a few years later.

The population estimate for mid-2001 is based largely on the Census of that year and can be considered more accurate than the 2007 population estimate which updates the census using registered births and deaths, and estimates of migration based on health patient re-registrations. The latter appear to capture a steady inflow of students to St Andrews but not to fully capture their outflow, leading to the increase in adults in their late twenties apparent in Figure 3. This possible flaw in the quality of population estimates helps to understand some of the later results for this unusual ward.

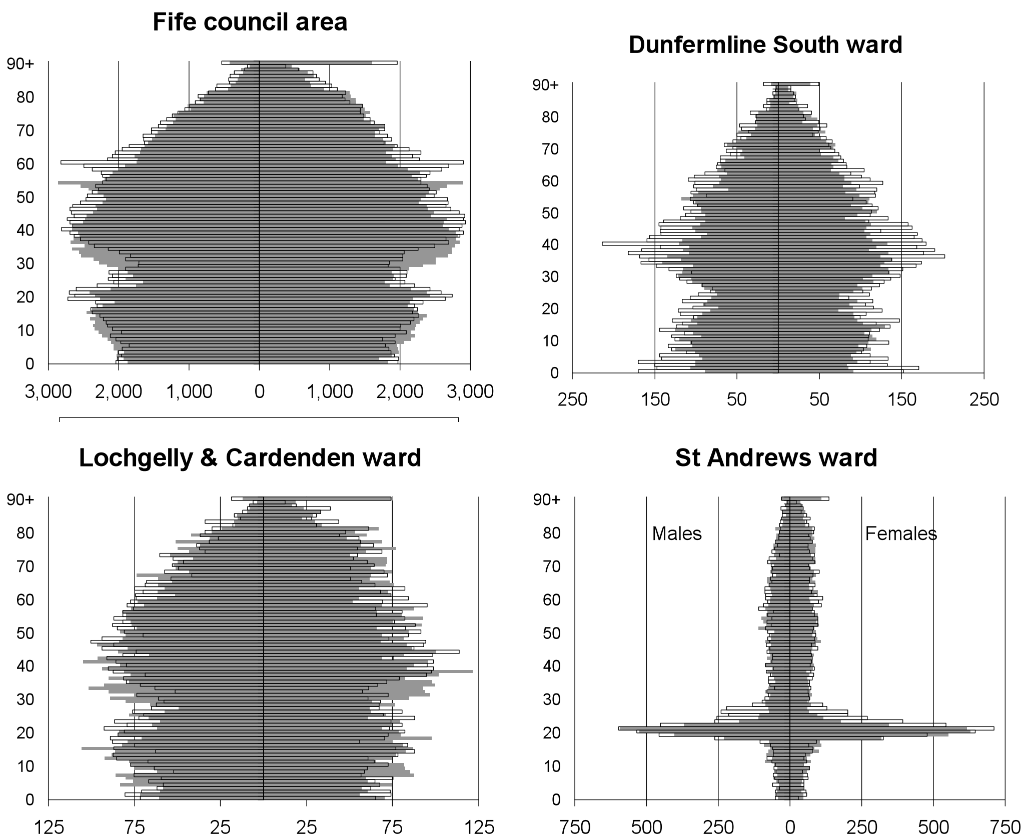


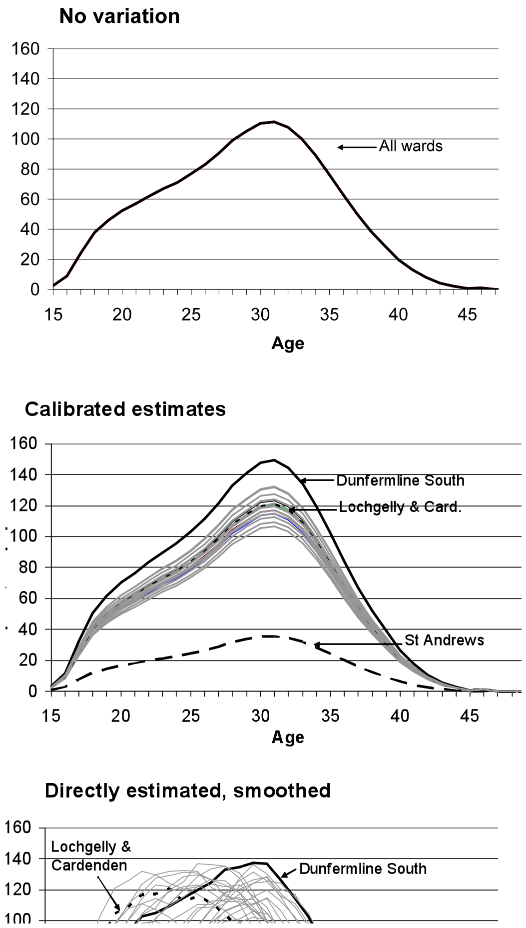
Figure 3. Population estimates 2001 and 2007, Fife Council Area and selected wards

