
Migration Assumptions in the UK National Population Projections: Methodology Review

Jakub Bijak¹

University of Southampton

Southampton, September 2012

¹ Social Sciences, Southampton Statistical Sciences Research Institute (S3RI) and ESRC Centre for Population Change, University of Southampton, Southampton, SO17 1BJ. Tel. 023 8059 7486; e-mail: J.Bijak@soton.ac.uk.

Contents

Executive Summary	2
1. Introduction	3
2. Background: State of the Art in Migration Projections.....	4
2.1. <i>Current practice: national statistical offices and international agencies</i>	4
2.2. <i>Migration projections and forecasts in the academic discourse</i>	6
3. Migration Assumptions in the UK National Population Projections.....	9
3.1. <i>Core methodology and its modifications</i>	9
3.2. <i>Discussion and evaluation</i>	12
4. The Way Ahead: Proposed Methodological Framework.....	14
4.1. <i>Sources of Information: Drivers and Trends, Data and Experts</i>	14
4.2. <i>Towards the New Framework: Scenarios, Probabilistic Forecasting and Decision Support</i> .	16
5. Discussion and Conclusions	24
Acknowledgements.....	26
References	27

Executive Summary

This review has two main aims. The first one is to evaluate methods currently used by the Office for National Statistics for setting the assumptions for extrapolating the migration component of the national population projections for the United Kingdom. The second objective is then to suggest new or improved methods for setting such assumptions. The scope of this analysis comprises migration flows at the national level, in the breakdown by age and sex and constituent country of the UK, as well as between the four constituent countries: England, Wales, Scotland and Northern Ireland.

An overview of current practice in setting migration assumptions for official projections prepared by statistical authorities of selected developed countries and by international agencies indicates a diversity of methods, with a dominant role of deterministic, argument-based scenarios. Nevertheless, the most advanced approaches involve probabilistic forecasting (**Section 2.1**). This is also the most promising direction of methodological development identified in the academic literature. Current state of the art includes clear guidance to separate modelling migration flows instead of net migration, and express outflows in terms of rates, rather than volumes (**Section 2.2**).

A review of migration assumptions in the current UK national population projections has identified exponential smoothing as the main method of extrapolating past trends towards an argument-based long-term target. The underlying data series come mainly from the International Passenger Survey and the National Health Service Central Register, and are augmented with some other sources for specific flows (**Section 3.1**). In discussion, these methods have been found to be in line with the current practice of many statistical agencies of the developed countries, but not necessarily with the main recommendations of the academic literature. In particular, modelling net migration levels is not recommended. Besides, a methodological shift from deterministic projections towards probabilistic forecasting, and dispensing with several ad-hoc solutions, has been found desirable (**Section 3.2**).

The proposed new methodological framework would involve a gradual shift from the deterministic towards the probabilistic paradigm. It would also make fuller use of different available sources of information: data and expert knowledge, next to a systematic, argument-based analysis of migration drivers and their possible trends (**Section 4.1**). Hence, a three-step evolution of the methodology has been proposed. In the first step, the existing argument-based scenarios would be streamlined and brought in line with the current recommendations in the academic discourse. The second step, involving a switch from deterministic projections to probabilistic (Bayesian) forecasts, would need to consider whole population dynamics, including fertility and mortality, rather than just the migration component. Finally, the third step would involve extending the forecasting process to an interactive decision support offered to the users (**Section 4.2**).

Finally, it is argued that the imminent release of 2011 census information offers an ideal timing for the implementation of a new methodology, which would position the UK official statistics amongst the world leaders. Nevertheless, in conjunction with methodological developments, the issue of quality of migration and population data needs to be addressed as well (**Section 5**).

1. Introduction

The overall objective of this methodological review (hereafter: the Review), commissioned by the Office for National Statistics (ONS), is twofold. Firstly, this document aims to evaluate methods currently used by the ONS for producing long-term and short-term assumptions concerning the migration component of the national population projections (NPP) for the United Kingdom (UK). The scope of this undertaking includes the projections of migration flows into, out of, and between the four constituent countries of the UK: England, Wales, Scotland and Northern Ireland. Secondly, the Review seeks to suggest new or improved methods for setting such assumptions at the national level, as well as for constituent countries, in the breakdown by age and sex. In order to achieve these two aims, the Review takes into account not only the existing academic literature and current practice of official statistics, but also the views and requirements of the ONS and the users of population projections, consulted within the framework of the NPP user forum. Since the current work is being conducted over 20 years after the previous review of migration assumptions took place (GAD/OPCS 1992), it is envisaged that its outcomes will contribute to methodological developments in setting assumptions for new sets of population projections for the UK, after the 2011 population census.

Throughout this Review, the exercise of computing plausible trajectories of future populations and their components of change, including migration streams, is referred to as ‘projection’. This is in line with the dominant practice of the official statistical agencies, and with the mainstream demographic tradition, whereby projection is conditional on the set of its underlying assumptions². Further, a single outcome of the projection exercise will be interchangeably referred to as a ‘variant’ or ‘scenario’. However, when unconditional rather than conditional statements about the future will be referred to, with some statement regarding their likelihood of occurrence, these will be clearly labelled as ‘forecasts’, in order to be consistent with the distinction made in the academic discourse (e.g. Keilman 1990, 2008). The notion of forecasts, increasingly more used in academic writing, as opposed to mere projections, preferred in the area of official statistics³, is important in particular with reference to public policy and planning, discussed in Sections 4 and 5 of this Review.

In terms of contents, next to the current Introduction, the Review has four further sections. In Section 2, an overview of the state of the art in migration projections is conducted. The focus is on a number of selected statistical offices of developed countries, international agencies tasked with preparing population projections, as well as on the academic discourse. The methodology for setting migration assumptions in the UK national population projections prepared by the ONS is summarised and evaluated in Section 3, together with several additions and modifications implemented since its inception. The suggested methodological changes and additions to the existing framework are discussed in Section 4. In particular, this report focuses on different sources of available background information (including data and experts), and ways of translating them into assumptions through analysing of drivers and trends. In this context, a framework for setting long-term migration scenarios is proposed, together with an extension into the areas of probabilistic

² Projection can be thus defined as “computation of future changes in population numbers, given certain assumptions about future trends in the rates of fertility, mortality, and migration” (The Population Reference Bureau [PRB] Glossary of Terms, www.prb.org/Educators/Resources/Glossary.aspx; 8/08/2012).

³ With a notable exception of Statistics Netherlands (see: <http://www.cbs.nl/en-GB/menu/themas/bevolking/methoden/begrippen/default.htm?ConceptID=934>; 8/08/2012).

forecasting and decision support. Finally, Section 5 discusses the main conclusions of the Review in the light of the proposed recommendations for the future rounds of ONS population projections.

2. Background: State of the Art in Migration Projections

This section presents a brief overview of methodological issues related to projecting migration as a component of population dynamics. The discussion is by necessity selective, and focuses on several examples of projections prepared within national statistical offices and international agencies, and on the debates within the academic community.

2.1. Current practice: national statistical offices and international agencies

The inclusion of migration assumptions in UK population projections dates back to 1965 (Shaw 2007), although more generally, the migration component has been incorporated in official population projections in a consistent manner only since early 1970s (for an overview, see Bijak 2010). Since then, despite on-going advancements in the area of population projections, the methodology for incorporating assumptions on particular components of demographic change remained largely focused on macro-level argument-based deterministic scenarios. Two decades ago, Keilman and Crujisen (1992) conducted a comprehensive survey of projection methods used in 30 developed countries. Back then, six countries did not include the migration component in the projections at all, and a majority of the remaining ones relied on some form of argument-based variants of population flows, whereby the numerical scenarios followed a qualitative description of plausible trends. Only a handful of countries used other methods, based on policy targets, trend-based extrapolations, interpolation of curves, or applying constant migration assumptions throughout the forecast horizon.

Since the overview of Keilman and Crujisen (1992), no other such a comprehensive study has been undertaken for official population projections, and filling this gap would be far beyond the scope of the current study. Even with respect to migration alone, a similar undertaking by Howe and Jackson (2005) discusses only five developed countries, all of which are covered in this Review. In that respect, Table 1 presents only a brief summary of selected features of international migration assumptions in recent population projections prepared by the statistical authorities of seven out of eight most populous EU countries (Germany, UK, France, Italy, Spain, Poland and The Netherlands)⁴, as well as six other developed nations (Australia, Canada, Japan, Republic of Korea, New Zealand, and the USA). Additionally, projections compiled by two pan-national organisations are shown; this includes the latest European Population Projections (EUROPOP 2010) of Eurostat, and the 2010 release of World Population Prospects of the United Nations Population Division.

In particular, Table 1 includes information on whether the assumptions are set in terms of *net migration* (immigration *less* emigration) or separately for the in- and outflows⁵; whether they concern *levels* of migration or *rates* expressed per population at risk of migrating; and what is the methodological basis for assumption-setting. With respect to the methodology, the assumptions presented in Table 1 are classified as *argument-based* (trying to quantify the underlying narrative), *expert-based* (explicitly mentioning the involvement of an expert panel), *target-based* (interpolating

⁴ At the time of preparing this Review, information for Romania was not publicly available.

⁵ The literature sometimes refers to immigration and emigration as *gross migration* flows, although since this term can be also used to denote the sum of both flows, it is not used in this Review.

Table 1. Migration assumptions in selected national and international population projections

Country	Base year	Net/flows	Rates/levels	Assumptions ^a	Variants ^b	Constancy assumption	Source
Australia	2006	Net	Levels	Argument/expert-based	3 variants	Throughout	[1]
Canada	2009	Flows	Rates	Target/argument-based	- 3 variants (immigration) - 1 variant (emigration)	- From 2012-13 (immigration) - Throughout (emigration)	[2]
France	2007	Net	Levels	Argument/expert-based	3 variants	Throughout	[3]
Germany	2008	Net	Levels	Target/argument-based	2 variants	From 2014 or 2020 (variants)	[4]
Italy	2011	Net	Levels	Target-based	3 variants	Never	[5]
Japan	2006	Net	Levels	Trend/argument-based	1 variant	From 2025	[6]
Korea (Rep. of)	2010	Net	Rates	Argument-based	3 variants	Never	[7]
Netherlands	2010	Flows	Levels / Rates ^c	Probabilistic argument-based	1 variant / 12 country groups	Never	[8]
New Zealand	2011	Net	Levels	Probabilistic argument-based	Derived variants ^d	From 2015 (distribution)	[9]
Poland	2008	Flows	Levels	Argument-based	2 variants	From 2020	[10]
Spain	2009	Flows	Levels	Trend/argument-based	1 variant	From 2018	[11]
UK	2010	Net	Levels	Trend/argument-based	3 variants	From 2016-17	[12]
USA	2008	Net	Levels	Trend-based	1 variant	Never	[13]
EUROSTAT EUROPOP	2010	Net	Levels	Argument-based + convergence	1 variant + 1 sensitivity test	Never	[14]
UN Pop. Division	2010	Net	Levels	Argument-based + convergence	1 variant	Never	[15]

Notes: ^a Argument-based assumptions include e.g. policy declarations and analysis of migration drivers; ^b Variants do not include 'zero migration' scenarios; ^c Levels for immigration, rates for emigration (returns); ^d Variants can be obtained e.g. from quantiles of probabilistic projections; Statistics NZ has also prepared two deterministic 'what-if' scenarios [9].

