J Pop Research (2011) 28:109–127
DOI 10.1007/s12546-011-9060-7

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

Ludi Simpson · Harvey Snowling

Published online: 2 April 2011

Abstract A cohort component projection of local populations based on sex and single year of age offers great value for planning local services, but demands data beyond the detail available. Local fertility, mortality and migration schedules by age and sex must be estimated sensitively to local variation if the results are to be of greater value than simpler methods of projection. Two approaches are compared, using data for the recent past: (a) direct estimation of local area age-specific schedules of fertility, mortality and migration based on data available to the national statistical agency; (b) graduation of national schedules using only local area population estimates by age, total numbers of births, and total numbers of deaths; age-specific migration is indirectly estimated from successive population estimates. These two approaches are compared with a projection using the same rates for each area. The three projections have been implemented for electoral wards in the Fife local government area of Scotland, using the flexible framework provided by POPGROUP software. Persuasive local population projections based on standard data for standard areas are feasible without the regular publication of migration flows.

Keywords Projection · Indirect estimates · Small area · Migration · Evaluation

L. Simpson (✉)
Cathie Marsh Centre for Census and Survey Research, University of Manchester,
Manchester M13 9PL, UK
e-mail: ludi.simpson@manchester.ac.uk

H. Snowling
General Register Office for Scotland, Ladywell House, Ladywell Road, Edinburgh EH12 7TF, UK
e-mail: Harvey.snowling@gro-scotland.gsi.gov.uk

Introduction

Demand for projections of the detailed age structure of the population in subnational areas arises from demands on government, and by government on other organizations, 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, highlighted by government preference for 'localism' over 'regionalism'. 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 vital statistics and migration data with sufficient 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 with single years of age and annual time periods provides the framework for this investigation, using the POPGROUP software (Andelin and Simpson 2007; Simpson 2006; www.ccsr.ac.uk/popgroup). A simplified version of the cohort component model, the Hamilton-Perry method (Smith et al. 2001; Swanson et al. 2010), is not considered here but can benefit from the results of this investigation.

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 parameterized 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 2000; Westerling 1995). Migration attracts a variety of approaches for subnational projection (Wilson and Rees 2005).

In this paper we test a practical approach to calibrate each component of change. Regional age schedules of fertility and mortality are scaled to reflect local levels of fertility and mortality, maintaining a common age-sex structure. For migration, we calibrate a regional age-sex schedule to be consistent with 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 in England, Wales and Scotland since 2001 by the publication of annual age-specific population estimates for small areas. The dependence of demographic methods on data availability is further discussed later in the paper.

We first set the context of Fife in Scotland for which we present small area projections, and outline 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 northeast of Edinburgh in Scotland. Its 360,000 population is divided between the major towns of Dunfermline, Glenrothes and Kirkcaldy which have expanded through housing developments in the past decade, several smaller towns, and rural areas. Three quarters of its boundary consists of 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 subareas in a number of ways. Here we use the 23 ‘multi-member ward’ electoral areas, henceforth called ‘wards’ (Fig. 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 number of internal and international 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. The younger outflow reflects mainly young adults leaving the parental homes for work and study outside Fife, while the slightly older outflow reflects mainly graduating university students located specifically in St Andrews. Estimating an equivalent of Fig. 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 organizations. 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

Fig. 1 The multi-member wards of Fife Council area. *Note* Named wards are examples used in the text

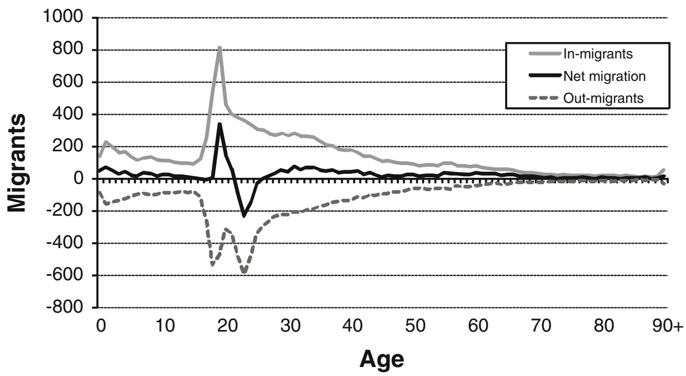


Fig. 2 Average annual migrants, 2002–2006, Fife. *Note* Migration flows combine internal and international migration. *Source* GROS