Number of people (horizontal axis) at each single year of age (vertical axis). 2001 shown in grey; 2007 shown in white.

The proportional allocation of population (and vital events) from data zones to the areas of interest is an important part of the methodology, as it allows the standard data to be used for any area of interest. Comparisons with the population of each ward estimated by allocation of whole data zones showed differences of up to 10% for neighbouring wards where data zones cross their boundaries. This estimation of indicators for non-standard areas from an aggregation of standard areal units is a common challenge, which is generally made more accurate by smaller standard units and by proportional allocation based on information closely related to the indicators. In this case the proportions were based on Fife Council's database of residential property which one would expect to be closely related to the location of population (Simpson and Yu, 2003).

Estimates of fertility and projection of births

Figure 4 clearly illustrates the differences between the three estimates of fertility schedules. The calibrated estimates stretch the overall fertility schedule estimated for Fife up or down, using the same multiplier at each age according to the number of births recorded in the area during mid-2001 to mid-2007, taking into account the population age structure of women in that ward. The calibrated estimates clearly identify areas with fertility different from the average, for example the very low fertility in St Andrews due to the substantial student population, and the higher



fertility of Dunfermline South. However, these calibrated estimates maintain the same relative fertility between each age, which does not reflect reality as estimated directly from registered births in the direct estimates. The direct estimates shown in Figure 4 are smoothed across three adjacent years of mother's age at birth to avoid very volatile schedules. The younger and older fertility schedules respectively of Lochgelly & Cardenden and Dunfermline South are evident, as well as the bimodal age structure of some wards' fertility. The calibrated estimate correctly estimates low fertility for St Andrews but overstates the fertility of its young adult women and understates the fertility of its older women.

The calibrated estimates reproduce the correct number of births in each ward in recent years equally accurately as do the direct estimates, since that is how they were calculated. However, if the age structure of the population changes, one would expect the number of births to be better projected by the direct estimates. For example, if Lochgelly & Cardenden or St Andrews were to have more young women in their late teens and early twenties, the calibrated estimates would under-estimate the births in Lochgelly & Cardenden and over-estimate in St Andrews, because its standard age-structure would not capture the unusual fertility at those ages in the two wards.

In order to focus on the impact of the fertility estimates alone, three projections are reported which differ only in that component; mortality and migration use the calibrated estimates. The number of births for Fife as a whole is stable at 100 thousand in each projection, but the number of births in each area varies considerably between the projections. As expected, the projection that assumes the same fertility curve for each area is often far from the other projections, averaging a 9.4% deviation from the projection using fertility schedules estimated directly for each area (Table 1). The calibrated estimates, using a different level of fertility for each area but the same age-pattern, is closer to the direct projection of births, averaging a distance of 2.3%. The full data shows that the calibrated projection of births is less than 2.5% from the direct projection of births for every ward except two. In Dunfermline South the calibrated projection is 6.4% above the direct projection of births, while in the university ward of St Andrews the calibrated projection is 19.9% below the direct projection of births.

	No Variation	Calibrated
Mean absolute deviation		
All 23 wards:	411 (9.4%)	99 (2.3%)
Deviation for single wards		
Dunfermline South	-1347 (14.3%)	+600 (6.4%)
Lochgelly & Cardenden	-103 (3.2%)	+18 (0.6%)
St Andrews	+4190 (125.1%)	-668 (19.9%)

Table 1. Projected births: deviation from the direct projection

Percentages are of the (mean) area births during 2001-2026.

However, the directly estimated curves can only provide an accurate projection if the population age structure is also accurately projected. Earlier it was apparent that the age structure of St Andrews is wrongly aged in official estimates for 2001-2007. The calibrated projection continues this false ageing on because its estimates of migration are based on the recent population estimates. The extra women projected in their twenties and thirties, who in reality would be ex-students who move away, create a high projected number of births when using the direct estimates because these emphasise fertility in the older ages (Figure 4). Figure 5 shows this clearly in the unrealistically high number of projected births for St Andrews from the direct estimates of fertility.

