An uncertain future of immigration in Europe

Document Version

Final published version

[Link to publication record in Manchester Research Explorer](#)

Citation for published version (APA):

Winiowski, A., Bijak, J., Kupiszewski, M., & Kupiszewska, D. (2012). An uncertain future of immigration in Europe: Insights from expert-based stochastic forecasts for selected countries. In M. Okólski (Ed.), *European immigrations: trends, structures and policy implications* (pp. 211-232). (IMISCOE Research). Amsterdam University Press.

Published in:

European immigrations

Citing this paper

Please note that where the full-text provided on Manchester Research Explorer is the Author Accepted Manuscript or Proof version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version.

General rights

Copyright and moral rights for the publications made accessible in the Research Explorer are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Takedown policy

If you believe that this document breaches copyright please refer to the University of Manchester's Takedown Procedures [<http://man.ac.uk/04Y6Bo>] or contact uml.scholarlycommunications@manchester.ac.uk providing relevant details, so we can investigate your claim.

Chapter Title: An uncertain future of immigration in Europe Insights from expert-based stochastic forecasts for selected countries

Chapter Author(s): Arkadiusz Wiśniowski, Jakub Bijak, Marek Kupiszewski and Dorota Kupiszewska

Book Title: European Immigrations

Book Subtitle: Trends, Structures and Policy Implications

Book Editor(s): Marek Okólski

Published by: Amsterdam University Press

Stable URL: <https://www.jstor.org/stable/j.ctt46mtbm.12>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <https://about.jstor.org/terms>

This content is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (CC BY-NC-ND 4.0). To view a copy of this license, visit <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

Amsterdam University Press is collaborating with JSTOR to digitize, preserve and extend access to *European Immigrations*

JSTOR

9 An uncertain future of immigration in Europe

Insights from expert-based stochastic forecasts for selected countries

*Arkadiusz Wiśniowski, Jakub Bijak, Marek Kupiszewski and
Dorota Kupiszewska*

9.1 A research paradigm for the uncertain world (Kupiszewski and Kupiszewska)

The theoretical underpinning of this volume is the analysis of changes in migration patterns over time, discovering regularities and formulating a theory describing the observed migration processes. Marek Okólski offered such a theory in chapter 1. Migration forecasting allows for the extension, albeit with a degree of uncertainty, of migration patterns into the future. In an ideal and reasonably predictable world, it would be possible to use forecasts to identify the moment of transition from an emigration to an immigration country for those countries that are in the transition phase. To do so, the IDEA project team (see note 9 in Introduction) tried to introduce some quantitative measure of the stage in the migration cycle and to develop some kind of ‘trajectory’ on which each country could have been located, depending on its migration system characteristics. However, operationalising the theory turned out to be an impossible task due to its long-term perspective. Also posing a challenge were the irregularity of the migration processes involved, complexity and inherent differences in the historical changes in migration observed in various countries (see chapter 2) and uncertainty embedded in forecasting (see later in this section). Substantial impact of historical and political processes (for example, the demise of communism in Central Europe and the 2010-2011 revolts sweeping Arab World countries), which have profound impact on migration flow counts, is impossible to account for in a theoretical framework, although it is very apparent in statistics. This failure is not a surprise: Sture Öberg and Babette Wils (1998) and Marek Kupiszewski (2002) noted that theories of migration are always difficult, and very often impossible, to operationalise and use for forecasting purposes.

Apart from purely theoretical considerations, migration forecasting serves utilitarian purposes by supporting the formulation of migration policies. The authors of this chapter feel strongly that migration policies are often formulated without a firm base of factual support and are rather predicated on qualitative information only. While chapters 3 through 8 present data on past trends, this forecasting exercise is done in an effort to provide some quantitative insight into the future and thus provide a better factual basis for policymaking. The Bayesian forecasting method, implemented in the IDEA project, allows for combining qualitative information with statistical data. Moreover, it has an added bonus: very clearly showing how much uncertainty is involved in the forecasts. As a consequence, it allows for the formulation of alternative policies in case some extreme scenarios are realised.

One of the central methodological problems of forecasting in general and of demographic forecasting in particular is how to handle forecasting error. In population forecasting, Nico Keilman (1990) specified several types of possible sources of errors in deterministic population forecasts. If we shift from thinking in terms of deterministic forecasts and forecasting errors calculated *ex post* to considering some measure of forecasting uncertainty¹ calculated *ex ante*, as posited by the proponents of the stochastic approach, the sources of forecast errors listed by Keilman could be interpreted as the sources of uncertainty of a forecast.

Even though the methodology of migration forecasting is much less developed than the methodology of population forecasting – as is clearly demonstrated by Jakub Bijak's (2010) study – we can attempt to create a similar, but not identical list of sources of uncertainty in migration forecasting. The forecasts of migration are uncertain because:

- We cannot uniquely define the migration phenomenon.
- We cannot measure migration precisely.
- We do not know how migration depends on explanatory variables or how to quantify the relationship.
- We do not know how migration, or the determinants of migration, will change in the future.

The key problems with the definitions of international migration include, but are not limited to, the lack of international harmonisation and the changes of definitions over time (see Kupiszewska & Nowok 2008; Nowok, Kupiszewska & Poulain 2006; Nowok & Kupiszewska 2005). The differences between the IDEA countries involve, among others, the duration of stay criterion used to determine which individuals are included in the migration flow count. In particular, the following categories of migration flow statistics can be observed: changes of permanent residence (e.g. in Poland), long-term migration for one year or more (e.g. the immigration of non-European Economic Area citizens to the Czech Republic and

Hungary), short-term and long-term migration for more than x months (e.g. three months in Austria) or changes of usual residence irrespective of the duration of stay (e.g. in Spain).