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 small area population estimates for data zones 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 National Health Service in the UK. International migration to and from data zones is less straightforward to estimate (GROS 2010a). Unfortunately, the published estimates of total migration to and from each zone cannot be aggregated to give flows for other areas of interest to local planning. This would require deduction of interzonal 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 and used in this paper are counts of births and deaths for each midyear annual period from mid-2001; and 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 effect 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 calculated using a single-area cohort component model 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 fertility, mortality and transition migration in- and outflows 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 as optional constraints (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; its algorithms are documented in its manuals and may be implemented on other technical platforms.

The Fife ward projections each begin in 2001 with six annual periods from mid-2001 to mid-2007 for which start and end populations, 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 single year age-sex cohort; it adjusts gross flows that are initially based on standard age-profiles in order to estimate in- and outflows of migration consistent with the changing population. These migration flows combine internal and international migration; the variants are described below. 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 19 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 subsequent 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 the GROS 2006-based projections (GROS 2008). Migration age schedules are the same for each ward. Fife's total of births, deaths and age-sex population as projected by GROS are set as constraints. Thus the annual net effect 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 from 2001–2002 to 2006–2007 using the following calculations (Andelin and Simpson 2007):

Total Fertility Rate (TFR) = Fife TFR × local births recorded/(births expected from Fife fertility rates applied to the local starting population structure).

Standardized Mortality Ratio (SMR) = Fife SMR × local deaths recorded/(deaths expected from Fife mortality rates applied to the local starting population structure).

Migration flow at age-sex group *a*: Migration expected from Fife migration rates applied to the 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 inflow and outflow.

In these ways, the training projection provides estimates of the TFR, SMR and single year of age-sex migration in- and outflows, 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 6 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 in- and outflows to a single external area.

(c) The Local Direct projection uses local age-sex schedules of fertility, mortality and migration estimated from the detailed data available to GROS. These are derived from mid-2001 to mid-2007 data for comparability with the local calibration projection. For migration, age-sex schedules of in- and outflows were estimated separately for movement within Fife and for all other movement, from the individual records of movement held by GROS. 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 Local Variation projection assumes implausibly equal demographic characteristics in each locality. The Local 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 Local 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 effect 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 Local Variation and Local Calibration approaches is to the direct approach, when the population is projected 19 years ahead. We examine the mean absolute deviation between the No Local Variation and the Local Direct projections across the 23 wards, for each of births, deaths, and population. If successful, the difference between the calibration and the direct projections would