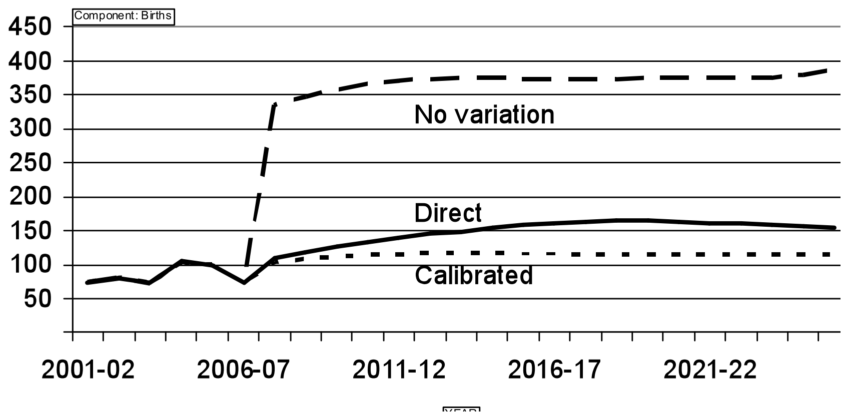
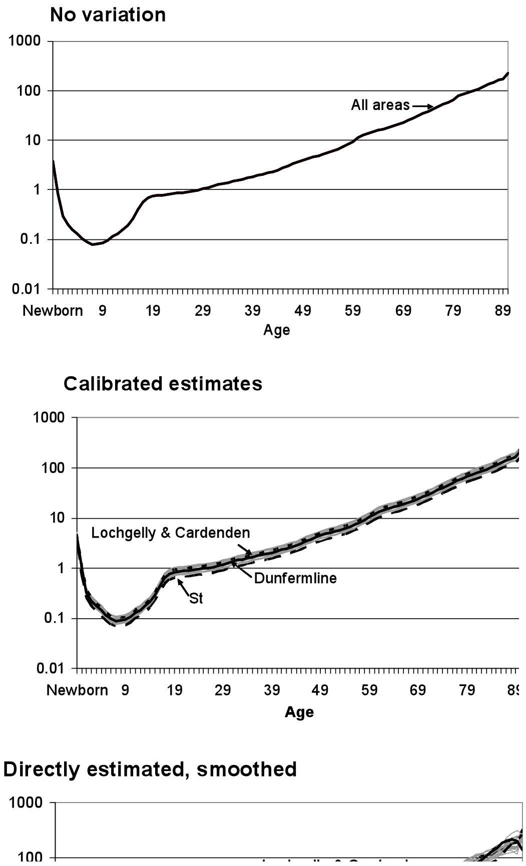


Figure 5. Births for St Andrews ward, three projections
Each projection differs only in the fertility schedules used. Mortality and migration use the calibrated estimates.

Estimates of mortality and projection of deaths

The three estimates of mortality schedules (Figure 6) show how the calibration captures the level of mortality in each area and the direct estimates capture the age-specific variation in area mortality. The direct estimates' variability is partly due to the small populations that each age-specific estimate is based on, which is very evident for the younger ages. There are few discernible differences in mortality age patterns between areas. The impact on population size will be greatest at older ages where mortality is higher, but it is here that population size for any particular age is small and itself estimated with error. Figure 6 uses smoothed mortality rates for the direct estimates, averaged from three neighbouring single years of age.



Three projections are reported which differ only in mortality; each uses the calibrated estimates of fertility and migration. The total number of deaths projected for Fife is about 95 thousand. While the number of deaths in each area varies between the projections, the variation is less than for births. The projection that assumes the same mortality curve for each area is on average 4.1% from the projection using fertility schedules estimated directly for each area. The projection using calibrated estimates, using a different level of mortality for each area but the same age-pattern, are much closer to the direct projection of births. Every ward is less than 2.0% away from the direct projection of deaths for every ward (not shown in the table) and only 0.6% away from it on average. The university ward of St Andrews is not an outlier.