Sources: (accessed on 8/08/2012)

- [1] <http://www.abs.gov.au/ausstats/abs@.nsf/mf/3222.0/>
- [2] <http://www.statcan.gc.ca/pub/91-520-x/2010001/part-partie1-eng.htm>
- [3] http://www.insee.fr/fr/publications-et-services/docs_doc_travail/docf1008.pdf
- [4] <https://www.destatis.de/DE/ZahlenFakten/GesellschaftStaat/Bevoelkerung/Bevoelkerungsvorausberechnung/Bevoelkerungsvorausberechnung.html>
- [5] <http://www.istat.it/it/archivio/48875>
- [6] http://www.ipss.go.jp/webj-ad/webjournal.files/population/2008_4/05population.pdf
- [7] <http://kostat.go.kr/portal/english/news/1/9/index.board>
- [8] <http://www.cbs.nl/NR/rdonlyres/.../0/2011k2b15p42art.pdf> and <http://www.cbs.nl/NR/rdonlyres/.../0/2011k2b15p51art.pdf>; see also Howe and Jackson (2005: 16).
- [9] http://www.stats.govt.nz/browse_for_stats/population/estimates_and_projections/NationalPopulationProjections_HOTP2011.aspx
- [10] http://www.stat.gov.pl/cps/rde/xbcr/gus/L_proгноza_ludnosci_na_lata2008_2035.pdf
- [11] <http://www.ine.es/metodologia/t20/t2030251.pdf>
- [12] <http://www.ons.gov.uk/ons/rel/npp/national-population-projections/2010-based-projections/stb-2010-based-npp-principal-and-key-variants.html>
- [13] <http://www.census.gov/population/www/projections/methodstatement.html>
- [14] http://ec.europa.eu/economy_finance/publications/european_economy/2011/pdf/ee-2011-4_en.pdf
- [15] http://esa.un.org/wpp/Documentation/pdf/WPP2010_ASSUMPTIONS_AND_VARIANTS.pdf

between observations and a pre-set future target), *trend-based* (extrapolating past trends from data), or a combination of these types. Additionally, two countries (the Netherlands and New Zealand) apply a probabilistic methodology to argument-based projections, whereas both pan-national agencies considered (Eurostat and the UN Population Division) include an assumption of long-term convergence of net migration. Finally, Table 1 shows the number of variants of migration trajectories, and an indication whether constant migration is assumed throughout the projection horizon, only from a certain year, or never.

As it can be seen from Table 1, in terms of setting the assumptions for projecting international migration, there is no agreed standard, even amongst the largest European and non-European developed countries and international agencies. The main commonalities that can be seen are, firstly, that a majority of projections shown in Table 1 involve an argument-based component, and secondly, that most of them acknowledge uncertainty in the projections in a deterministic fashion, by producing more than one variant of future migration. Moreover, countries presented in Table 1 tend to set the assumptions in terms of net migration rather than specific flows (immigration and emigration), and more often in terms of levels than rates. All these elements are discussed in Section 2.2 with reference to the state of the art of the academic literature concerning assumptions on the migration component of population projections.

2.2. Migration projections and forecasts in the academic discourse

Over the past three decades, several authors presented various methodological overview studies related in general to population projections and forecasts, and in particular to the assumptions on their migration components. Many of these works include suggestions as to the directions, in which the projection methods should be ideally developed. Thus, for example, Willekens (1990: 23) suggested subdividing the population under study into homogenous strata, for which separate assumptions can be set, data availability permitting. Specifically for migration, Willekens (*idem*: 36) proposed distinguishing not only different strata of migrants (e.g. by different nationalities or by legal/illegal status), but also different directions of flows, and different types of migration (economic, family-related, political, etc.). Finally, Willekens argued that argument-based migration assumptions should ideally also take into account policy factors (e.g. targets and quotas in different countries), and should be based on reliable and comparable data and sound methods, including different forms of judgemental forecasting⁶, in order to produce different scenarios of future migration (*idem*).

More recently, O'Neill et al. (2001) suggested that migration assumptions in population projections need to carefully consider current and historical trends, migration theories, as well as feedback effects, which can manifest themselves via the relationships between human populations and their natural environment (see also Findlay 2011). Whereas the historical and more recent trends and patterns are acknowledged as an important element of migration assumptions, at least in the short term (e.g. Willekens 1990; Bijak 2010; de Beer 2011), the use of theories is much more contentious. In particular, Arango (2000) criticised the existing theories for being too vague and too fragmented along disciplinary boundaries for allowing predictions of such a multi-dimensional phenomenon as migration. Given that these theories are at best being capable of providing useful explanations *ex post*, Bijak (2010) suggested that migration predictions should be model-based rather than theory-

⁶ The judgemental forecasting path has been later explored in several forecasts of migration, which used multi-stage “Delphi” surveys amongst migration experts either directly (Drbohlav 1996) or as a way of obtaining prior distributions in the Bayesian framework (Bijak and Wiśniowski 2010; Abel et al. 2011).

based. Still, the challenge remains to identify a good model, which would integrate many available, yet disjointed sources of information, including data and expert judgement.

The most recent methodological overviews of population projection and forecasting include comprehensive studies by Wilson and Rees (2005), Booth (2006), and de Beer (2011). At a very general level, Booth (2006) distinguishes extrapolation-based, expectation-based and theory-based projection methods, the second one including expert-based approaches, and the third one covering also structural modelling. Despite the caveats on theories mentioned above, Booth (2006) argues that structured projections are useful for policy purposes, by allowing the construction of various “what-if” scenarios. Specifically with respect to migration, Wilson and Rees (2005) point to several emerging issues, such as the need for integrating internal and international migration components, and for including the sub-national distribution of the latter. De Beer (2011) stresses the role of distinctions between different groups of migrants for the purpose of setting the projection or forecast assumptions, and points to the inter-dependencies between immigration and (later) emigration of some of the returning migrants. Besides, de Beer (2011) emphasises the key role of transparency in assumption setting, as a key factor of accountability of the authorities tasked with preparing projections or forecasts. Notably, all of the mentioned authors are arguing for including methods for assessing the uncertainty of projections or forecasts, the most natural of which seem to be rooted in the probabilistic approach.

Specifically in the context of migration, the academic literature also addresses several important dilemmas that the authors of projections are faced with, some of which have already been flagged in Section 2.1. This Review explicitly deals with four of them. The first dilemma is related to the level of analysis, and in particular to the question whether the assumptions are to be set at the micro (individual) or macro (population) level. The second dilemma is concerned with modelling separate migration flows or, alternatively, net migration. The third dilemma, somewhat related to the second one, is whether migration should be modelled and projected or forecasted in terms of volumes or rates. Finally, the fourth dilemma is related to the choice of the underlying methodology. This is related to such questions as: Should assumptions be set in a deterministic or probabilistic fashion? Should expert opinion be explicitly included? These issues are discussed in turn below.

With respect to the micro versus macro dilemma, Willekens (1990) argued for applying the micro-level life course perspective to population projections and forecasts, with focus on individuals and their transitions between different demographic *states*, some of which are related to geographical space, as it is the case with migration. However, despite recent advances in multi-level multi-state modelling (cf. Courgeau 2007), which in principle allow for developing sophisticated projections based on micro-simulations, the uptake of this approach up to date was very limited, especially in the realms of official statistics (Wilson and Rees 2005). Among the reasons for this *status quo*, data limitations are often to blame (*idem*), which generally constitute a serious obstacle for building micro-simulation models with high-dimensional state space. This is the case for example with disaggregating the required transition (migration) probabilities by age, sex, region, etc.

As to the net migration versus specific flows dilemma, one of the most prominent voices in the discussion was the one of Rogers (1990), who argued that net migration is an artificial category, which should not be used in models of population dynamics. In that regard, the main flaws of net migration are twofold. Firstly, net migration obfuscates the difference between changes in migration

propensities and in population stocks. Secondly, it blurs the regularities in the age profiles of migration intensity $m(x)$, which in the case of specific flows can be successfully modelled for example with double-exponential curves of Rogers and Castro (1981). An example of the relevant formula for the thirteen-parameter variant of the curve can be written as (e.g. Wilson 2010: 196):

$$(1) \quad m(x) = c + a_1 \exp(-\alpha_1 x) + a_2 \exp\{-\alpha_2 (x - \mu_2) - \exp[-\lambda_2 (x - \mu_2)]\} + \\ + a_3 \exp\{-\alpha_3 (x - \mu_3) - \exp[-\lambda_3 (x - \mu_3)]\} + a_4 \exp(\alpha_4 x),$$

where x denotes age, c is constant, whereas a_i , α_i , μ_i and λ_i are model parameters. The successive components of the function (1) are related respectively to the overall level of migration intensity, early-childhood decline, labour force peak, retirement-peak, and late-life increase in migration⁷.

Unlike for specific flows, in a general case there is no such regularity in age patterns for net migration. On the other hand, there are some pragmatic arguments for the use of net migration for example when data on specific flows are unavailable and net migration can be obtained from census-based population estimates and vital events statistics, through population balance equation (Smith and Swanson 1998). Net migration can be also argued to offer a simple and inexpensive summary measure of the third component of population change beyond fertility and mortality (*idem*). Clearly, whereas the arguments against the use of net migration in projections are of more fundamental nature, the ones in its defence are chiefly practical.

Similarly to the previous dilemma, there are different arguments for modelling migration as volumes and as rates (for an overview and detailed discussion, see e.g. McDonald and Kippen 2002). The advantage of rates is that they clearly relate the process under study to the population at risk, and hence the projections cannot yield implausible values, such as negative population stocks, when projected emigration is in excess of the initial population size. On the other hand, for immigration it is usually impossible to determine the appropriate population at risk. For this reason, a compromise solution is to model migration as rates whenever possible – for example for flows within a closed migratory system with possible subsystems (see e.g. Kupiszewski and Kupiszewska 2011), as well as out of it – and as volumes for inflows to the system from the outside world.

Finally, in terms of the choice of projection or forecasting methodology, in the last twenty years the mainstream academic debate seems to have been increasingly shifting from deterministic to probabilistic projections or forecasts (a succinct overview can be found for example in Alho and Spencer 2005)⁸. In official statistics, the probabilistic approach was first applied by Statistics Netherlands in 1998 (de Beer 2011), but it did not find many followers since then. On the other hand, in recent academic literature, examples of probabilistic population projections abound. Examples

⁷ Wilson (2010) proposed an extended sixteen-parameter version, which additionally caters for the study-related peak in the late teens and early 20s by adding yet another double-exponential term for the student migration, whilst reducing the number of parameters for the retirement peak from four to three.