The second source of uncertainty is the measurement of migration. The problems related to measurement include, among others, underestimation due to registration avoidance. This is probably the largest source of uncertainty in migration statistics (for legal flows, especially emigration) and applies to all the countries where population registers are the source of migration flow statistics. In addition, illegal migrants in the IDEA countries are usually not included in official statistics, with the exception of Spain. In France, immigration flow data are available for non-EEA citizens only, while data on emigration flows are not available at all. Greece started to provide data on immigration flows to Eurostat in 2007 (data on residence permits issued in 2006). In the countries where data on total immigration flows are available, some disaggregated statistics are sometimes missing, for example, there is no data on immigration by country of previous residence in Hungary and France and no data until 2005 on immigration by citizenship to Poland. Even if such disaggregated data are produced, there is sometimes a problem of a large value in the 'unknown' category (e.g. a large number of immigrants with unknown country of previous residence in Austria).

Other problems include the delays in the production and publication of statistics, differences in data published in various secondary sources and the lack of relevant metadata. The numbers published in various sources may be provisional or final, may or may not include administrative corrections, may be produced using different methodologies or may cover different categories of migrants (Kupiszewska & Nowok 2008; Nowok et al. 2006).

In explanatory models, where migration is modelled as a function of exogenous variables, the randomness of the parameters of a forecasting model itself or the uncertain relationship between the migration and exogenous variables constitute a separate source of forecast uncertainty. In addition, problems with definitions and the measurement of the exogenous variables can be very similar to those specified above with respect to migration.

Finally, we have to deal with the last source of uncertainty, which is closely linked with the lack of knowledge of future migration processes. Migration is very volatile and can rise to extremely high levels or drop very quickly, as shown by migration from Poland to Germany at the turn of the 1980s and 1990s and post-enlargement migration from new EU member states to the United Kingdom and Ireland. These changes are dictated by external factors, not usually accounted for in the forecasting models, such as wars, significant political changes and modification of migration policies. If we look back to 1980 and inspect the forecasts of migration between the European states 25 years on (around 2005), we would – firstly –

have an incorrect set of states; no one at that time predicted the fall of the Soviet Union and the creation of an independent Latvia, Lithuania, Estonia, Ukraine and Moldova, not to mention the Central Asiatic Republics, and the dissolution of Czechoslovakia and Yugoslavia – the latter in a violent manner. Second, at that time nobody, expected the 2004 and 2007 enlargements of the European Union to accept, among others, ten new countries formerly in the Communist bloc and such large-scale flows from new to selected old EU member states after 2004. The discontinuity of demographic processes is one of the main problems for migration forecasters.

The above considerations lead us to the conclusion that, as far as migration is concerned, we live in a very uncertain world and uncertainty is a key factor shaping the development of migration processes. However, many researchers and policymakers are under the illusion that the world is predictable. In economics, Roman Frydman and Michael Goldberg (2007) noted in their influential book that we are unable to precisely understand and describe the future; therefore, forecasting error is a natural phenomenon. The authors made a plea for recognising the limitations of economic knowledge. In demography, in parallel to economics, similar thinking has developed, which aims to incorporate uncertainty – both as a concept and as a quantifiable variable – into research, and in particular into forecasting. Philip Rees and Ian Turton (1998) offered a simple classification of possible ways of dealing with uncertainty in population forecasting.

- To ignore it by constructing a single variant population forecast and pretending that uncertainty does not exist.
- To incorporate it by constructing a multi-variant population forecast (typically with central, high and low variants).
- To create a stochastic forecast which allows for the provision of some *ex ante* measures of uncertainty.

The solution to depart from deterministic population forecasts and devise a forecast where the uncertainty of the results could be somehow quantified – that is, stochastic forecasts of population – has been considered for well over a decade. The methodology of such forecasts has been proposed by Ronald Lee and Shripad Tuljapurkar (1994), Wolfgang Lutz, Warren Sanderson and Sergey Scherbov (1996), Juha Alho and Bruce Spencer (1997) and many others. Stochastic forecasts have been prepared for many developed countries, but they are not yet part of the standard toolkit of national statistical institutions and international organisations. However, the development of the new sub-field was advanced enough to allow pioneers to write a monograph-cum-textbook (Alho & Spencer 2005).

Interestingly, the development of stochastic approaches in population forecasting has led to a change in professional self-identification. Alho and Spencer (2005: 3), quoting other prominent demographers, describe this change in the following way: James Vaupel, they say, defined a demographer

as ‘someone who knows Lexis [diagram]’. Joel Cohen defined a demographer as ‘someone who forecasts population wrong’. Finally, Alho and Spencer (2005) themselves define a statistical demographer as ‘someone who knows Lexis, forecasts population wrong, but can at least quantify the uncertainty’. While somewhat anecdotal, this change of definition shows the evolution of modern demography; perhaps a similar process will arise in the future in other social sciences.

The development of stochastic population forecasting had quite a limited impact on the way the components of population change have been forecasted. As Bijak (2010) demonstrated, there were some attempts at modelling and forecasting migration using probabilistic and stochastic methods, but this was not in existing mainstream research. Bijak (2010) proposed a new methodology, in which Bayesian statistics was applied to produce forecasts of international migration flows between Germany and three European countries: Italy, Poland and Switzerland.