	No Variation	Calibrated
Mean absolute deviation		
All 23 wards	167 (4.1%)	25 (0.6%)
Deviation for single wards		
Dunfermline South	+8 (0.2%)	-7 (0.2%)
Lochgelly & Cardenden	-133 (3.8%)	-3 (0.1%)
St Andrews	+475 (13.9%)	-10 (0.3%)

Table 2. Projected deaths: deviation from the direct projection

Percentages are of the (mean) area deaths during 2001-2026.

Estimates of migration and projection of migrants

The three sets of migration estimates are constructed in different ways as described above, using the software's flexibility to include not only schedules of migration rates specific to age, sex and area, separately for in- and out-flows from one or two external regions, but also constraining migration counts and population targets.

The no-variation projection uses a single national schedule of migration rates drawn from the 2001 Census for both in- and out-flows, and provides the projected single year of age population of Fife as a whole as a constraint. The migration rates are the same for each area.

The calibrated projection derives age-sex counts of migrants from a training projection for the annual periods mid-2001 to mid-2007, using the GROS population estimates for each local area as constraints. These indirect estimates of annual migrant flows by age and sex are consistent with the estimated population in that period. The calibration is thus entirely governed by the quality of the population estimates in the training period. After mid-2007, the projection uses the mean number of migrants estimated for mid-2001 to mid-2007 as a constraint for quinary age-sex groups.

The direct estimates of migration are extracted from a variety of sources (GROS, 2010). Two sets of health service administrative records provide information based on patient re-registrations within Scotland and with the rest of UK. Immigration from outside the UK is measured by the International Passenger Survey which measures both in and out flows and is supplement by health service records. Several approaches to measuring rates for each local area were attempted, limited by the data available and the very low numbers of international migrants for most local areas. The most robust set of age schedules specific to sex and ward were based on moves within Fife and moves with the rest of the world, smoothed across three adjacent ages. The smoothed estimate only replaces the unsmoothed estimate when based on fewer than 200 migrants, to avoid dampening the extreme migration rates.

Figure 7 shows the three estimates; the direct estimates for Fife and outside Fife are added for comparison with the other projections which use only one external area. Both the calibrated and the direct estimates capture similar differences in migration patterns which are ignored by the single overall schedule, including the inflow of students to St Andrews aged 17. However there are significant discrepancies that are caused by the different ways in which the estimates are created; high in-migration to Dunfermline South at most ages is evident in the calibrated estimates but not in the direct estimates, and the peak in-migration to St Andrews at age 17 is higher in the calibrated estimates.