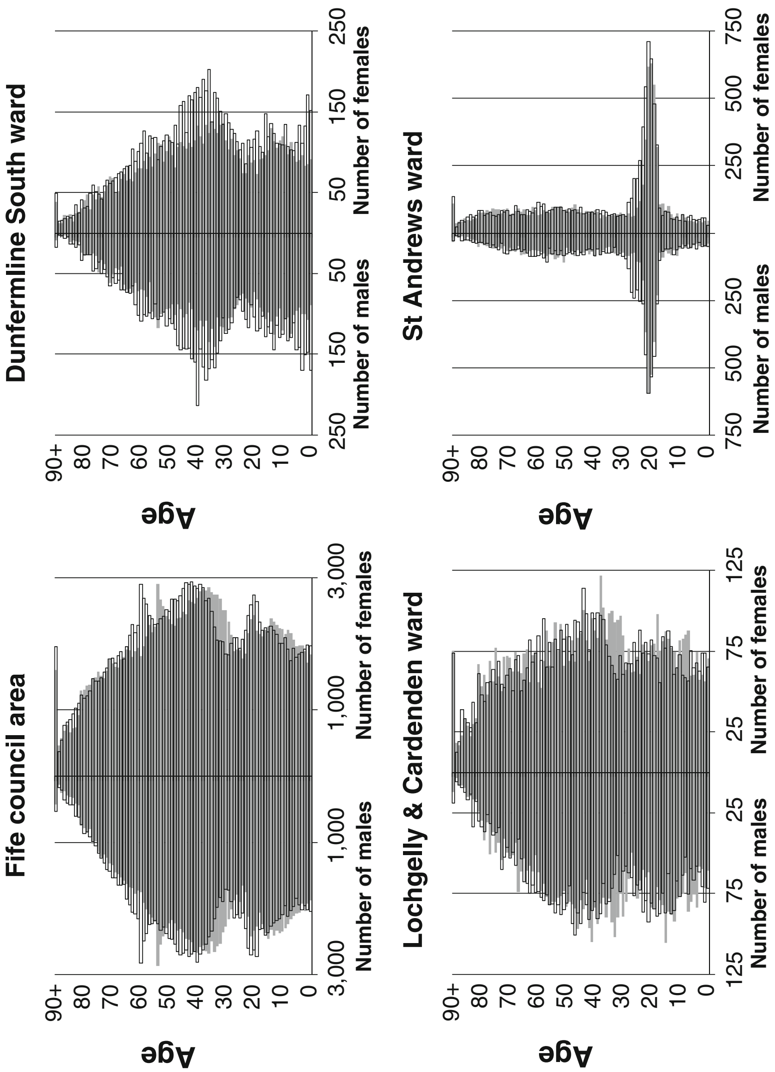


Fig. 3 Population estimates 2001 and 2007, Fife Council Area and selected wards. 2001 shown in *grey*; 2007 shown in *white*. *Source* GROS

involve much lower mean deviations, to indicate that the calibration projection captures variation between wards. In order to isolate the separate effects 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 Fig. 3 to further indicate the variety of local circumstances. Fife's pyramid shows a growing population with ageing, the postwar baby boom reaching age 60 in 2007. There are reductions among school-age children 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, in spite of amendments made by GROS using Higher Education Statistics Agency data of student residence, leading to the increase in adults in their late twenties apparent in Fig. 3. This possible flaw in the quality of population estimates helps to understand some of the later results for this unusual ward.

The proportional allocation of population and vital events from data zones to the areas of interest is an important part of the method, as it allows 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

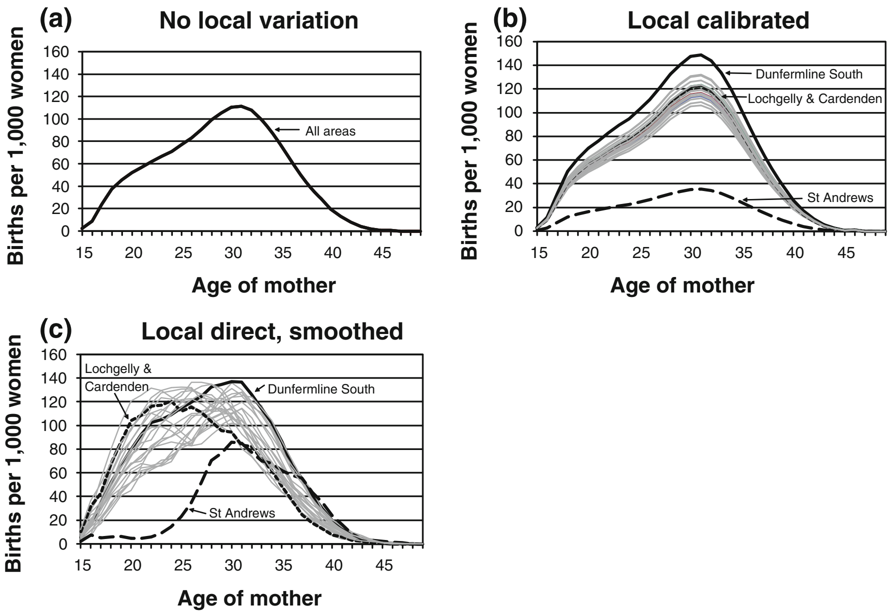


Fig. 4 Fertility schedule, three estimates

Table 1 Projected births, 2001–2026: deviation from the local direct projection

	No local variation	Local calibrated
Mean absolute deviation		
All 23 wards:	411 (9.4%)	99 (2.3%)
Deviation for single wards		
Dunfermline South	−1,347 (−14.3%)	600 (6.4%)
Lochgelly & Cardenden	−103 (−3.2%)	18 (0.6%)
St Andrews	4,190 (125.1%)	−668 (−19.9%)

Percentage deviation = (Annual average births in projection shown − Annual average births in local direct projection)/Annual average births in local direct projection × 100%

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 Fig. 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.