A methodology similar to the one proposed by Bijak, extended through the application of the Delphi method to complement information on immigration reported by national statistical offices, was used in the current study. The forecasts were prepared for seven European countries: Austria, the Czech Republic, France, Hungary, Italy, Poland and Portugal.² The forecasts concerned total immigration flows (nationals and foreign citizens together), except for France, where immigration of foreigners was forecasted (in line with the data availability). In all cases, the definitions used by national statistical offices in their demographic statistics were applied.

There were four reasons why immigration, rather than net migration, was forecasted. First, the data concerning emigration (required to calculate net migration) are less reliable than data concerning immigration, mainly because people often do not report their departure, even if they register their arrival at the destination. Second, the elicitation of knowledge through the Delphi method was deemed to be easier when the questions referred to immigration rather than to net migration. Moreover, it was expected that using a variable referring to unidirectional flows (rather than a difference of two flows, as is the case for net migration) would reduce the uncertainty, which was expected to be (and was indeed) very large. Finally, from the policy point of view, immigration is the variable that policymakers are keen to look at, whereas net migration somehow blurs the picture from the perspective of the receiving country. The methodology and results of this study, presented in subsequent sections of the current chapter, are described in more detail and in more technical terms in Bijak and Wiśniowski (2009, 2010).

9.2 Bayesian forecasting of migration: A methodological outline (Wiśniowski and Bijak)

9.2.1 *Why forecast migration?*

Forecasting migration is a very difficult research task for the reasons outlined above, in particular, due to the lack of comprehensive migration theories and the lack of data, including short time series (Willekens 1994; Kupiszewski 2002). On the other hand, migration forecasts with suitable uncertainty assessments are crucial for obtaining credible population predictions, as well as for other policy and planning purposes.

It has to be noted that, in general, the primary aim of socio-economic forecasting is not to predict the future with 100 per cent accuracy, as this is impossible. The forecasters' goal is rather to provide input in order to guide the political decision-making process. Therefore, the key issue in the forecasting process is not to offer a (what is improbable) point estimate of the future values of some variables, but rather to provide a reliable, quantitative assessment of the related uncertainty span. In the current study, this was attempted by estimating predictive probability distributions (i.e. probability distributions of the future volumes of immigration flows). Such distributions are a way of illustrating how probable particular immigration scenarios are in our view.

With these considerations in mind, we sought to provide forecasts of immigration into seven European countries up to the temporal horizon of 2025. To do so, we relied both on quantitative data and on knowledge of country-specific migration experts.

9.2.2 *Forecasting methodology*

Frans Willekens (1994) noted that the problems with the migration predictions mentioned above call for including expert knowledge in the forecasting exercise. The methodology that allows for a combination of subjective expertise and hard data in a coherent and structured manner is the so-called Bayesian approach, dating back to the theorem proposed by an eighteenth century English statistician, Rev. Thomas Bayes (1763). This approach allows for combining expert judgements or opinions, treated as 'prior' knowledge, with data.

The variables forecasted in the current study were immigration inflows, both total and from up to three of the most important sources of immigration (countries of previous residence or – in the case of immigration to France and Hungary – countries of citizenship). It should be noted that no harmonisation of migration data was envisaged for the project. As a result, the predictions of flows obtained for various countries were virtually incomparable. Rather than providing comparisons of the flows forecasted for different countries, we concentrated on the consistency of the forecasted

immigration volumes and population stocks for a given country, within its own demographic balance equation.

9.2.3 *Statistical models*

The statistical tools employed included the following simple time series models.

- For the total immigration flows: the autoregressive model (AR) of order 1, stationary (i.e. exhibiting stable characteristics over time) and, for the sake of comparison, the so-called random walk with drift (not stationary). The models also allowed for the capture of the changing variability of migration (known as stochastic volatility).
- For modelling the most important sources of immigration: vector autoregressive (VAR) models of order 1 were applied to log-ratio transformed variables, ensuring the summation of the shares to the total.
- Additionally, for assessing the impact of the selected demographic and economic variables, VAR models of order 1 were used.

In general, the AR models assume that the number of immigrants in a given year depends on their number one year before and on some purely random change that reflects all the uncertainty (on average, this change is assumed to be zero). Generally, in an AR process, the number of immigrants in a given year is the sum of all these random components from the very beginning of the process (in practice, it is the first observation in the series). This specification allows us to describe three different patterns of behaviour: stationary, non-stationary (random walk) and explosive.

In the case of a stationary AR process, the older the random component, the less influence it has on today's immigration flow. The expected value of the immigration flow is then constant, which means that the process is, on average, stable over time. Any shock (e.g. radical liberalisation of the immigration policy) will change the number of immigrants only temporarily (say, for a few years) and, in the long run, this number will return to the average observed before the shock. Hence, the forecasting based on an AR model is relatively precise. The addition of a trend term (e.g. linear, logistic or logarithmic) introduces an increasing (or decreasing) deterministic tendency along which the process develops.

In a random walk model (which is a special case of AR), all random components have equal influence on today's value. As a result, the process is non-stationary, which means that it is unstable in time and the precision with which we are able to predict the number of immigrants becomes smaller with the forecasting horizon (e.g. such a model may be used to describe a situation where we are unable to predict the behaviour of immigration flows after a liberalisation of the policy in the long-run). A drift term

(i.e. a constant value added in each step) introduces increasing (or decreasing) tendencies along which these characteristics hold.