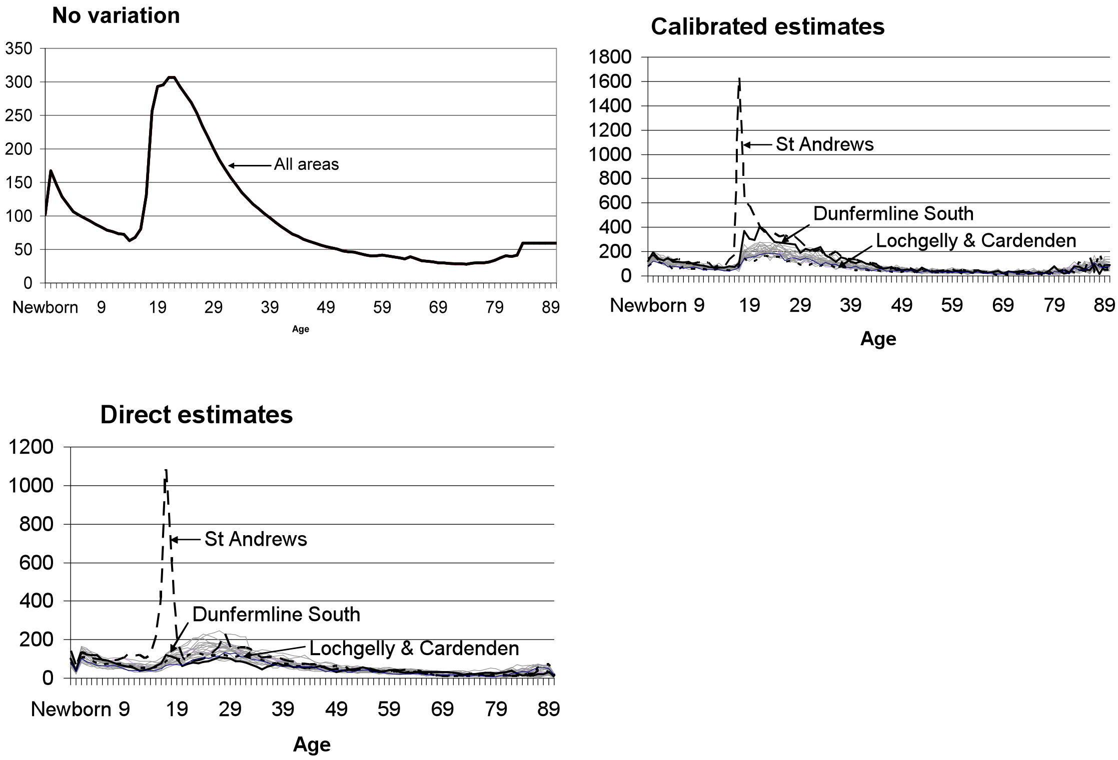


Figure 7. Male in-migration schedule, three estimates
Migration rate, migrants per 1,000 males by single year of age. NB. Each chart uses a different vertical

Unlike fertility and mortality, the calibration of migration is age-specific with the single year of age detail contained in the population estimates provided for 2001-2007. However, the population estimates determine only the net impact of migration; the division into in and out-flows is based on the initial migration schedules from the previous census, using an assumption that any discrepancy is resolved evenly between in and out flows. For this reason, and because the impact on a projection of migration, however estimated, is a *net* impact, the most relevant comparison between the three estimates is of net migration, which is shown in Figure 8 for St Andrews ward. The net estimated migration 2007-08 is expressed as a percentage of the population of the relevant age group at the start of the period.

The *no variation* estimate is clearly wrong, capturing neither the extreme in-migration at age 17 nor the out-migration of residents aged in their twenties, both identified by the direct and calibrated estimates. However, the calibrated estimate has in this ward estimated a much larger in-migration at age 17 than the direct estimate. The reason for the difference results from the different sources of migration, but may

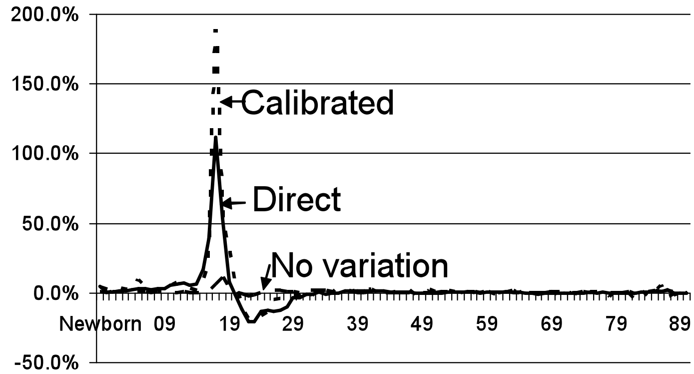


Figure 8. Net migration, three estimates for 2007-08, St Andrews ward
In-migration minus out-migration, as a percentage of mid-2007 population; age at start of year.

be also partly a difference in age-time plan of the projection model and the migration data. The impact on the projection is magnified in each year of the projection as the in-migrants at age 17 become older and are subject to further in-migration, as this is specified as proportions of the current population.

Migration adds about 50 thousand people to Fife's population over the projection period. On average of the 23 wards, the calibrated estimate of migration across the projection period is 3.8% away from the direct projection, less than one third the deviation of the projection with no variation between wards (Table 3). The distance between the calibrated and direct projections is particularly great in St Andrews (14%).

Not surprisingly, the distance between the projections is greater when age is considered. The calibrated projection is on average 11.5% away from the direct projection of an age-group's net migration, compared to 3.8% for the total. For Dunfermline and St Andrews, age group estimates are on average 16% and 40% from the direct projection, but while large these discrepancies are very much less than for the projection with no variation between wards.

15-19 year olds in St Andrews provide the largest discrepancy. The projection with calibrated estimates projects an addition of 25 thousand to the 15-19 year old through migration over the period, to a cohort that remains between one and two thousand as students age out of that age group and are replenished by newcomers. The projection with direct estimates indicates an addition of 18 thousand. Out-migration at ages 20-24 remove almost all the in-migrating students.