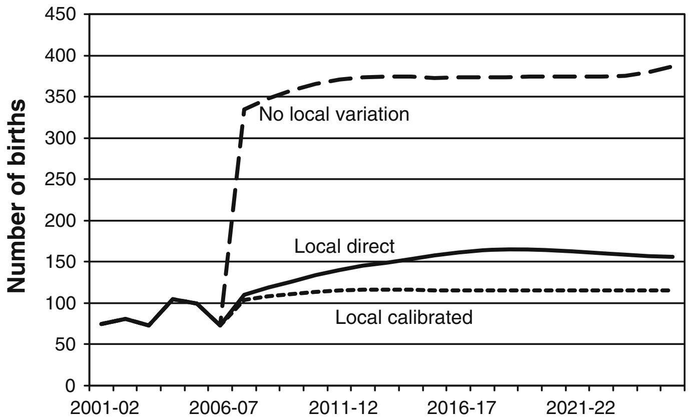


Fig. 5 Births for St Andrews ward, three projections. Each projection differs only in the fertility schedules used. Mortality and migration use the calibrated estimates

The calibrated estimates reproduce the correct number of births in each ward in recent years as 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 underestimate the births in Lochgelly & Cardenden and overestimate 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 effect 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 during 2001–2026 is stable at 100,000 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 quite different 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 difference of 2.3%. The full data show 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.

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

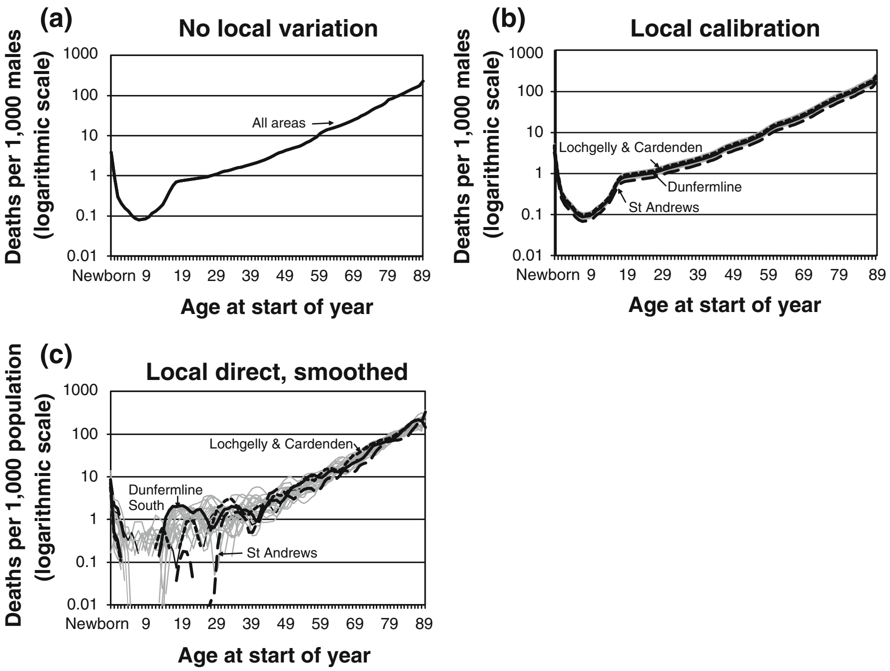


Fig. 6 Male mortality schedule, three estimates. Mortality rate (logarithmic scale), deaths per 1,000 males by single year of age

these emphasize fertility in the older ages (Fig. 4). Figure 5 shows this clearly in the unrealistically high number of projected births for St Andrews from the direct estimates of fertility. In this case the relatively large gap between the Local Calibration and the Local Direct estimates projections probably does not indicate a failure in the calibrated projections, but points instead to an advantage in the use of a smoothed fertility schedule.