The last possibility, as far as AR models are concerned, is the ‘explosion’ of the process under study. In this scheme, the more distant the past changes are, the more influence they have today. In the current application, this would lead to the number of immigrants growing exponentially to achieve very large values. After some period of time, the predicted values would become simply implausible.

Finally, the VAR models are multidimensional generalisations of the AR model, including not only the variable of interest (here, migration flow or share of flow from a specified country), but also other variables, such as the considered migration determinants or shares of flow from other countries. In the VAR models, the potential instantaneous and lagged relationships among all the variables under study are analysed jointly.

9.2.4 *Prior expert knowledge elicitation*

Prior knowledge, a very important element of analysis in the Bayesian framework, was obtained by means of a Delphi survey from country-specific migration experts selected by the national teams participating in the IDEA project.

The Delphi survey is a technique that acquires anonymous opinions and judgements from selected respondents and provides a means of reaching an informed consensus among them. Originally, the usual applications of the Delphi survey involved forecasting. In the current study, we used it as a tool for obtaining expert knowledge that served only as part of the input into the model that produced the forecasts. One of the most important characteristics of the Delphi survey is that the answers are obtained iteratively. After each round, an anonymised statistical summary of the preceding round is provided to the respondents. In the consecutive rounds, they can reformulate their judgements and views and reach an agreement.

There are two major issues that require attention during preparation of the Delphi survey: proper formulation of the questions and the selection of the experts. The questions should take into account the hints provided by cognitive psychology, so as to ensure correct interpretation and unambiguous answers. Apart from this problem, two additional problems were identified in this study.

First, the questions relate to the characteristics of the immigration process, and were then transformed into the characteristics of the models that served as forecasting tools. It could have been the case, however, that the underlying experts’ model – which led them to formulate such an opinion – was different and could have brought about inconsistencies. The second issue relates to the survey questions, which contain subjective probabilities. The literature on the subject (see e.g. Kadane & Wolfson 1998; Rowe &

Wright 2001) points out that the respondents tend to perceive probabilities as frequencies, and are also overconfident with respect to their beliefs. As a result, they tend to provide exceedingly narrow uncertainty spans. Inconsistent answers can be given to differently formulated questions. Yet another issue is that the group of experts can sometimes comprise specialists untrained in statistics and hence unfamiliar with the terms and concepts used for model formulation.

As far as the selection of experts is concerned, in order to avoid complications, their joint knowledge should be heterogeneous and should cover various aspects of the problem. The group of experts should not be too large or too small, according to J. Scott Armstrong (1985) and Gene Rowe and George Wright (2001).

The anonymous survey in question here was made up of two stages. Answers were obtained from between six and fourteen experts per country. The survey was given to those with various professional backgrounds – from demography, economics and sociology, to political science and law, and public administration – thereby ensuring the heterogeneity of their joint knowledge on migration. The questions focused on the characteristics of immigration processes in the period between 2007 and 2025: namely, trends, types of processes and variability, as well as the uncertainty of the processes' future developments. It should be stressed that there was no explicit question concerning the stage of the migration cycle of a particular country, which was meant to largely emerge from the qualitative analysis of the results.

The experts were also asked about the impact of economic variables (GDP per capita growth and unemployment rates in the receiving countries) and demographic variables (natural population growth rate and share of the productive-age population in the receiving countries) on immigration. Additionally, they were asked to identify three (in the case of Poland, two) of the most important sources of immigration to their countries in the future.

The answers were transferred to a form required by the model (e.g. expressed as prior probability distributions or deterministic trends) and then combined with data on immigration flows produced by the national statistical institutions (mostly available in the Eurostat database) in order to produce forecasts.

9.3 Bayesian forecasting of immigration: Results *(Wiśniowski and Bijak)*

A key aim that accompanies the interpretation of Bayesian forecasting results is measuring the change in expectations, as formulated in subjective expert knowledge, due to quantitative and objective, yet still inherently flawed data.

As far as the underlying forecasting model is concerned, the data exhibited a clear preference for the non-stationary – and thus hardly predictable – random walk process in most of the countries. This conclusion is in line with that of several other studies (for more discussion, see e.g. Pijpers 2008).

The predictions are presented in terms of central tendencies (medians) with uncertainty spans, which are based on symmetric quantiles of predictive distributions. Hence, the 50 per cent ranges are based on the predictive quantiles (lower and upper), the 80 per cent ranges on the quantiles of rank 0.1 and 0.9, and the 90 per cent ranges on the quantiles of rank 0.05 and 0.95. These uncertainty assessments reflect our conclusions, based on expert knowledge and statistical data, that future migration inflows will fall within a given interval with a pre-defined probability. Of special note is the fact that the probability of migration not being higher than the upper values of the 50 per cent, 80 per cent and 90 per cent intervals is 0.75, 0.90 and 0.95, respectively.