⁸ This shift can be also noticed for example in the recent Eurostat/UNECE Work Sessions on demographic projections, which are the main meetings of academics and practitioners dealing with population projections from Europe (and beyond). Whereas during the 2005 session, held in Vienna, probabilistic projections were barely mentioned, during the 2010 session, held in Lisbon, a majority of presentations on the projections of particular components of population change involved probabilistic models of some sort. See:

<http://www.unece.org/stats/documents/2005.09.projections.html>;

<http://www.unece.org/stats/documents/2007.10.projections.html>;

<http://www.unece.org/stats/documents/2010.04.projections.html> (as of 10/08/2012).

include the Uncertain Population of Europe (UPE) project, with methodology based on the extrapolative model for error (Alders et al. 2007), and expert-based probabilistic population projections prepared in the International Institute for Applied Systems Analysis (IIASA; see Lutz et al. 2004). Specifically for migration, Bijak (2010) argues for utilising a Bayesian approach, whereby extrapolations and expert knowledge can be combined in a joint probabilistic model in a coherent fashion. This approach has been later extended to include proper elicitation of expert opinion by the means of two-stage Delphi surveys (Bijak and Wiśniowski 2010; Abel et al. 2011). A detailed overview of different methods for forecasting and projecting migration is offered by Bijak (2010).

From the brief overview presented above, it is clear that the current practice of projecting or forecasting migration (or, wider, population) by the statistical offices of various developed countries and by international agencies does not necessarily follow the most recent advances in the projection methodology. In this context, Section 3 presents and assesses the methods currently used in the national population projections for the United Kingdom.

3. Migration Assumptions in the UK National Population Projections

This section discusses and evaluates the current methodology for setting migration assumptions for the UK national population projections. Since its implementation twenty years ago (GAD/OPCS 1992), the method has undergone several modifications, which are mentioned below. In the final part of this section, an assessment of the presented methodology is offered, taking into account the position of NPP assumptions on the methodological spectrum discussed in Section 2.

3.1. Core methodology and its modifications

The current methodology for setting the migration assumptions for national population projections for the United Kingdom dates back to the 1992 review prepared by the Government Actuary's Department (GAD) and the Office for Population Censuses and Surveys (OPCS) for the 1991-based projection round (GAD/OPCS 1992). At that time, the responsibility of preparing official population projections for the UK rested with GAD⁹. The GAD/OPCS review scrutinised the data available for historical migration series, with net migration possible to derive for the whole 20th century only decade-by-decade, based on census-based population estimates and vital events statistics. However, the International Passenger Survey (IPS), launched in 1964, offered a possibility of applying various methods of time series analysis to setting the migration assumptions.

This idea was further explored in detail in the GAD/OPCS review, which distinguished three groups of time-series methods: subjective (corresponding to argument-based assumptions in the terms of this Review), univariate, focused exclusively on migration, and multivariate, involving various economic covariates of migration (GAD/OPCS 1992: 3). Amongst univariate methods, the methods discussed included trend-fitting, ARIMA (autoregressive integrated moving average) models and exponential smoothing, which was found preferable due to its perceived simplicity. Nonetheless, the authors of the review noted two major limitations of the exponential smoothing method: an arbitrary choice of

⁹ ONS became responsible for producing population projections in January 2006, with the first projections prepared by the ONS being the 2008-round. See: http://www.gad.gov.uk/Demography%20Data/Population_projections_background.html (16/08/2012).

the smoothing parameter and overreliance on past data. Multivariate methods were dismissed altogether on the grounds of problems with forecasting the economic variables¹⁰.

In result, the GAD/OPCS review made seven substantive and two procedural recommendations for setting the migration assumptions (GAD/OPCS 1992: 22). For international migration, assumptions were to be set separately for UK nationals and non-nationals, and for each category four regions of origin/destination were to be considered: Old Commonwealth (with the USA), New Commonwealth, European Economic Community and the rest of the world. For each of those, a modified variant of the exponential smoothing was to be applied to the IPS data since 1975, whereby the projected value of migration in year $t+1$, m_{t+1}^p , is obtained via the following recursive equation (*idem*: 23):

$$(2) \quad m_{t+1}^p = \alpha m_t + (1 - \alpha) m_t^p,$$

with m_t denoting the actually observed values and α being a pre-set smoothing parameter. The GAD/OPCS review examined a weighted average of two models (2), with $\alpha = 0.2$ (long memory) and $\alpha = 1$ (short memory). Additionally, on the grounds of “common sense” (GAD/OPCS 1992: 6), the inclusion of a limiting factor was recommended in order to cap net migration increase or decline after ten years¹¹. As a result, for migration the trend factor obtained from a weighted average of models (2) would be reduced by one-tenth every year into the projection period, starting two-years after the jump-off year¹². The average net migration for these ten years is then taken as a basis for a long-term assumption. It has to be noted that the flows from and to the Republic of Ireland, as well as those of asylum seekers and members of the Armed Forces, which were not covered by the IPS, were to be treated separately, although no specific detail was provided in the 1992 review in that respect. In order to align with the population estimates, the projection assumptions are specified as mid-year to mid-year, rather than for full calendar years.

With respect to the sub-national migration, the GAD/OPCS review recommended allocating international migrants to the constituent countries according to the intentions recorded in the IPS. For flows between England, Wales, Scotland and Northern Ireland, the estimates were to be based on the records of the National Health Service Central Register (NHSCR). The age and sex distributions were to be assumed as empirical averages from the past ten years of the IPS and NHSCR data for the respective flows. Finally, in the long run, it was suggested that after the initial extrapolation for 15 years, all migration flows should decrease to zero in the subsequent five to ten years’ horizon. The procedural recommendations included a postulate of keeping the method flexible and open to modifications, as well as of continuing the consultation process between GAD and relevant government departments and respective statistical authorities of the UK constituent countries.

¹⁰ The problem of cumulating uncertainty stemming from the predictions of different variables in multivariate migration models was later confirmed in the context of vector autoregression (VAR) processes (Bijak 2010).

¹¹ At that time, a net migration value of +100,000 persons per year was considered too high – as noted in the GAD/OPSC review, even though “it is possible that the trend of the 1980s of a growing excess of inward over outward migration will continue for many years, generally *it is wisest with demographic projections* [emphasis added – JB] to assume that such trends will not continue forever” (*idem*: 6). By a way of comparison, a recent ONS Migration Statistics Quarterly (May 2012) reported net migration of +252,000 between October 2010 and September 2011, a decline by 3,000 from the preceding year (<http://www.ons.gov.uk/ons/rel/migration1/migration-statistics-quarterly-report/may-2012/msqr.html>; 16/08/2012).

¹² Personal communication with the ONS; July 2012.

Since its inception for the 1991-based projections, the methodology for setting the migration assumptions, as described above, has undergone several modifications. Firstly, already for the 1996-based projections the averaging of two versions of the model (2) was given up due to too much sensitivity of the outcomes to atypical last observations (ONS 1999: 34, after ONS 2012). Since then, the extrapolation has been based only on the long-memory variant of the smoothing procedure, with $\alpha = 0.2$ (*idem*). In the most recent (2010-based) edition of projections, long-term level of net migration has been assumed as +200,000 persons per year¹³, broken down into +175,000 persons corresponding to IPS estimates, and the remaining +25,000 to various adjustments (ONS 2012: 14). The last-mentioned number can be broken down into: (a) +15,000 for those who change their intentions reported in the IPS and become long-term migrants rather than short-term visitors ('visitor switchers'); (b) -5,000 to allow for change in the opposite direction ('migrant switchers'); (c) +15,000 for asylum-seekers; and (d) 10,000 migrants each way between the UK and the Republic of Ireland, a net of zero (*idem*). In the short term assumptions, an additional correction has been incorporated in the projections, declining from +40,000 in 2011-12 to +8,000 in 2015-16, in order to allow for an increase in migration exchange with ten new European Union member states, which joined the EU in 2004 and 2007 (with the exception of Cyprus and Malta). Moreover, for the years 2012-13 to 2015-16, a further allowance for 2,000 members of the Armed Forces returning from Germany has been included (*idem*: 23-24).

In the 2010-based projection, the allocation of international migrants to particular constituent countries has been based on the IPS averages for 2000-2009 for England, Wales and Scotland, and on five-year (2006-2010) averaged estimates of the Northern Ireland Statistics and Research Agency (NISRA)¹⁴. In result, out of the figure of +200,000 net migration assumed in the long term for the whole UK, +172,500 would be related to England, +10,000 to Wales, +17,500 to Scotland, and net international migration exchange with Northern Ireland will remain at zero (ONS 2012: 21). For intra-country flows, the respective long-term assumptions, based on recent data from the NHSCR, result in England losing 15,500 migrants net per year to Scotland (a gain of 8,500) and Wales (+7,000), with Northern Ireland again breaking even. In the short term (2010-11 to 2012-13), a smooth transition from recently-observed levels to the ones envisaged for the long-term has been assumed (*idem*: 22).

The age and sex distributions of international migration for England, Wales and Scotland were assumed to follow the ONS estimates of the corresponding components of total net migration: IPS, switchers, asylum-seekers and migrants to/from the Republic of Ireland (ONS 2012: 25). For migration between Northern Ireland and the rest of the UK, as well as for flows from/to ten new EU member states, age and sex distributions were assumed to follow the overall IPS-based average for the whole of the United Kingdom. For the remaining intra-UK migration, the NHSCR data for particular constituent countries was used (*idem*). For all flows, the distributions were calculated separately for immigration and emigration based on averaged values for the last five years of available data (2005-2009), and only then re-calculated into net migration proportions (*idem*).

Specifically with respect to emigration from Northern Ireland, an additional adjustment to total net migration was required for 18-19-year-olds, in order to reflect significant study-related emigration to

¹³ The exception is the first year of projection, 2010-11, with a value of +222,000 assumed (ONS 2012: 23).

¹⁴ In order to preserve the total values of IPS-based estimates for the whole of the UK, any differences between these totals and the sums of IPS sub-totals for England, Scotland and Wales, plus the NISRA totals for Northern Ireland, are being allocated to England (ONS 2012: 3).

other UK constituent countries, as reported by NISRA. This correction was done by applying the empirical net migration for this age group observed in 2008-09 throughout the projection horizon, and adjusting the remaining age groups in order to ensure the correct net migration total. This solution was clearly labelled as temporary, ideally to be replaced with a rates-based model in subsequent NPP editions¹⁵.

The two remaining assumption variants – the ones of high and low migration – were obtained by respectively adding 60,000 to the values from the main variant described above, or subtracting the same 60,000 from it, from the year 2011-12 onwards¹⁶.