Estimates of mortality and projection of deaths

The three estimates of mortality age schedules (Fig. 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 variability in the direct estimates is partly due to the small populations on which each age-specific estimate is based and is very evident at the younger ages. There are few discernible differences in mortality age patterns between areas. The effect on projected population 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. For the Local Direct estimates of mortality rates in Fig. 6, single-year-of-age rates were smoothed using the average of three.

Three projections are reported which differ only in mortality; each uses the calibrated estimates of fertility and migration. The total number of deaths projected

Table 2 Projected deaths 2001–2026: deviation from the local direct projection

	No local variation	Local 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%)

Percentage deviation = (Annual average deaths in projection shown − Annual average deaths in local direct projection)/Annual average deaths in local direct projection × 100%

for Fife during 2001–2026 is about 95,000. While the number of deaths in each area varies between the projections, the variation is less than that for births. The projection that assumes the same mortality curve for each area is on average 4.1% different from the projection using mortality schedules estimated directly for each area (Table 2). The projections based on 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 within 2.0% of the direct projection of deaths (not shown in the table) and within 0.6% on average. Unlike the situation with fertility, the university ward of St Andrews is not an outlier.

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 outflows from one or two external regions, but also constraints to migration counts and population targets. Where migration rates are referred to, they use the local area population as denominator; thus for in-migration they are not true demographic rates which would use the origin area's population as denominator.

The No Local Variation projection uses a single national UK schedule of migration rates drawn from the 2001 Census for both in- and outflows, and provides the projected single-year-of-age population of Fife as a whole as a constraint. The resulting migration rates are the same for each area, and refer to migration with one external area, the rest of the world.

The Local Calibration projection derives age-sex counts of migrants from a training projection for the annual periods mid-2001 to mid-2007, using the GROS single-year-of-age 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. Migration is with one external area, the rest of the world.

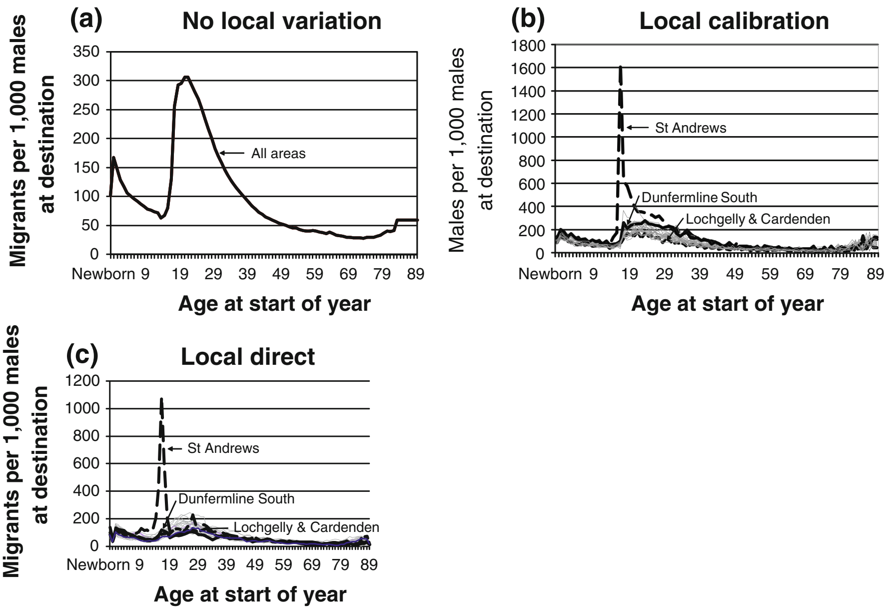


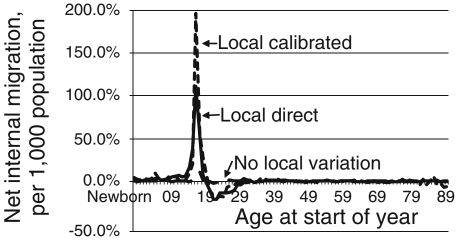
Fig. 7 Male in-migration rate age schedules, three estimates. *Note* Each chart uses a different *vertical scale*

The Local Direct estimates of migration are extracted from a variety of sources (GROS 2010a). Two sets of health-service administrative records provide information based on patient re-registrations within Scotland and with the rest of the UK. Immigration from outside the UK is measured by the International Passenger Survey which measures both in- and outflows, supplemented by health-service records for foreign nationals. 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 to and from the rest of the world, smoothed across three adjacent ages. The smoothed estimate replaces the unsmoothed estimate only when based on fewer than 200 migrants, to avoid dampening the extreme migration rates. Thus for the Local Direct projections, two external areas were used, Fife excluding the local area, and the rest of the world outside Fife.