9.3.1 *Forecasting total immigration flows*

Examples of quantile-based ex ante assessments of predictive uncertainty for Poland and Italy are illustrated in Figure 9.1. The forecasts for Italy are characterised by a relatively large level of uncertainty. The median expectations of immigration flows follow historically increasing tendencies: from 305,000 in 2005, to 370,000 in 2016, to 434,000 in 2025. Broad 50 per cent intervals indicate large uncertainty levels inherent in the expected future flows. The upper bound amounts to 839,000 people in 2016 and 1.4 million in 2025. These values should be thought of as only barely plausible. We thus advise forecast users to interpret such extreme values only as an indication of orders of magnitude. There are two main reasons for such high levels of uncertainty: first, the steadily increasing inflows of immigrants in the past; second, prior input into the models based on experts' expectations of the explosive patterns of future immigration to Italy. These factors, together with the data that were affected by regularisations in the past, suggest that these outcomes should be considered an indication of the extremely high levels of uncertainty surrounding future immigration inflows.

Compared to Italy, the inflows forecasted for Poland are relatively moderate and stable. The predictive median tendency indicates an increase in permanent immigration to Poland from an initial 15,000 people in 2007, to 28,000 in 2016, to 53,000 people in 2025. The 50 per cent intervals are relatively narrow, yet exhibit an increasing predictive uncertainty: between 16,000 and 54,000 people estimated to immigrate in 2016 and between 21,000 and 158,000 in 2025.

Figure 9.1 *Immigration to Italy and Poland*

Note: The frame indicates a ten-year forecast horizon.
Source: Data until 2006 (Italy) or 2007 (Poland): Eurostat and NSIs with own re-calculations; forecast: own computations

The forecasts for the rest of the countries are presented later (left-hand side, Figure 9.3). The general conclusion is that immigration flows are expected to increase in all the countries (in terms of the median variant). Although uncertainty appears to be growing over time in all the cases, its magnitude differs among different countries. The widest uncertainty spans are observed for Austria, the Czech Republic and Italy, whereas the narrowest ones are visible for France and Hungary. In all the cases – including Italy and Poland though with the exception of Portugal – incorporating expert knowledge increased the uncertainty of the forecasts. This result may stem from the fact that experts were knowledgeable about past events not reflected by the data and, based on that knowledge, they formulated their opinions and judgements for the future and thereby increased data-based uncertainty.

9.3.2 *Forecasts for the top source countries of immigration*

Figure 9.2 presents the median forecasts of cumulative source country shares (cumulative shares of migrants from countries of previous residence within immigration totals), again for Italy and Poland. For the sake of clarity, the uncertainty spans were omitted here. However, the predictive uncertainty of the shares appeared to be so high that the analysis of different combinations of quantile-based predictions became meaningless. The forecasted numbers of migrants coming from major source countries are conditional upon the median values of the predicted shares (presented in Figure 9.2) and overall immigration volumes (presented in Figure 9.1).

For Italy, the three most important immigrant countries of origin were Romania, Albania and Morocco. The available data series were very small, covering only ten years (1995-2004). During the last observation year, out of the total registered inflow of 415,000 people, 15 per cent came from Romania (64,000), 9 per cent from Albania (37,000) and 7 per cent from Morocco (31,000). According to the median forecasts, the share of inflows from Morocco is projected to remain stable, at just below 8 per cent, throughout the forecast horizon – which would correspond to 28,000 immigrants in 2015 and 34,000 in 2025. At the same time, the share of inflows from Romania would rapidly decline to about 10 per cent and remain relatively constant; while the inflow from Albania is thought likely to stabilise at just around 12 per cent starting in 2008. In absolute terms, however, these inflows would follow an increasing trend: from Romania, 38,000 people in 2015 and 45,000 in 2025; from Albania, 44,000 in 2015 and 52,000 in 2025.

For Poland, experts identified two of the most important sources of future inflows as Ukraine (immigration) and the UK (return migration). In 2007, out of 15,000 immigrants newly registered for permanent residence, 777 people (5 per cent) came from Ukraine and 3,900 (26 per cent) from

Figure 9.2 *Immigration to Italy and Poland, cumulative source country shares, median forecasts*

Notes: Italy data until 2004, forecast for 2005 to 2025

Rest refers to the rest of the world.

Source: Data until 2004 (Italy) or 2007 (Poland): Eurostat and NSIs with own recalculations; forecast: own computations

the UK. The median forecasts, prepared on the basis of 1990-2007 data, indicate that the share of immigrants from Ukraine would first increase to 10 per cent in 2008 and then exceed 11 per cent in 2013, only to decline below 10 per cent in 2016 and to ultimately 6 per cent in 2025. Return migration from the UK would follow a similar pattern: from below 15 per cent in 2008, to a peak of over 18 per cent in 2014, to below 17 per cent

in 2016, down to below 10 per cent in 2025. Both these trajectories are the result of including experts' opinions in the forecast. The shares for 2016 would correspond to 2,800 immigrants from Ukraine and 4,800 thousand return migrants from the UK or around 28,200 immigrants altogether. For 2025, the respective numbers would amount to 3,300 for Ukraine and 5,100 for the UK, out of the total 53,100 permanent immigrants.

Figure 9.3 (see right-hand side) presents the forecasts of cumulative source country shares for the other countries. As mentioned earlier, the predictive uncertainty levels were too high to draw any meaningful conclusions from them. The forecasted shares are aligned with the trends suggested by the experts. For instance, in Austria the future shares of

Figure 9.3 *Immigration (left) and cumulative source country shares (right) for Austria, Czech Republic, France, Hungary and Portugal*

Notes: The frame indicates a 10-year forecast horizon (left), and the vertical lines mark the end of the data series (right).