3.2. Discussion and evaluation

In terms of setting the assumptions for the migration component of the NPP, there are several other considerations that have been acknowledged by the ONS in the documentation of projections. Firstly, with respect to the methodology, the results of the exponential smoothing method are very sensitive to the choice of the parameter α , which so far has been addressed only in a trial-and-error fashion (cf. ONS 1999: 34). Secondly, in terms of data, assumptions on migration between the UK and the Republic of Ireland are based on the information from the Irish Central Statistical Office (for Great Britain) and NISRA (for Northern Ireland), despite the IPS figures being available for these flows since 2008. Due to the apparent discontinuity in numbers, it was decided to wait before including the IPS figures in the assumptions until a longer time series becomes available (ONS 2012: 3). Moreover, with respect to setting the assumptions for the future, several issues were purposefully excluded from the analysis, such as: the impact of the recent EU enlargements beyond 2015-16; illegal migration; the impact of points-based system for non-EU immigration; and the effect of increasing university fees on student migration (*idem*: 24).

With reference to the current practice of official statistics and state of the art of the academic discussion on migration assumptions in population projections, presented in Section 2, the following observations can be made. Firstly, in terms of methodology and composition of the assumptions, the UK projections based on past trends and arguments about the future do not visibly stand out from similar endeavours undertaken in other developed countries. Applications of novel methodological approaches, such as probabilistic forecasting, are so far relatively rare; however, given the increasing attention to these methods on international forum, this is likely to change in the coming years.

Secondly, as most of the other statistical offices, with the clear exception of Statistics Netherlands, the ONS labels the insights into future population dynamics as ‘projections’, rather than ‘forecasts’. This approach has been criticised in the demographic literature for referring in fact to a purely arithmetic exercise of translating assumptions into results (Keilman 2008), instead of attaching a judgement of likelihood to the outcomes, which is what is needed for the users to make informed decisions (cf. Shaw 2007). As noted by Keyfitz already 40 years ago, even though “a demographer makes a projection, and his reader uses it as a forecast” (Keyfitz 1972: 363). In a similar vein, Booth (2004: 10) observed that “demography appears to be alone among disciplines in its insistence that its projections are somehow not forecasts of the future”. In this light, the terminology uniquely adopted by Statistics Netherlands merely constitutes an acknowledgement of the *status quo*.

¹⁵ Personal communication with the ONS; July 2012.

¹⁶ For 2010-11, only half of that correction has been applied. Source: (16/08/2012).

Thirdly, setting the assumptions in terms of net migration levels, despite being a common practice of many national statistical offices, goes counter to the suggestions in the mainstream academic literature. From the methodological point of view, the most coherent solution would be the one distinguishing both inward and outward flows, which treats immigration in terms of volumes, while emigration and inter-regional migration is best described in terms of rates. In this way, the characteristic features of different processes can be addressed and described more clearly and transparently. Besides, for outward flows, this also helps avoid conflating the effects of variation in the assumed propensity to migrate with the changes in population at risk.

Fourthly, with respect to the constancy of long-term migration assumptions, it can be argued that such a solution is methodologically neutral and anyway constitutes a common practice in many official population projections (Shaw 2007, see also Table 1). The assumption stems from a virtual impossibility to foresee future migration beyond the horizon of five to ten years (cf. Bijak and Wiśniowski 2010), and seems to acknowledge the ignorance of the authors of projections in that respect, rather than to envisage constancy as such. As noted by Shaw (2007: 19), the constancy assumption can be a way of admitting that, in the words of a US demographer Michael S Teitelbaum, “projecting immigration is impossible, but unavoidable”.

Lastly, owing to historical developments, the methodology includes several ‘patches’, which can be seen as arbitrary. Their inclusion was driven either by a need to respond to unpredicted changes in trends (e.g. for migration from Central and Eastern Europe following the EU enlargement), or by specific data situation (e.g. migration assumptions for Northern Ireland or for flows between the Republic of Ireland and the whole of the UK). Whereas they are obviously justifiable on the grounds of the argumentation presented in the NPP documentation (ONS 2012), at the same time they do not enhance the cohesion of the whole system of assumptions, and sometimes necessitate making further ‘tweaks’, as in the case of the age structure of migration from Northern Ireland.

A theoretical discussion aside, another important question is how well the methodology applied to migration assumptions reflects the demographic reality of the UK. In this context, an empirical evaluation of the UK population projections (here interpreted as forecasts) from the period 1954–2005 has been provided by Shaw (2007). With respect to migration it was noted that errors, unlike those for fertility or mortality, have increased over the years. This observation can be linked to the increasing volatility of the underlying data series. Besides, until the early 1990s, net migration assumptions remained negative, following a history of dominant outflows from the UK, but the situation significantly reversed in the last two decades (*idem*). Overall, throughout this period, the mean absolute error in net migration assumed ten years ahead totalled around 60,000, increasing to 100,000 after twenty years and 150,000 after thirty years (*idem*: 19).

Comparing with other Western European countries, Keilman (2007, 2008) observed that for the UK the mean absolute error in the net migration component for projections made in the second half of the 20th century was somewhat lower than the average for 14 countries included in the comparison. On the other hand, the mean error, allowing for the sign, was very much in line with the 14-country average. Also a statistical cross-country analysis indicated that in the case of the UK the country-specific effect for the relative error in projected net migration was almost in the middle of the ranking, although generally the patterns in ex-post errors of the migration component were found very difficult to explain (Keilman 2007: 27).

As a final point, given the difficulties with setting migration assumptions and explaining the source of their errors, it can be questioned whether precise projections or forecasts are at all possible, or if their accuracy should be even desirable (Bijak 2010). With such a volatile process as migration, failing to acknowledge the immanent uncertainty can provide the users of projections or forecasts with a false sense of security, and consequently lead to incorrect policy or planning decisions (*idem*). Some authors even argue that too ‘orderly’ assumptions of future migration are a tool of political influence rather than a valuable source of information (Pijpers 2008). An example of forecasts of immigration from the new EU member states to the UK (Dustmann et al. 2003), which were off the mark by over an order of magnitude, illustrates the point. Bearing that in mind, Section 4 provides some suggestions on moving beyond the pure issue of accuracy of projections or forecasts, and concentrating instead on providing the users with additional information to support the process of decision-making and planning.

4. The Way Ahead: Proposed Methodological Framework

In this section, a methodological framework for future NPP migration assumptions is suggested. Firstly, different sources of information for formulating migration assumptions, and hence for population projections in general, are discussed. Then, a possible evolution of the methodology is sketched, including a gradual shift from deterministic, argument-based variants, to probabilistic projections with derived scenarios, which can directly serve as tools of user-specific decision support.

4.1. Sources of Information: Drivers and Trends, Data and Experts

With respect to the UK, data for both directions of international migration flows (including the ones from and to the Republic of Ireland) are currently available from the IPS. The data can be obtained in a breakdown by constituent country or by broad groups of countries of origin/destination, although for more detailed distributions (e.g. by age) the IPS samples may be too small to allow for a meaningful inference. Additionally, for Northern Ireland, NISRA produces estimates based on data from Health Cards, which are required to register with a local GP. Separate series for the asylum seekers can be obtained from the Home Office¹⁷, whereas for the intra-UK flows, the NHSCR data are available. All these sources should provide series with at least a few observations, which could either indicate a broad trend, or in the absence thereof, in the case of short series, at least an approximate level of the processes under study. For calculating historical rates of outward migration flows, the series for populations at risk would need to be taken from the ONS mid-year population estimates.

Furthermore, the NPP Expert Advisory Panel, consisting of ten members representing the academic community (ONS 2012: 30), can provide additional insights into the expected future migration streams. The input of the Panel is largely based on a review of possible drivers and their tendencies of change, and this information is currently used mainly to provide arguments and justifications behind the long-term migration assumptions (ONS 2012). Currently, the expert information is not utilised to directly inform the assumptions (e.g. as in the case of probabilistic projections of Lutz et al. 2004), or to be combined with data in a common, Bayesian framework (Bijak and Wiśniowski 2010; Abel et al. 2011). On the contrary, as noted in the NPP documentation, “[t]he role of this expert panel is strictly advisory; responsibility for final decisions on the assumptions remains with ONS and

¹⁷ See <http://www.homeoffice.gov.uk/science-research/research-statistics/migration/> (23/08/2012).

the statistical offices of the devolved administrations.”¹⁸ From this perspective, the potential of the expert opinion seems to be underutilised in the NPP, especially that the existing methods allow for exploring the whole richness of expert information, rather than just concentrating on the central tendencies, such as means (cf. Abel et al. 2011).

With respect to the analysis of migration drivers underlying the argument-based assumptions, the information is largely available, yet may be quite dispersed. There are examples of reviews of drivers of human mobility and their trends, for example with reference to environment-related migration (Findlay et al. 2011) or different forms of human mobility across the external borders of the EU (Ariely et al. 2011). The latter study attempted a systematic “Survey of the Borders Universe” based on literature and other public sources (*idem*: 5), with the aim to identify the drivers and trends contributing the most to cross-border flows. An undertaken taxonomy allowed for distinguishing three ‘tiers’ of drivers and trends, according to the magnitude of their impact for the overall streams of future population flows across the EU borders. The subsequent scenario generation for the future concentrated on the ‘top-tier’ processes underlying particular forms of cross-border mobility (*idem*). With respect to the analysis of drivers and trends, the caveats concerning migration theories, mentioned in Section 2.2, are still applicable in that their review needs to be comprehensive enough to cover the most important facets of migration. From the point of view of NPP assumptions, there is a scope for applying a similar systematic process within the remit of the Expert Advisory Panel, especially as a part of the second step of the proposed framework described in Section 4.2.

A further possibility for enhancing the assumptions would be to explicitly include policy targets. In that respect, the methodological framework for projections with targets has been provided by Simpson (2005), and migration seems to be one of few demographic processes, where UK-wide government targeting is present. Currently, there exists a target to reduce net migration to the level of “tens of thousands” (i.e. below +100,000) by 2015¹⁹. However, so far the targeting remains far from being successful. As reported in the ONS Migration Statistics Quarterly from May 2012, for the last two years net migration remained above +250,000, chiefly as a result of a relatively stable immigration and a decrease of emigration since late 2008, the latter clearly beyond the control of the UK government (see Footnote 11). There are also concerns that the effects of reducing specific flows of migrants, especially students, who come to the UK for a limited period, may be short-lasting and aimed at artificially suppressing net migration in the run-up to the 2015 general elections²⁰. For these reasons, government targets, whilst remaining useful as a part of the analysis of migration drivers and informing the expert opinion, do not seem to be efficient enough to be included in the migration assumptions as a standalone component.

This discussion suggests that the assumptions should ideally be based on a mix of different sources of information, from available data to expert judgement, the latter taking into account – although not zealously – migration theories, drivers and targets. The framework for a proposed evolution of the current methodology is presented in Section 4.2 in the context of future NPP developments.