(a) Net migration, summed across all ages	No Variation	Calibrated
Mean absolute deviation		
All 23 wards	1,975 (12.6%)	602 (3.8%)
Deviation for single wards		
Dunfermline South	+12746 (68.7%)	-1194 (6.4%)
Lochgelly & Cardenden	1045 (8.6%)	-357 (2.9%)
St Andrews	-748 (4.1%)	+2593 (14.1%)

(b) Net migration, quinary ages	No Variation	Calibrated
Mean absolute deviation		
All 23 wards	242 (29.4%)	94 (11.5%)
Deviation for single wards, mean absolute deviation across age groups		
Dunfermline South	736 (75.4%)	153 (15.6%)
Lochgelly & Cardenden	88 (13.9%)	61 (9.5%)
St Andrews	1442 (148.9%)	389 (40.2%)

Table 3. Projected net migration, deviation from the direct projection

Percentages calculated from the (mean) population in mid-2007.

Discussion

Here we summarise what has been achieved, and discuss extensions to the calibrated small area projections.

A practical calibration approach has been described which provides useful local indicators of demographic characteristics without the requirement of direct statistics of migration. A training projection for a recent period integrates demographic schedules for a reference area with the local number of births and deaths and population estimates. This training projection provides a summary view of demographic differences between local areas by estimating the TFR, SMR and age-structured migration that are consistent with the local constraints. Used in a projection, the results are plausible and much closer to a projection based on direct estimates of demographic schedules for each local area, than a projection which assumes the same local schedule for every area. The approach provides age-structured population estimates that have high value in planning local services. Avoidance of direct migration statistics makes the approach suited to projections for any local area by aggregating standard data for standard areas.

Direct estimates of demographic schedules are not usually available for local areas due to the political sensitivity of information based on few events, but tend also to be poor estimates of underlying demographic characteristics again because they are based on few events. The calibrated estimates smooth the schedules. The paper has shown that this smoothing loses the detail of fertility in particular, not distinguishing early, late or bimodal fertility; however the projected number of births is nonetheless similar to projections based on direct estimates of schedules. More generally, discrepancies between projections based on calibrated and direct estimates should be expected only when the population age structure is changing. In some circumstances the calibrated estimates probably perform better than a direct estimate. Direct estimates are often based on small numbers of past events or small populations (especially the very elderly) and so are not necessarily the most plausible for projections.

As a result of this investigation, the statistical agency in Scotland has provided further standard datasets to local authorities so that they may make independent demographic projections [ref to be completed when published during summer 2010]. These datasets for standard small areas (data zones in Scotland) provide a time series of vital events and population estimates that are intended to be aggregated to the area boundaries of most relevance to the changing planning and political priorities in each local authority. The local authorities have been provided with a rich dataset for demographic work, while the statistical agency avoids production of direct estimates for non-standard areas, which would be particularly onerous for migration, and avoids the thorny issues of confidentiality and statistical reliability associated with small areas. GROS will also provide age-specific fertility and mortality tables for data zones, summed across five year periods to avoid confidentiality concerns, so that local age-specific schedules may be calculated and used for more refined projections.

The calibration at the heart of this approach may be extended and adapted to data available. If local births and deaths are available with age detail of mother or deceased, then further detail of local characteristics can be estimated. In the POPGROUP software this estimation is currently automated for quinary age detail of deaths. The population estimates used in this example were available by sex and single year of age, and therefore provided indirect estimates of migration in this detail. Should population estimates be available by quinary age groups, or in total only, then the same approach (and this software) provides estimates of migration in correspondingly less detail.

The training projection provides a time series of demographic indicators for the recent past which might be of further use to projection work. While the simple mean of the recent past was used to distinguish localities in this paper, an analyst might study each local time trend and judge whether these should be reflected in the projection through assumptions about varying directions and rates of change in fertility, mortality or migration.

A major student population may be handled other than by a cohort component projection, as also other large populations such as armed forces that replenish themselves through migration to maintain their age structure without contributing much to births or deaths. This paper has shown how such extreme migration can be indirectly estimated but others may consider projecting these 'special populations' separately from the main cohort component projection (as for example in the projections for English local authorities, Office for National Statistics, 2010).

One consequence of using a single region model, albeit a set of 23 single-region models in tandem with constraints on their total, is that in-migration 'rates' are related to the destination population rather than the population at risk in the rest of the world.

It is standard practice to then specify flows of migrants rather than rates for future migration, to avoid implausible exponential change in population.

The software used in this work does not at present support a multi-regional demographic structure, which would allow the proper specification of in-migration rates in relation to populations at risk. Consideration of the use of multi-regional models for small areas would contend with greatly increased data demands, in exchange for an improvement to population projections which is debated because population change is in the end affected by the net impact of migration (Rogers 1990; Smith and Swanson 1997).