Sources: Eurostat and NSIs with own recalculations; forecast: own computations

immigrants arriving from former Yugoslav countries and from Turkey are expected to increase. For France, an increase in the share of Chinese immigrants is predicted, as is the share of natives of Serbia and Montenegro in inflows to Hungary.

9.3.3 Impact of economic and demographic variables

An analysis of the impact of additional economic and demographic variables demonstrates the emergence of a clear pattern. The two macroeconomic covariates examined (GDP growth rates and unemployment rates in the receiving countries) appeared to have hardly any influence on migration, even when data were supported by expert judgement. The only exceptions were economic growth in France (instantaneous impact) and Portugal (a

one-year time lag). That result does not confirm the hypotheses of Heinz Fassmann and Ursula Reeger, put forward in chapter 3. However, it should be emphasised that our analysis concerned total immigration, not net migration, which was assumed to be influenced by GDP growth. Secondly, unemployment rate was used in the forecasts as a proxy covariate for a labour market situation and it might have not reflected the overall situation of a given country. Alternatively, the impact of the labour market situation could have been already captured by the demographic covariates, especially the productive-age population in the destination country. A tentative conclusion might also be that the available data generally do not allow for formal inferences to be drawn about the interrelations between immigration and its two most important (e.g. according to Jennissen 2004) economic determinants.

The impact of the two demographic variables (natural population growth rates and the share of the productive-age population in the destination countries) was significant in most cases, except for Portugal. For Austria, significance applied to the lagging share of the productive-age population only, while for the five remaining countries various combinations of lagging and/or instantaneous influence of both demographic covariates were found. A tentative conclusion is thus that the two demographic factors indeed play some role in shaping immigration flows.

Despite the explanatory potential of the demographic migration determinants in this study, the conditional forecasts they helped generate remain unsatisfactory from a policymaking point of view. In cases with the 'best' test results (France and Italy), the forecasts quickly reached implausible values with high uncertainty spans. It thus seems that, even if the demographic covariates have a significant impact, the price to pay for including them in the forecasts is extremely high predictive uncertainty.

In many cases, despite the variables' significant impact on migration inflows, the interpretation of the model parameters was counterintuitive. This means that the positive outcomes of significance tests alone do not guarantee reliable scenario-based predictions and that the latter need to be considered with all due scepticism.

9.3.4 *Migration cycle in the light of the forecasting exercise*

It is difficult to draw a meaningful conclusion concerning confirmation or disproof of the migration cycle stage of a given country, as put forth in previous chapters, just by looking at the results of the forecasting exercise presented above. There are two main reasons for that. The first and most important is the exclusion from the analysis of the emigration flows from the countries under study due to reasons described in section 9.1, primarily the lack of data or their poor, even worse than for immigration, quality. This automatically precludes using immigration forecasts alone for

falsifying or validating the migration cycle theory outlined in previous chapters. Hence, all results should be examined with caution, only as potential symptoms indicating the possible position of a given country in the cycle. Secondly, the differences between definitions of migrants in the countries under study, the quality of the available data and their scarcity, as outlined in section 9.1, additionally hampered comparisons between the analysed countries.

Let us consider, for example, the case of Austria, an ‘old’ immigration country. In chapter 3 of this volume, Fassmann and Reeger state that since 1994 Austria has been in the last stage of the cycle, i.e. the adaptation phase. However, the results of the forecasting exercise do not seem to support that thesis unequivocally. The median trend of immigration is increasing, reaching more than 200,000 migrants by 2025, with a very wide uncertainty span. Should the adaptation phase hypothesis hold, this would have to be balanced with increasing outflow (due to e.g. increased mobility within Europe) in order to maintain the predicted stable net migration. Another speculative explanation is a return to the initial stage of migration, as suggested by Godfried Engbersen in chapter 4 in this volume.

The comparison of the four Southern European countries is hindered by not having Greece and Spain in the analysis. For Italy and Portugal, median tendencies and uncertainty spans (relative to the migration levels) are similar. However, these two countries are distinct in their migration context, as they apply different definitions of migrants and duration-of-stay criteria, use different migration policies and have different colonial histories, as described in chapters 5 through 7. This further obstructs the comparison of their position in the migration cycle.

In the case of ‘future’ immigrant countries – the Czech Republic, Hungary and Poland – the results are ambiguous. All three share the same source country of immigrants suggested by the experts: Ukraine. It may be expected that flows from Ukraine will be largely comprised of labour migration. This would be in line with the historic experiences of the ‘old’ migration countries, as stated by Okólski in chapter 1 and Joaquín Arango in chapter 2 of this volume. On the other hand, there are many differences between the results for these three countries, which show that there is a considerable heterogeneity of migration patterns within this group.

In chapter 8 in this volume, Dušan Drbohlav concluded that the Czech Republic and Hungary are in their transition period of the cycle, yet considerably further on the path to becoming mature immigration countries than Poland. This seems to be supported for Hungary, whose median tendency is only slightly increasing with comparatively very narrow uncertainty, suggesting the stability of future migrant inflows and relative ‘maturity’ of the migration cycle. On the other hand, immigration to the Czech Republic is characterised by increasing tendency and very wide uncertainty. However, it should be noted that the data for this country were

subject to definitional change in 2001. Drbohlav's conclusions concerning Poland being at the early stage of the transition period of the migration cycle seem to be approved by the experts. They suggested that flows would start to increase even faster until around 2016, when the increase would start to slow down and immigration will stabilise. It must be kept in mind that forecasting for Poland was carried out using official data based on the permanent duration of stay criterion. This might have precluded from the analysis all temporary stay and labour migrants, i.e. liquid migration as outlined by Engbersen in chapter 4 and Drbohlav in chapter 8. Moreover, interpolation of the recent post-enlargement emigration patterns, as presented by Okólski in chapter 1, may prevail over even increasing immigration and further aggravate the positioning of Poland in the migration cycle.