¹⁸ Source: online version of the 2010-based NPP Reference Volume; Chapter 1: Background and Methodology, <http://www.ons.gov.uk/ons/rel/npp/national-population-projections/2010-based-reference-volume--series-pp2/background-and-methodology.html> (23/08/2012)

¹⁹ See for example the speech of PM David Cameron on government’s migration policy from 14 April 2011: <http://www.bbc.co.uk/news/uk-politics-13083781> (22/08/2012).

²⁰ See the comment from the Institute for Public Policy Research: <http://www.ippr.org/articles/56/9118/another-voice-how-ministers-are-gaming-the-net-migration-target> (22/08/2012).

4.2. Towards the New Framework: Scenarios, Probabilistic Forecasting and Decision Support

The proposed changes in the methodology for setting the NPP migration assumptions could be scheduled to take place in three separate steps, which are described below. The first step would consist in streamlining and simplifying the existing argument-based scenarios and aligning it better with the current state of the art in the academic debate. The second step would then involve a switch to fully probabilistic forecasting, and the third one – a development of tools for user-specific decision support. The advantage of a step-wise approach is that not all the changes need to be implemented at the same time, which can be important from the point of view of resource allocation within the ONS. The steps differ with respect to their level of priority: whereas the first one is of an immediate concern, since it would directly feed into the next NPP round, the second and third ones need to be seen as more long-term strategic goals to strive towards. At the same time, the proposed framework gradually builds up the methodological sophistication, with the aim to produce one of the world's leading methodologies by the end of the process.

First step: Streamlining argument-based scenarios

To begin with, within the methodological framework there is a scope for reorganising the current assumptions underlying argument-based scenarios and minimising the number of 'fudge factors' in order to increase the transparency of the method. Firstly, it is suggested that assumptions are set in terms of **separate flows**, rather than net migration, especially given that the latter can be easily derived as a difference between immigration and emigration. This solution would be very much in line with suggestions from the academic literature, especially since argumentation and expertise are far more likely to concern specific flows rather than their difference.

Secondly, again in line with the current best practice, assumptions for immigration should be set in terms of **volumes**, whereas for emigration, and for flows between the constituent countries – of **rates**. This would ensure that the dynamics of most of the flows is correctly distinguished from changes in the underlying populations at risk. In setting the assumptions, the historical rates would be derived using mid-year population estimates, and for the purpose of presentation, the projected migration rates could be easily recomputed into volumes. Hence, even though **net migration** as such will no longer form an assumption in its own right, it could be easily calculated and presented as an outcome of two independent processes: immigration and emigration. With respect to the latter, the numbers of emigrants will be derived from rates by taking into account step-by-step changes in projected population, in the same way as it is the case for projected births and deaths.

Thirdly, there is a clear user demand for distinguishing flows for specific **groups of countries** of origin²¹, not least with respect to the reporting requirements of the Eurostat, as stipulated by the Regulation (EC) No. 862/2007 of the European Parliament and of the Council on Community Statistics on Migration and International Protection²². This distinction would be next to the one between the UK nationals and non-nationals. In the academic discourse, this solution has been advocated for example by de Beer (2008, 2011) and is reflected in the practice of Statistics Netherlands. For the UK, a minimum of two groups of countries should be distinguished: EU and

²¹ Feedback from the national population projections user forum; obtained via the ONS (16/07/2012).

²² Official Journal OJ L 199, 31.07.2007, pp. 23–29; available via <http://eur-lex.europa.eu> (24/08/2012).

non-EU, but the former could be further sub-divided into ‘old’ (pre-2004) and ‘new’ (post-2004) Member States, and the latter could for example follow the traditionally used division into Old Commonwealth, New Commonwealth and the rest of world. An additional option could be to distinguish flows by purpose (work, formal study, family reasons, etc.), as suggested by de Beer (2008, 2011), but given that the estimates will be largely based on the IPS, and need to be allocated to four constituent countries on that basis, sample-size considerations would not allow too fine-grained disaggregation to be used.

From this point of view, two options are available: (a) to apply a broad division of countries (EU/non-EU), distinguishing purpose of migration for these two groups; or (b) to keep a more detailed division in terms of geography, treating all migration flows jointly. Given the need for more detailed geographic breakdown, option (b) is recommended, whenever feasible given the sample sizes. In particular, for immigration, volumes of flows would ideally be modelled by groups of countries of origin, separately by constituent country of destination, and separately by citizenship (UK versus non-UK nationals); for emigration, where data are less reliable, emigration rates would be modelled only by constituent country of origin. Also for within-UK flows, extrapolations would be based on rates, without the breakdown by citizenship. In the last two cases, there is not enough data on resident population by citizenship to allow for making such a split.

As to the methodology of extrapolation, a switch to formal time series models is recommended, again in line with the best practice (Statistics Netherlands and Statistics New Zealand) and also as a way of preparing ground for an implementation of probabilistic forecasting in step 2. In order to increase transparency of projections, in line with de Beer’s (2011) postulate, short-term alterations and modifications of the assumptions would become obsolete. On the other hand, given the consensus on impossibility of predicting migration over the longer period, the **constancy** assumption beyond the horizon of five to ten years is to be retained. In this way, the **short-term** assumptions would be derived from extrapolations, and the **long-term** assumptions on immigration volumes and emigration rates would be fixed²³. The exact horizon, whether of five, ten, or an intermediate number of years, would be determined through consultations with the NPP user forum. At this step, **deterministic variants** of low or high migration would be preserved, only to be replaced with derived variants from probabilistic forecasts, once the latter are implemented in the second step. For the time being, the variants would thus involve subjective migration levels below or above the main extrapolated volumes (for immigration) or rates (for emigration).

The standard approach to time-series extrapolation would be to apply various **ARIMA models** (for an overview, see e.g. Wilson and Rees 2005 or Bijak 2010), which can have either longer or shorter memory, depending on the parameters of the process. Generally, the ARIMA model of order (p, d, q) has the following form (Greene, 2000: 776–777):

$$(3) \quad \left(1 - \sum_{i=1}^p \varphi_i L^i\right) [(1 - L)^d x_t] = c + \left(1 - \sum_{j=1}^q \theta_j L^j\right) \varepsilon_t .$$

In the above equation, x_t denotes the variable under study, and ε_t is the error term, which is usually assumed to be normally distributed. Further, L is the *lag operator*, such that $L^k x_t = x_{t-k}$, and thus

²³ Note that in this way net migration will not necessarily remain constant in the longer term, as it will also partially depend on the overall population dynamics. However, given the inertia of population processes, the impact of changes in the population size on net migration will be relatively slow.

$(1 - L)^d x_t$ denotes the d -th difference of x_t . The left-hand side of (3) includes therefore the *autoregressive part* of the model for the d -th differences of x_t , and the right-hand side is the *moving average part* related to the error term. In demographic applications, the order of ARIMA models applied to forecasting usually does not go beyond (1, 1, 1) (see e.g. Keilman et al. 2001); hence, for the extrapolation it is recommended to explore only the models within this parameter range.

In particular, a simple benchmark long-memory model for extrapolating migration five to ten years ahead could be a **random walk** with drift, a simple special case of ARIMA of order (0, 1, 0), whereby the logarithm of predicted migration, m_{t+1} , depends on the value from the preceding period, m_t , a drift constant c , and error term, ε_t , usually assumed to follow independent Normal distributions:

$$(4) \quad \ln(m_{t+1}) = c + \ln(m_t) + \varepsilon_t.$$

In migration forecasting applications, the random walk model, given its simplicity, was found to be favoured by formal model selection criteria (Bijak 2010; Bijak and Wiśniowski 2010). Besides, its statistical features, such as non-stationarity, seem to better reflect ‘unorderly’ character of contemporary migration flows (*idem*). The gradual increase of predictive uncertainty over the forecast horizon, as well as the permanent effects of shocks to the migration system, such as policy changes (e.g. with respect to visas for different migrant tiers or increasing tuition fees for students), also seem to be desirable features of this simple benchmark model.

Clearly, a random walk is only one of many possibilities offered by the ARIMA family, and the model selection can be conducted by using formal statistical methods (Bijak 2010). It may happen, especially with relatively short series, that these methods cannot detect any patterns in data besides a white noise – ARIMA(0, 0, 0). In such cases, this would effectively mean extrapolating the average from the sample into the future at a constant level. In general, the choice of a specific model remains an empirical issue, but one of the further possibilities which may be worth exploring consists in **averaging** of projections or forecasts yielded by different models or methods (*idem*). Recently, Rees et al. (2012) noted that for two population projections prepared within the framework of the UPTAP project²⁴ – one with assumptions formulated for emigration rates and the other one for emigration volumes – their average was reproducing the 2011 census estimates more closely than any of them separately. However, since this idea is best implemented within a probabilistic (Bayesian) framework (Bijak 2010), it is recommended that it is explored in step 2.

At this stage, it is recommended that the choice of an appropriate ARIMA model for extrapolation is based on the **goodness of fit** of the models to historical data series, which can be easily assessed and compared within many standard statistical packages. In order to allow for known shifts in levels of migration from certain origins, such as from the new EU member states after 2004 and 2007, different constants can be used for different sub-periods of the data sample; effectively this would mean using **indicator (0-1) variables** for selected sub-periods (e.g. after 2004 for the A8 migration).