Figure 7 shows the three estimates for male in-migration; 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 around age 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.

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

Fig. 8 Three estimates of net internal migration rates, St Andrews ward, 2007–2008. In-migration minus out-migration, as a percentage of mid-2007 population



2001–2007. However, the population estimates determine only the net effect of migration; the division into in- and outflows is based on the initial migration schedules from the previous census, using an assumption that any discrepancy is resolved evenly between in- and outflows. For this reason, and because the effect 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 Fig. 8 for St Andrews ward. The net estimated migration 2007–2008 is expressed as a percentage of the population of the relevant age group at the start of the period.

The No Local 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 Local Direct and the Local Calibrated estimates. However, the calibrated estimate in this ward shows much larger in-migration at age 17 than the direct estimate.

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

	No local variation	Local calibrated
(a) Net migration, summed across all ages		
Mean absolute deviation		
All 23 wards	1,975 (12.6%)	602 (3.8%)
Deviation for single wards		
Dunfermline South	12,746 (68.7%)	−1,194 (−6.4%)
Lochgelly & Cardenden	1,045 (8.6%)	−357 (−2.9%)
St Andrews	−748 (−4.1%)	2,593 (14.1%)
	No variation	**Calibrated**
(b) Net migration, quinary ages		
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	1,442 (148.9%)	389 (40.2%)

Percentage deviation = (Annual average net migration in projection shown − Annual average net migration in local direct projection)/Annual average net migration in local direct projection × 100%

The difference, for this single year of age, results from the different sources of migration, directly measured and indirectly estimated respectively. It may be also partly due to a difference in age-time plan of the projection model (age at start of the annual period) and the migration data (age at migration). The effect 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 50,000 people to Fife's population over the projection period. On average across the 23 wards, the calibrated estimate of migration across the projection period is 3.8% away from than the direct projection, less than one third the deviation of the projection with no variation between wards (Table 3). The difference between the calibrated and direct projections is particularly great in St Andrews (14%).

Not surprisingly, the difference between the projections is greater when age is considered. The calibrated projection is on average 11.5% away from than 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% different from the direct projection, but while large these discrepancies are very much less than for the projection with no variation between wards.

The 15–19-year age group in St Andrews registers the largest discrepancy (not shown in the table). The projection with calibrated estimates projects that 25,000 15–19 year olds will move to the area over the period, joining a cohort that nonetheless remains between 1,000 and 2,000 as students age out of that age group and are replenished by newcomers. The projection with direct estimates indicates in-movement of 18,000 15–19 year olds. Out-migration at ages 20–24 removes almost all the in-migrating students.

Discussion

Here we summarize 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 for direct statistics on migration. A training projection for a recent period integrates demographic age 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-specific migration that are consistent with recent local data. 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-specific 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 data for standard areas.

Direct estimates of demographic schedules are not usually available for local areas owing to the sensitivity of information based on few events, but tend also to be poor estimates of underlying demographic characteristics. 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 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 direct estimates. 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.

The quality of age-specific migration when estimated by calibration against successive population estimates is highly dependent on the quality of those population estimates. There is a large literature on the accuracy of population estimates. In the current UK context, small area estimates rely on health registration which may lag behind rapid population change consequent on housing developments, and more generally underestimates student post-study movement as noted for St Andrews.

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 (Scottish Government 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. It also 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 the available data. If local births and deaths are available with age detail of mother or age at death then further detail of local characteristics can be estimated. In the POPGROUP software this estimation is 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 at this level of detail. Should population estimates be available by quinary age groups, or in total only, then the same approach would provide estimates of migration in correspondingly less detail.