9.4 Conclusions

(Wiśniowski and Bijak)

The outcomes of the forecasting exercise carried out within the IDEA project can be summarised as follows (for details, see Bijak & Wiśniowski 2009, 2010). First, migration appears to be a hardly predictable demographic process, strongly in accordance with previous research findings (see e.g. Pijpers 2008). Moreover, assuming stationarity of immigration, which is in fact a non-stationary variable, seems reckless from a methodological standpoint, as it can bring about very serious forecasting errors and lead astray forecast users. A recent example is the study prepared for the British Home Office (Dustmann, Casanova, Fertig, Preston & Schmidt 2003), in which the stationarity of the underlying model was assumed explicitly (Dustmann et al. 2003: 28, 68). The forecasted volumes of immigrants to the UK after the 2004 EU enlargement were underestimated by more than an order of magnitude (see IPPR 2008).

Second, precise forecasting of the exact values of immigration flows is virtually impossible, especially in light of the above. Predictive uncertainty, however, can be accounted for by using the stochastic approach (Keilman 1990; Lutz et al. 2004). The randomness can then be quantified and expressed in probabilistic terms. It should be stressed that the stochastic methodology of migration forecasting does not impact the nature of the process itself. It simply addresses the fact that migration is difficult to predict. Consequently, the use of deterministic models is not a wholesale solution. Without an appropriate warning, it serves only to blur the picture, giving forecast users and decision makers the false impression that migration can be predicted.

As for the subjective information obtained using the Delphi survey – which served as an input for the forecasting models in the current study –

its impact on final outcomes was diverse. Expert knowledge appeared to play a very important role in the estimation of the model parameters, especially with respect to the forecasts' precision. Model selection, and thus determining the very nature of the process, seemed almost completely unaffected by the experts' opinions. Given the prevailing selection of random walk models – often independent from experts' expectations – we argue that the uncertain and barely predictable character of migration flows is not just characteristic of a particular forecasting model, but also a fundamental feature of migration. Again confirming earlier suggestions (e.g. Holzer 1959 or Keyfitz 1981), this implies that migration should be forecasted for only a relatively short period: five to ten years at most. The uncertainty and inherent unpredictability of migration flows grows over time, mainly due to its non-stationary characteristics. Using horizons of more than a decade into the future will necessarily lead to very broad uncertainty spans, which will ultimately constitute a meaningless piece of information for policymakers.

Finally, the impact of migration covariates proved to be difficult to quantify. Due to the brevity of the time series and hardly predictable nature of migration itself, the economic variables' impact, if any, was very limited. On the other hand, despite the significant impact of demographic covariates, in many cases the counterintuitive results of the parameter estimation and subsequent largely meaningless forecasts rendered the interpretation of the outcome dubious at best. It should be borne in mind that these results may stem from the usage of an inexhaustive series of data or from the narrow scope of measures chosen for the study. That is, the uncertainty of the forecasts had at least three sources of uncertainty: of the migration processes, of the covariates and of the interactions and interrelations among them. These outcomes support earlier suppositions that additional, theory-based determinants of migration – although extremely helpful in explaining the processes *ex post* – are of very limited use when it comes to forecasting (see Kupiszewski 2002). The results of the current study seem to indicate that, while demographic factors play a significant role in shaping migration processes, they simultaneously fail to contribute to more meaningful and precise migration forecasts – even despite the very important support of expert judgement.

The most important implication of the above conclusions for would-be users of forecasting models is that migration is indeed a vastly uncertain phenomenon. This may be seen as an undesirable feature. However, the degree of uncertainty can nevertheless be quantified and evaluated using statistical data and additionally enhanced by expert knowledge. This sort of assessment is in itself an important piece of information for policymakers.

Migration forecasts such as the ones presented in the current study may serve as an informational basis for political leaders and for informed and rational migration policymaking, as outlined by Magdalena Lesińska in

chapter 11 in this volume. Their actions will strongly depend on the nature, strategic objectives and constraints of a particular decision-making problem, as well as on their own preferences and views. A proper assessment of the consequences of possible underestimation or overestimation of future migration flows is crucial.

Our results should be viewed through the lens of the overall findings of the IDEA project. In particular, and with respect to interpreting forecast results, the phase of the migration cycle in a particular country, as described in chapters 1 and 2, along with the specificities of the countries belonging to one of the three groups ('old', 'new' and 'future' immigration countries), should all be taken into account. However, it must be emphasised that the forecasting concerned only immigration flows, which shows only half of the picture of future migration, thus all results can only serve as symptoms indicative of a country's position in the migration cycle. Data-related issues, such as differences in definitions and collection schemes, regularisation programmes and the unending problem of estimating the number of illegal immigrants (see chapter 5), should also be acknowledged. Moreover, identification and quantification of the migration cycle in the forecasting exercise is affected by the definition of migrants in the specific country, as well as the country-specific characteristics, economic and demographic. Thus, formal comparisons between, for instance, 'old' and 'future' immigration countries were hardly possible.