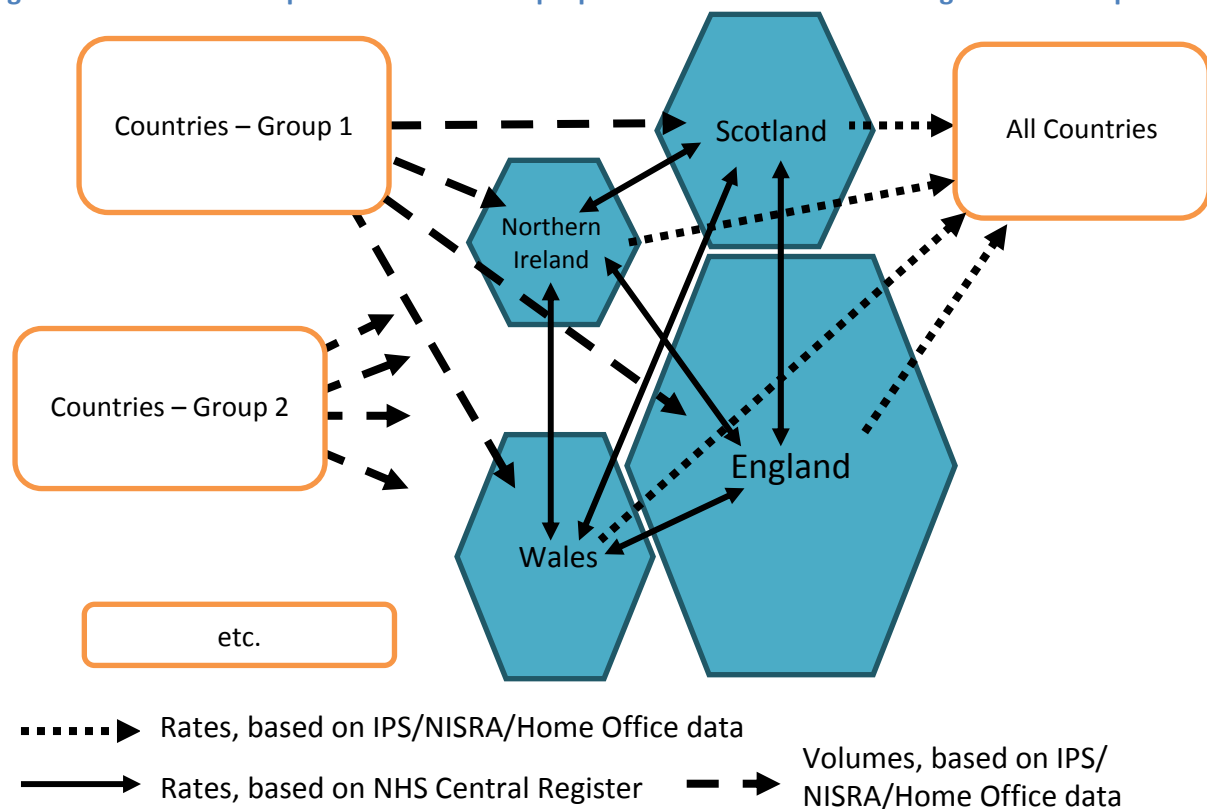
In all cases some differences between ARIMA-based extrapolations and the current method, based on exponential smoothing, will be inevitable. Nevertheless, it is recommended – again for the sake of the transparency of the approach – that these differences are explained *ex post*, rather than

²⁴ UPTAP: Understanding Population Trends and Processes; project “What happens when international migrants settle? Ethnic group population trends and projections for UK local areas” <http://www.uptap.net/project36.html> (24/08/2012).

attempted to be avoided by adding further modifications to the underlying model. It is worth remembering that no method is perfect – and there is irreducible uncertainty in the migration assumptions, regardless of which model is being used. In any case, this issue will be addressed comprehensively in step 2 of the process, after switching to full probabilistic projections.

The methodology proposed above could be based on three main **sources of data**: (a) IPS figures for all flows, including the ones previously not accounted for (such as migration from/to the Republic of Ireland); (b) NISRA estimates on international migration for Northern Ireland, based on Health Card information; (c) Home Office data on asylum seekers; and (d) NHS Central Register for flows between the four constituent countries. In this case, the estimates for (a), (b) and (c) would be required for the groups of countries discussed above; given the specific nature of asylum flows, it is recommended that the two categories of data are modelled separately and only added up in the final stage of the assumption-setting process. It is assumed here that the respective IPS totals already include the adjustments for visitor switchers and migrant switchers, as well as the flows to and from the Republic of Ireland (here, to be aggregated together with other EU countries). In all cases, the relevant time series should be as long as possible; where only relatively short series are available, these will be augmented with expert information in the second step of the process. In the future, should additional data become available, for example from the e-Borders exercise, these would supplement the survey-based estimates mentioned under (a). The scheme of the proposed framework for setting NPP migration assumptions is summarised in Figure 1.

Figure 1. A schematic representation of the proposed framework for NPP migration assumptions



Note: Flows of asylum seekers are treated as a special case of migration, to be modelled separately on the basis of the Home Office data, and only then re-aggregated with the IPS estimates

In order to reconcile the joint distributions e.g. by a countries of origin and the constituent country of destination, the **log-linear framework** for estimating contingency tables based on their marginal distributions is recommended before making the extrapolations, following the methodology advocated for example by Raymer and Rogers (2007). In particular, this may involve obtaining the joint distribution proportionally to both margins at the same time. This is especially recommended for all IPS-based series, where sample size can be a critical issue – since within-UK migration will be modelled using register-based data, such reconciliation might not be needed. Additionally, should IPS samples still be too small to allow for meaningful inference on their basis, the initial data could be first pooled over time. In all cases, since the processes under study are very high-level and relate to large geographic units (UK constituent countries), smoothing of age patterns and contingency table data would help reduce the noise in the data, whilst at the same time preserving the idiosyncrasies of particular flows²⁵.

With respect to other dimensions, the distribution by age should be smoothed, for example by using a variant of the **Rogers-Castro model** (1). For within-UK flows, given high volume of student migration, Wilson’s (2010: 199) version involving an additional university peak will be apt, whereby:

$$(5) \quad m(x) = c + a_1 \exp(-\alpha_1 x) + a_2 \exp\{-\alpha_2 (x - \mu_2) - \exp[-\lambda_2 (x - \mu_2)]\} + \\ + a_3 \exp\{-[(x - \mu_3)/\sigma_3]^2\} + a_4 \exp(\alpha_4 x) + a_5 \exp\{-\alpha_5 (x - \mu_5) - \exp[-\lambda_5 (x - \mu_5)]\},$$

with the third curve (here assumed to be symmetric) describing retirement migration, and the fifth one – student migration. For each flow, a visual inspection of the data should enable choosing the proper version of the Rogers-Castro model, with an appropriate number of peaks, for example for non-EU flows, retirement and post-retirement migration is unlikely to be significantly high.

Finally, in line with the recommendations of Wilson and Rees (2005), the assumptions should be propagated into projections within a joint, **coherent framework**, linking international and sub-national migration flows. An example is provided by Kupiszewski and Kupiszewska (2011) in their MULTIPOLES model, which treats different migration levels within a common multi-regional setting.

The workflow for assumption setting would be thus top-down, and comprise of the following stages:

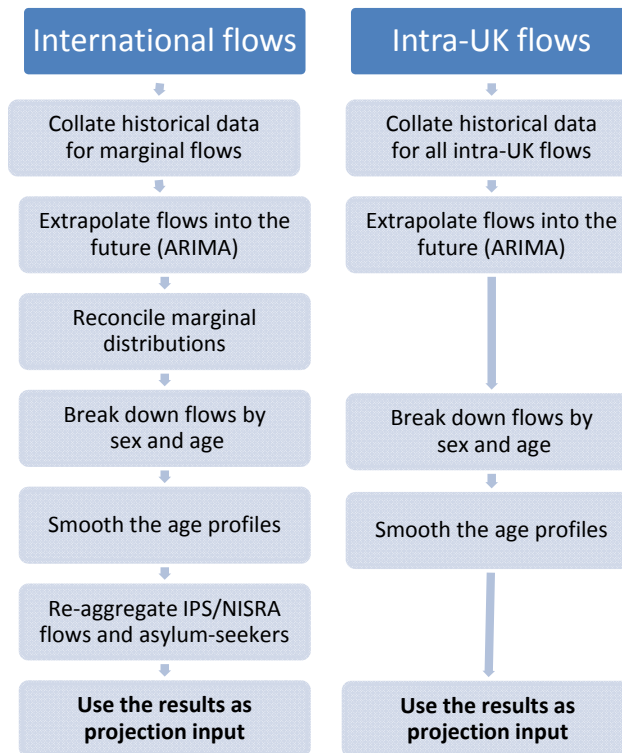
1. Collate information on marginal flows for the categories of interest (e.g. inflows by groups of countries, and by constituent country of destination; outflows by constituent country), as well as for the matrix of intra-UK flows between the constituent countries;
2. For each flow, extrapolate the series with a chosen ARIMA model (3) for five to ten years ahead, and fix the obtained final values as constant for the remaining part of the projection horizon;
3. For all IPS-based estimates and projections of flows, reconcile the extrapolated marginal information on different breakdowns by using log-linear models for contingency tables to estimate joint distributions by the characteristics of interest;
4. Separate the data by sex by using proportions from the most recent five to ten years of data, and smooth the age patterns in data for each flow by applying an appropriate variant of the Rogers-Castro model to averaged data from the same period;
5. Apply the age and sex distributions derived in stage 4 to the series extrapolated in stage 2, in order to obtain assumptions on specific migration flows (volumes or rates, as appropriate), by flow, sex and single-year age groups;

²⁵ This is different from the analysis at lower levels of spatial aggregations, where there is an increased risk of ‘oversmoothing’ and the resultant loss of important information on the local specificities.

6. Re-aggregate the IPS-based flows and those of the asylum seekers. Enter the overall migration assumptions into the underlying model of population dynamics in order to produce the ultimate projections. Derived quantities of interest (e.g. net migration) will be also obtained at this stage.

The workflow is illustrated in Figure 2 below:

Figure 2. Suggested workflow for the first step of the proposed procedure



Second step: Probabilistic forecasting

The rationale for moving towards probabilistic forecasting is mainly twofold. Firstly, the users are offered **much more information** embodied in the full probability distributions. Lawrence et al. (2006) cite several studies that suggest that forecast users actually prefer interval forecasts to point forecasts, when it is made clear how the former provide more information. Secondly, the probabilistic approach increases **awareness of the surrounding uncertainty** and can potentially safeguard the users against more extreme decisions. The same authors suggest that, in practical applications, in order to avoid extreme policy responses the users should focus on predictive intervals related to moderately high probabilities (e.g., 0.8 or 0.667), as a way of avoiding the potentially dangerous “illusion of control” resulting from undervaluing the uncertainty (*idem*). An extension of this argument in the context of migration has been provided by Bijak (2010: Chapter 11).

The second step in the evolution of the NPP methodology would thus consist in catching up with the standards set by the most advanced statistical offices (Statistics Netherlands and Statistics New Zealand) by implementing **probabilistic methods**. In this way, through an explicit acknowledgement of uncertainty via whole probability distributions, the outcomes would become forecasts rather than

just projections²⁶. As indicated in the first step, these distributions, obtained from the statistical extrapolation of the historical time series for volumes or rates, would then serve as means to **derive variants**, for example based on their central values (means or medians) or different quantiles. In the longer horizon (beyond five to ten years), the **constancy** assumptions could be then applied to the distributions, and hence to their derived variants, following the argumentation mentioned before.

At this step, different sources of information mentioned in Section 4.1 can be brought together by the means of **Bayesian statistics**, for example along the suggestions of Bijak (2010). In particular, this concerns the historical trends and expert judgement, which can be combined in a probabilistically-coherent way, with a correct description of uncertainty stemming from different sources – data, models and expert knowledge. Additionally, in the same way, **auxiliary data** for example from various administrative sources (e.g. Department for Work and Pensions or NHS) can be incorporated into the model. As mentioned above, the averaging of results of different forecasting models can be also explored at this step.