The training projection provides a time series of demographic indicators for the recent past which may 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 large student population may be handled other than by a cohort component projection, in the same way as for other large populations such as the 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 these ‘special populations’ may also be projected separately from the main cohort component projection, as for example in the projections for local authorities by the Office for National Statistics (2010), and by the General Register Office for Scotland (2010b). For areas that are particularly difficult to estimate, a mixed approach may be advisable, combining the results from calibrated, direct and possibly other estimates too, in order to reduce the worst errors of each (Armstrong 2001).

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 multiregional demographic structure, which would allow the proper specification of in-migration rates in relation to populations at risk. Multiregional models 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 effect of migration (Rogers 1990; Smith and Swanson 1997). If multiregional demography is at one end of a scale of complexity in cohort projections, the Hamilton-Perry method mentioned earlier is at the other end. Its use of local net cohort change ratios is a further simplification of the calibration approach evaluated here, which is particularly suitable for localities without vital statistics.

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 POPGROUP 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 age schedules.

Acknowledgments This paper is based on work financially supported by Fife Council and the General Register Office for Scotland. The paper was improved using comments from referees.

References

Andelin, J., & Simpson, L. (2007). POPGROUP: Population estimation and forecasting system, reference manual version 3.0. Manchester: University of Manchester. HYPERLINK http://www.ccsr.ac.uk/popgroup/about/manuals.html. Accessed December 9, 2010.

Armstrong, J. (2001). Combining forecasts. In J. Armstrong (Ed.), *Principles of forecasting: A handbook for researchers and practitioners* (pp. 417–439). Norwell: Kluwer.

Bray, H. (Ed.). (2008). *National population projections 2006-based*. Titchfield: Office for National Statistics, Series PP2 No 26 http://www.statistics.gov.uk/statbase/Product.asp?vlnk=8519. Accessed December 9, 2010.

Congdon, P. (1992). Multi-regional demographic projections in practice: A metropolitan example. *Regional Studies,* 26(2), 177–191.

General Register Office for Scotland GROS. (2008). *Population projections Scotland (2006-based) Population projections by sex, age and administrative area (revised 9 October 2008)*. Edinburgh. 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 December 9, 2010.

General Register Office for Scotland GROS. (2009). *Fife and Forth Valley migration report*, http://www.gro-scotland.gov.uk/files2/stats/migration/fife-forth-valley-migration-report.pdf. Accessed December 9, 2010.

General Register Office for Scotland GROS. (2010a). *Methodology used to produce GROS official migration estimates*. http://www.gro-scotland.gov.uk/statistics/theme/migration/methodology.html. Accessed December 9, 2010.

General Register Office for Scotland GROS. (2010b). *Sub-national population projections for Scotland: Methodology guide*. http://www.gro-scotland.gov.uk/files2/stats/population-projections/sub-national-pop-proj-methodology.pdf. Accessed December 9, 2010.

Johnson, F. A., Chandra, H., Brown, J. J., & 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., & McLeod, A. (2000). *Mortality in England and Wales, 1979–1992*, Occasional Paper No 1. Glasgow: MRC Social & Public Health Sciences Unit. http://www.sphsu.mrc.ac.uk/library/occasional/OP-001.pdf. Accessed December 9, 2010.

Marshall, A., & Simpson, L. (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: Oxford University Press.

Scottish Government. (2010). *Scottish neighbourhood statistics*. http://www.sns.gov.uk. Accessed December 9, 2010.

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

Simpson, L. (2006). *Integrating estimates within population forecasts*. Research Methods Programme Working Paper 18, Manchester: University of Manchester. http://www.ccsr.ac.uk/methods/publications/documents/WP18.pdf. Accessed December 9, 2010.

Simpson, L., & Yu, A. (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*(3), 283–307.

Smith, S. K., & Swanson, D. A. (1997). In defense of the net migrant. *Journal of Economic and Social Measurement, 24*(3/4), 249–264.

Smith, S. K., Tayman, J., & Swanson, D. (2001). *State and local population projections: Methodology and analysis*. New York: Kluwer Academic/Plenum Publishers.

Swanson, D., Schlottmann, A., & Schmidt, R. (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.

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*(2), 214–221.

Wilson, T., & Rees, P. (2005). Recent developments in population projection methodology: A review. *Population Space and Place, 11*(5), 337–360.