Finally, the production of projections for local areas with detail of age and sex is the platform for a number of derived forecasts for characteristics closely related to age and sex, which extend the utility of projections to strategic planning. Marshall and Simpson (2009) implement household and labour force projections for small areas in Britain. The POGPROUP software provides for such 'derived forecasts' of the labour force, households, disabled population or other characteristics, using user-defined age-sex plans suited to available data. Such projections require a local age-sex schedule of the characteristic (economic activity, headship rates, disability, and so on). These are often calibrated from reference schedules for larger areas constrained to local counts of varying detail, in a method equivalent to that used here for demographic schedules.

Acknowledgements

This paper is based on work financially supported by Fife Council and the General Register Office for Scotland.

References

Andelin, J. and L. Simpson (2007). POPGROUP: Population estimation and forecasting system, reference manual version 3.0. Manchester, University of Manchester. http://www.ccsr.ac.uk/popgroup/about/manuals.html. Accessed 7th July 2010.

Bray, H (ed) (2008). National population projections 2006-based, Office for National Statistics, Series PP2 No 26 http://www.statistics.gov.uk/statbase/Product.asp?vlnk=8519. Accessed 1st July 2010.

General Register Office for Scotland (2007). 2001-2004 Small Area Population Estimates – Methodology, http://www.gro-scotland.gov.uk/statistics/publications-and-data/population-estimates/special-area/sape/0104sape/0104sapemethodology.html. Accessed 1st July 2010.

General Register Office for Scotland (2008). Population Projections Scotland (2006-based) Population projections by sex, age and administrative area (revised 9 October 2008). Edinburgh, General Register Office for Scotland. http://www.gro-scotland.gov.uk/files2/stats/population-projections/population-projections-scotland-2006-based-population-projections-by-sex-age-and-administrative-area.pdf. Accessed 6th July 2010.

General Register Office for Scotland (2009). Fife and Forth Valley Migration Report, http://www.gro-scotland.gov.uk/files2/stats/migration/fife-forth-valley-migration-report.pdf. Accessed 1st July 2010.

General Register Office for Scotland (2010). Methodology used to produce GROS official migration estimates. http://www.gro-scotland.gov.uk/statistics/migration/methodology.html. Accessed 11th July 2010.

Johnson F.A., Chandra H., Brown J.J. and Padmadas S.S. (2010) District-level Estimates of Institutional Births in Ghana: Application of Small Area Estimation Technique Using Census and DHS Data, Journal of Official Statistics, 26(2): 341–359.

Leyland A.H. and McLeod A. (1999). Mortality in England and Wales, 1979-1992, Teaching resources and materials for social scientists, http://tramss.data-archive.ac.uk/documentation/MLwiN/mortalit.asp, (accessed 23 July 2010).

Marshall, A. and L. Simpson (2009). Population Sustainability in Rural Communities: the Case of Two British National Parks, Applied Spatial Analysis and Policy 2(1): 107-127.

Office for National Statistics (2010). 2008-based subnational population projections for England: methodology guide. Titchfield, Office for National Statistics Centre for Demography.

Rogers, A. (1990). Requiem for the net migrant, Geographical analysis, 22: 283-300.

Rowland, D.T (2003). Demographic methods and concepts, Oxford University Press, Oxford.

Siegel, J.S. (2002). Applied demography. Academic Press, San Diego.

Simpson, L (2006) Integrating estimates within population forecasts. Research Methods Programme Working Paper 18, University of Manchester. http://www.ccsr.ac.uk/methods/publications/documents/WP18.pdf. Accessed 1st July 2010.

Simpson, L. and A. Yu. (2003). Public access to conversion of data between geographies, with multiple look up tables derived from a postal directory, Computers, Environment and Urban Systems 27: 283-307.

Smith, S. K. and D. A. Swanson (1997). In defense of the net migrant, Journal of Economic and Social Measurement 24: 249-264.

Smith, S. K., J. Tayman and D. Swanson (2001). State and Local Population Projections: Methodology and Analysis. New York: Kluwer Academic/Plenum Publishers.

Westerling, R. (1995). Components of small area variation in death rates: a method applied to data from Sweden, Journal of Epidemiology and Community Health, 49, 214-221

Wilson T. and Rees P. (2005) Recent developments in population projection methodology: a review. Population, Space and Place, 11, 337-360.