A separate issue, which will require attention at this stage, is a formal **elicitation of expert knowledge**, that could be then applied to Bayesian forecasts through the *prior distributions* of different parameters, which are then subject to updating by the data. A natural group of experts to extract the prior information from would be the NPP Expert Advisory Panel. An additional advantage of Bayesian methods in this context is that, due to including expert knowledge next to statistical data, this approach is particularly well-suited in the situation when time series are short (Bijak 2010), such as for migration between the UK and the Republic of Ireland.

In operational terms, however, the second step should ideally be implemented jointly for all three components of **full population forecasts** (rather than projections), not only migration. Otherwise, the resulting distributions would be conditional on the assumed trajectories of the remaining components of population change – fertility and mortality – and the predictive uncertainty would be thus artificially suppressed by not including the uncertainty of these components. One example here is the set of 2010 UN World Population Prospects, whereby fertility is forecasted using probabilistic (Bayesian) methods, yet mortality and migration remain deterministic²⁷. On the other hand, a prototype Bayesian methodology for full population forecasts has been recently implemented by Statistics New Zealand (Bryant and Graham 2011), and the ONS is currently exploring possible ways of implementing their methodology to the UK context. The results of the related feasibility study are expected to become available by mid-2013 (Elkin 2013, forthcoming).

Third step: User-specific decision support

Once the probabilistic framework for population projections is in place, the final step of the methodological evolution would consist of building mechanisms for providing bespoke **user support**. Probabilistic forecasts offer much greater flexibility here than the traditional deterministic variants

²⁶ From a purely conceptual point of view, especially in the light of Keilman's (1990, 2008) definitions, "probabilistic projections" can be seen as an oxymoron, since attaching a statement of likelihood turns a projection into a forecast. On the other hand, the very term "projection" is well rooted in the tradition of official population statistics. Hence, even though this report recommends a terminology change, the ultimate decision on this matter will be down to the ONS.

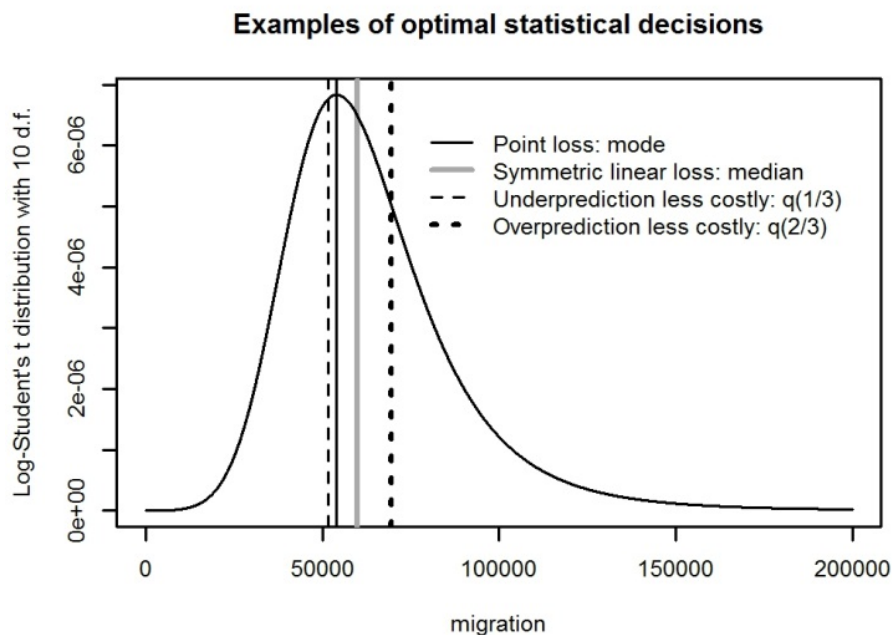
²⁷ See e.g. http://esa.un.org/wpp/P-WPP/htm/PWPP_Total-Population.htm (24/08/2012).

of argument-based projections, since they offer much more information than single trajectories. In this context, generic probabilistic forecasts can be made conditional on certain assumed future developments of e.g. some of the underlying components of population change. This would allow for generating probabilistically-coherent **derived ‘what-if’ scenarios**, which could correspond to the specific needs and questions of the forecast users. In that respect, the probabilistic forecasts would retain all the features of the current NPP, while at the same enhancing them to provide more information through probability distributions.

The bespoke derived scenarios could be additionally enhanced by the elements of formal **statistical decision theory**, which has been discussed in the demographic context by Alho and Spencer (2005), and specifically for migration by Bijak (2010). In this framework, the aim is to provide the forecast user with support for their decision-making on the basis of the underlying probability distribution. Here, the suggestion for the choice of a particular scenario is additionally founded on the assumed **loss function**, which describes how large losses would be incurred as a result of underpredicting or overpredicting the variable in question.

The loss functions, and thus solutions to the decision problem (if they exist), are specific to a particular context of decision-making (Alho and Spencer 2005; Bijak 2010). Hence, the same probability distribution can yield several optimal decisions, depending on the particular situations of the decision makers and their loss function. For example, medians or means reflect symmetric loss functions (linear or quadratic), where the direction of error does not matter for the decision maker. In turn, other quantiles reflect asymmetric linear losses: e.g. lower and upper tertiles (quantiles of the rank 0.333 and 0.667) relate to underestimation being twice less, respectively twice more costly than overestimation (see a stylised example in Figure 3).

Figure 3. A stylised example for the decision analysis concerning migration predictions



Source: Bijak (2010: 226)

Note: The graph in the original source contains an obvious mistake, which has been corrected above: lower quantiles correspond to underprediction being **less** costly than overprediction, and vice versa

Given this specificity of loss functions, Bijak (2010) suggested that the decision support should be an iterative process, whereby the forecast makers and users remain in close collaboration: the users provide insights into their decision setting (loss functions) and the forecasters provide a tailor-made decision advice on that basis.

5. Discussion and Conclusions

From the sceptical point of view, it may be debateable to what extent changes to the methodology for setting the NPP migration assumptions are presently needed. The current method, producing several argument- and trend-based deterministic variants is fully operational, and is in line with the standard used in many other developed countries. On the other hand, during the twenty years since the previous review was undertaken, the methodology accumulated several arbitrary ‘tweaks’ as a result of sudden changes in trends or data availability. In this way, what was an elegant solution proposed in the GAD/OPCS (1992) review, has lost some of its simplicity and transparency. Recently, methodological developments in the area of population projections and forecasting have gained momentum. The most advanced statistical agencies, such as those in the Netherlands and New Zealand²⁸, have shifted to probabilistic forecasting, with the UN Population Division to follow suit.

Thus, the imminent release of the results of the 2011 population census for the UK provides a unique opportunity to develop a new, state-of-the-art methodology, which could count amongst the most advanced in the world. As argued above, the new method could be based on relatively simple solutions, which would fully retain the features of the current approach, and add new aspects to them. The information which is already available, such as the one from the NPP Expert Advisory Panel, could be then utilised more fully in the second step of the process, in conjunction with data, in order to produce a coherent picture of expectations of future migration (and, more broadly, population). The users would also benefit by being provided with more information – embodied in the probability distributions – rather than just three deterministic variants. As mentioned above, the advantages of this approach could be additionally strengthened by offering tailor-made decision support. Hence, the most important recommendations for change can be summarised as follows.

Summary of recommendations

Step 1. Streamlining argument-based scenarios

- Switch to assumption-setting for separate migration flows (volumes for immigration, and rates for emigration and flows between constituent countries), rather than net migration
- Separate assumptions to be set for large groups of countries of origin (immigration) and for overall outflows (emigration)
- Sources of data: IPS, NISRA, Home Office (asylum seekers), NHS Central Register (intra-UK flows); to be supplemented by e-Borders statistics once they become available
- Smoothing of irregularities in data by using the Rogers-Castro model (with student peak for intra-UK flows) and log-linear framework for contingency tables, based on table margins
- ARIMA-based extrapolations for short-term, and constancy for long-term assumptions

²⁸ In the overview prepared by Howe and Jackson (2005), the approach of Statistics Netherlands was presented as a model example of how migration assumptions for official population projections should be set.

Step 2. Probabilistic forecasting

- Switch to a fully probabilistic methodology, ideally Bayesian to allow expert judgement
- Elicitation of expert information, for example from the Expert Advisory Panel
- Calculation of derived variants, based e.g. on the predictive medians or other quantiles

Step 3. User-specific decision support

- Elicitation of specific settings (loss functions) of decision problems from the users
- Calculation of bespoke ‘what-if’ scenarios, including solutions to the decision problem

The changes proposed in this Review would address many of the postulates being echoed across the academic literature. Firstly, already fifty-seven years ago, John Hajnal suggested that population forecasting should “involve less computation and more cognition than has generally been applied” (Hajnal 1955: 321). While in preparing quantitative projections or forecasts there is no escape from computations as such, a mix of trend-based and expert-based approaches seems to offer a compromise solution. Secondly, as argued by de Beer (2011), transparency and simplicity of applied methods of population forecasting are their desirable features, so as to avoid the ‘black box’ outcomes which are difficult to explain. This view is seconded by Howe and Jackson, who stated that “[a]t every turn, the theoretical and empirical basis for modelling assumptions would have to be spelled out.” (Howe and Jackson 2005: 25). Thirdly, the adopted distinction between short-term and long-term assumptions aligns with the suggestions made by Nico Keilman, who argued for the use of probabilistic forecasts in the short run, and scenarios thereafter (in Willekens 1990: 42–44). As argued in Section 4.2, in the proposed approach, various ‘what-if’ scenarios can be derived from the underlying probability distributions and take into account various user requirements. Finally, even a terminological shift from ‘projections’ to ‘forecasts’ could be significant; it would bring about the possibility of analysing the ex-post accuracy of the outcomes in a straightforward fashion, with the aim of informing the successive rounds of the forecasting exercise in an iterative way. As noted by de Beer (2011), the question of accuracy is what distinguishes projections from forecasts. The accountability of statistical offices is thus not related to – anyway unattainable – perfect accuracy of the predictions, but rather to the transparency and soundness of the underlying methods.

An issue which remains beyond the scope of this Review, yet will be crucial for the success of any adopted method of population forecasting, is the one of adequate data quality²⁹. In the UK, the problem is especially acute in the case of migration. The reasons include the lack of adequate data on emigration and underutilisation of administrative sources in producing migration estimates. Ideally, changes in the methodology for population projections or forecasts should be accompanied by further developments in the area of migration statistics. In general, comprehensive suggestions for improving the UK population and migration statistics have been recently provided in the conceptual framework by Raymer et al. (2012a). It is also hoped that the quality statistical information will radically improve with the introduction of the e-Borders scheme, which however is unlikely to provide the required statistical information before the end of 2017 (ONS n.d.). Until then, alternative solutions can be sought – for example those based on a comprehensive modelling of migration for pan-national systems, such as the Integrated Model of European Migration (IMEM) for 31 countries of Europe and the rest of the world (Raymer et al. 2012b). Such estimates could be at least utilised to validate the IPS-based statistics on UK migration.

²⁹ This paragraph is largely based on personal communication with Prof. Phil Rees, to whom I am very grateful for sharing his views on this issue.

In summary, the present moment seems to be ideal for implementing changes not only to migration assumptions, but also to the whole UK national population projections. On the other hand, the proposed changes have an evolutionary character and their implementation can be stretched over time. Of course, the new methods, once implemented, would need to be monitored and adjusted with respect to their performance and meeting the user needs. There are also many avenues, along which the methods could be developed further, e.g. by exploring micro-simulation approaches and analysing different dimensions of population processes in the light of available data (cf. Wilson and Rees 2005). Hence, the next comprehensive Review would also ideally need to be carried out sooner than the current one – perhaps after ten years, rather than 20. This is not only to take advantage of the newly-available data sources, such as e-Borders, but also to make sure that the methods used remain fit for purpose and follow the world-wide best practice in population forecasting.

Acknowledgements

The Review was carried out within the general framework for cooperation between the Office for National Statistics (ONS) and the Southampton Statistical Sciences Research Institute (S3RI). The author is very grateful for helpful comments of colleagues from the ONS and members of the national population projections user forum. Notwithstanding, all opinions expressed in this report are those of the author only and do not necessarily express the views of the ONS, the user forum, or any of the institutions with which the author is affiliated.

References

- Abel GJ, Bijak J, Wiśniowski A, Findlay A and McCollum D (2011) Model specification and results of Bayesian forecasts of total and environmental immigration to the UK, 2010–2060. Modelling review MR8, Migration and Global Environmental Change project. London: GoS Foresight. <http://www.bis.gov.uk/assets/bispartners/foresight/docs/migration/modelling/11-1168-mr8-model-bayesian-forecasts-total-immigration-to-uk> (27/07/2012)
- Alders M, Keilman N and Cuijssen H (2007) Assumptions for long-term stochastic population forecasts in 18 European countries. *European Journal of Population* 23(1): 33–69.
- Alho JM and Spencer BD (2005) *Statistical Demography and Forecasting*. New York: Springer.
- Arango J (2000) Explaining Migration: A Critical View. *International Social Science Journal* 52(165): 283–296.
- Ariely G, Warnes R, Bijak J and Landesmann R (2011) Futures of Borders. A Forward Study of European Border Checks. Warsaw: Frontex. http://www.frontex.europa.eu/assets/Publications/Research/Futures_of_Borders.pdf (27/07/2012)
- Beer J de (2008) Forecasting International Migration: Time Series Projections vs Argument-Based Forecasts In: J Raymer and F Willekens (eds.) *International Migration in Europe: Data, Models and Estimates*. Chichester: John Wiley (pp. 283–306).
- Beer J de (2011) *Transparency in population forecasting: Methods for fitting and projecting fertility, mortality and migration*. NIDI books 83. Amsterdam: Amsterdam University Press.
- Bijak J (2010) *Forecasting international migration in Europe: A Bayesian view*. Springer Series on Demographic Methods and Population Analysis, vol. 24. Dordrecht: Springer.
- Bijak J and Wiśniowski A (2010) Bayesian forecasting of immigration to selected European countries by using expert knowledge. *Journal of the Royal Statistical Society Series A* 173(4): 775–796.
- Booth H (2004) On the importance of being uncertain: Forecasting population futures for Australia. *People and Place*, 12(2): 1–12.
- Booth H (2006) Demographic forecasting: 1980 to 2005 in review. *International Journal of Forecasting* 22(3): 547–581.
- Bryant J and Graham P (2011) Demographic estimates and projections using multiple data sources: A Bayesian approach. 58th World Statistics Congress of the International Statistical Institute (ISI), Dublin, 21–26 August. <http://isi2011.congressplanner.eu/pdfs/950832.pdf> (21/08/2012)
- Courgeau D (2007) *Multilevel Synthesis: From the Group to the Individual*. Springer Series on Demographic Methods and Population Analysis, vol. 18. Dordrecht: Springer.
- Drbohlav D (1996) The Probable Future of European East-West International Migration – Selected Aspects. In: F. W. Carter, P. Jordan, and V. Rey (eds.), *Central Europe after the Fall of the Iron Curtain; Geopolitical Perspectives, Spatial Patterns and Trends*. Wiener Osteuropa Studien 4. Frankfurt: Peter Lang & Co. (pp. 269–296).
- Dustmann C, Casanova M, Fertig M, Preston I and Schmidt CM (2003) The impact of EU enlargement on migration flows. Home Office Report 25/03. London: Home Office.
- Elkin M (forthcoming 2013) Applying Statistics New Zealand's Bayesian framework to demographic estimates and projections for England and Wales. MSc Dissertation in Official Statistics, University of Southampton. *In progress*.
- Findlay AM (2011) Migrant destinations in an era of environmental change. *Global Environmental Change* 21(Suppl. 1): S50–S58.

- GAD/OPCS (1992) National population projections: a new methodology for determining migration assumptions. Occasional Paper No. 42. London: Government Actuary's Department / Office of Population Censuses and Surveys. <http://www.ons.gov.uk/ons/rel/npp/national-population-projections/occasional-paper-no--42/index.html> (17/07/2012)
- Greene WH (2000) *Econometric Analysis*. Prentice Hall, Upper Saddle River, NJ.
- Hajnal J (1955) The Prospects for Population Forecasts. *Journal of the American Statistical Association* 50(270): 309–322.
- Howe N and Jackson R (2005) Projecting immigration: A Survey of the Current State of Practice and Theory. Washington DC: Centre for Strategic & International Studies. http://csis.org/files/media/csis/pubs/0504_howe_jacksonprojimmigration.pdf (17/08/2012)
- Keilman N (1990) *Uncertainty in national population forecasting: Issues, backgrounds, analyses, recommendations*. Amsterdam: Swets & Zeitlinger.
- Keilman N (2007) UK national population projections in perspective: How successful compared to those in other European countries? *Population Trends* 129: 20–30.
- Keilman N (2008) European Demographic Forecasts Have Not Become More Accurate Over the Past 25 Years. *Population and Development Review* 34(1): 137–153.
- Keilman N and Cruijsen H (eds.) (1992) *National population forecasting in industrialised countries*. NIDI & CBGS publication 24. Amsterdam: Swets & Zeitlinger.
- Keilman N, Pham DQ, and Hetland A (2001) *Norway's Uncertain Demographic Future*. Statistics Norway, Oslo.
- Keyfitz N (1972) On future population. *Journal of the American Statistical Association* 67(338): 347–363.
- Keyfitz N (1982) Can Knowledge Improve Forecasts? *Population and Development Review* 8(4): 729–751.
- Kupiszewski M and Kupiszewska D (2011) MULTIPOLES: A Revised Multiregional Model for Improved Capture of International Migration. In: J Stillwell and M Clarke (eds.) *Population Dynamics and Projection Methods*. Understanding Population Trends and Processes, 2011, Vol. 4. Dordrecht: Springer (pp. 41–60).
- Lawrence M, Goodwin P, O'Connor M and Önkäl D (2006) Judgemental forecasting: A review of progress over the last 25 years. *International Journal of Forecasting* 22(3): 493–518.
- Lutz W, Sanderson WC and Scherbov S (eds.) (2004) *The End of World Population Growth in the 21st Century. New Challenges for Human Capital Formation and Sustainable Development*. London: Earthscan.
- McDonald P, and Kippen R (2002). Projecting future migration levels: should rates or numbers be used? *People and Place* 10(1): 82–83.
- O'Neill BC, Balk D, Brickman M and Ezra M (2001) A Guide to Global Population Projections. *Demographic Research* 4(8): 203–288.
- ONS (1999) National population projections: 1996-based: Series PP2, No. 21. London: The Stationery Office.
- ONS (2012) National Population Projections, 2010-based reference volume: Series PP2 No. 28. Titchfield: Office for National Statistics. <http://www.ons.gov.uk/ons/rel/npp/national-population-projections/2010-based-reference-volume--series-pp2/index.html> (17/07/2012)
- ONS (n.d.) Delivering statistical benefits from e-Borders. Titchfield: Office for National Statistics. <http://www.ons.gov.uk/ons/guide-method/method-quality/imps/latest-news/delivering-statistical-benefits-from-e-borders/index.html> (28/08/2012)
- Pijpers R (2008) Problematizing the 'orderly' aesthetic assumptions of forecasts of East–West migration in the European Union. *Environment and Planning A* 40(1): 174–188.

- Raymer J, Rees P and Blake A, et al. (2012a) Conceptual Framework for UK Population and Migration Statistics. Titchfield: Office for National Statistics. <http://www.ons.gov.uk/ons/guide-method/method-quality/imps/latest-news/conceptual-framework/a-conceptual-framework-for-population-and-migration-statistics---download-file.pdf> (10/08/2012)
- Raymer J, Forster JJ, Smith PWF, Bijak J and Wiśniowski A (2012b) Integrated Modelling of European Migration: Background, specification and results. NORFACE Migration Discussion Paper No. 2012-04. http://www.norface-migration.org/publ_uploads/NDP_04_12.pdf (28/08/2012)
- Raymer J and Rogers A (2007) Using age and spatial flow structures in the indirect estimation of migration streams. *Demography* 44(2): 199–223.
- Rees P, Wohland P, Norman P and Boden P (2012) Ethnic population projections for the UK, 2001–2051. *Journal of Population Research*, 29(1): 45–89.
- Rogers A (1990) Requiem for the net migrant. *Geographical Analysis* 22(4): 283–300.
- Rogers A and Castro LJ (1981) Model Migration Schedules. IIASA Report RR-81-30. Laxenburg: International Institute for Applied Systems Analysis. <http://www.iiasa.ac.at/Admin/PUB/Documents/RR-81-030.pdf> (27/07/2012)
- Shaw C (2007) Fifty years of United Kingdom national population projections: How accurate have they been? *Population Trends* 128: 8–23.
- Simpson L (2005) Integrating estimates and targets within a population projection. ESRC Research Methods Programme Working Paper No 18. Manchester: University of Manchester. www.ccsr.ac.uk/staff/documents/Simpson2005IntestsinpopprojRMPWP18.pdf (10/08/2012)
- Smith SK and Swanson D (1998) In defense of the net migrant. *Journal of Economic and Social Measurement* 24(3–4): 249–264.
- Willekens F (1990) Demographic Forecasting; State-of-the-Art and Research Needs. In: CA Hazeu and GAB Frinking (eds.) *Emerging Issues in Demographic Research*. Amsterdam: Elsevier (pp. 9–66).
- Wilson T (2010) Model migration schedules incorporating student migration peaks. *Demographic Research* 27(8): 191–222. <http://www.demographic-research.org/Volumes/Vol23/8/>
- Wilson T and Rees P (2005) Recent Developments in Population Projection Methodology: A Review. *Population, Space and Place* 11(5): 337–360.