As far as the impact of social, economic and demographic variables is concerned, it should be borne in mind that their influence upon migration may be subtle and not immediate. The ambiguous results and data-related shortcomings described above, along with the uniqueness of the experiences of the particular countries (e.g. in Central and Eastern Europe; see chapter 8) indicate that the detection of these impacts and interrelations can be strongly hampered, as, of course, can be future implications for migration movements.

Finally, the authors strongly recommend that forecast-based, mature migration policies should be crafted with extreme caution. The central tendency (median) trajectories are well suited only for situations in which an underestimation of future migration would entail the same costs as an overestimation of the same amount; they may not work well at all for other decision-making scenarios or to inform other subjective priorities of migration policymakers. In some cases, the underestimation of future migration inflows could be more costly than their overestimation, as, for example, in assigning budgets to migrant integration programmes. In such cases, it would seem rational to use the above-median variants, if only in order to err on the side of caution. In other instances, however, overestimation can be more problematic – as in the case of the inflows required to fill local labour market shortages. The use of below-median forecast variants would be recommended in these cases. Irrespective of which variants are used,

however, the level of predictive uncertainty must always be kept in mind. Above all, researchers and policymakers, alike, must remember that simply using the median (or any other) trajectories and treating them as universal, deterministic predictions is both unwise and imprudent. They will almost certainly never come true.

Notes

- 1 In this chapter, uncertainty will be understood as a state of not having enough knowledge about a process or phenomenon now or in the future.
- 2 Forecasting migration was impossible for Greece, where no adequate migration data exists. For Spain, the unavailability of a second round of the Delphi questionnaires also rendered the forecasting exercise impossible.

References

- Alho, J.M. & B.D. Spencer (2005), *Statistical demography and forecasting*. New York: Springer.
- Alho, J.M. & B.D. Spencer (1997), 'The practical specification of the expected error of population forecasts', *Journal of Official Statistics* 13: 203-225.
- Armstrong, J.S. (1985), *Long-range forecasting* (2nd edition). New York: John Wiley.
- Bijak, J. (2010), *Forecasting international migration in Europe: The Bayesian approach*. Springer Series on Demographic Methods and Population Analysis 24. Dordrecht: Springer.
- Bijak, J. & A. Wiśniowski (2009), 'Forecasting of immigration flows until 2025 for selected European countries using expert information', IDEA Working Paper 7. www.idea6fp.uw.edu.pl/pliki/IDEA_Working_Paper_No.7.pdf.
- Bijak, J. & A. Wiśniowski (2010), 'Bayesian forecasting of immigration to selected European countries by using expert knowledge', Forthcoming in *Journal of the Royal Statistical Society, Series A* 173.
- Bayes, T. (1763), 'An essay towards solving a problem in the Doctrine of Chances', *Philosophical Transactions of the Royal Society* 53: 370-418.
- Dustmann, C., M. Casanova, M. Fertig, I. Preston & C.M. Schmidt (2003), 'The impact of EU enlargement on migration flows', Home Office Report 25/03. London: Home Office.
- Frydman, R. & M.D. Goldberg (2007), *Imperfect knowledge economics: Exchange rates and risk*. Princeton: Princeton University Press.
- Holzer, J.Z. (1959), *Prognoza demograficzna Polski na lata 1960-1975 wg województw* [Demographic forecast for Poland for 1960-1975 by voivodships]. Warsaw: PWG.
- IPPR (2008), *Floodgates or turnstiles? Post-EU enlargement migration flows to (and from) the UK*. London: Institute for Public Policy Research.
- Jennissen, R. (2004), *Macro-economic determinants of international migration in Europe*. Amsterdam: Dutch University Press.
- Kadane, J.B. & L.J. Wolfson (1998), 'Experiences in elicitation', *The Statistician* 47 (1): 3-19.
- Keilman, N. (1990), *Uncertainty in national population forecasting: Issues, backgrounds, analyses, recommendations*. Amsterdam: Swets & Zeitlinger.
- Keyfitz, N. (1981), 'The limits of population forecasting', *Population and Development Review* 7 (4): 579-593.

- Kupiszewska D. & B. Nowok (2008), 'Comparability of statistics on international migration flows in the European Union', in J. Raymer & F. Willekens (eds.), *International migration in Europe: Data, Models and Estimates*, 41-71. Chichester: John Wiley and Sons.
- Kupiszewski, M. (2002), 'Modelowanie dynamiki przemian ludności w warunkach wzrostu znaczenia migracji międzynarodowych' [The role of international migration in the modelling of population dynamics], *Prace Geograficzne* 181. Warsaw: Institute of Geography and Spatial Organization, Polish Academy of Sciences.
- Lee, R. & S. Tuljapurkar (1994). 'Stochastic population forecasts for the United States: Beyond high, medium, and low', *Journal of the American Statistical Association* 89: 1175-1189.
- Lutz, W., W.C. Sanderson & S. Scherbov (1996), 'Probabilistic population projections based on expert opinion', in W. Lutz (ed.), *The future population of the world: What can we assume today? Revised edition*, 397-428. Laxenburg: IIASA; London: Earthscan.
- Lutz, W., W.C. Sanderson & S. Scherbov (eds.) (2004), *The end of world population growth in the 21st century: New challenges for human capital formation & sustainable development*. London: Earthscan.
- Nowok, B. & D. Kupiszewska (2005), 'Official European statistics on international migration flows: Availability, sources and coverage', CEFMR Working Paper 5/2005.
- Nowok, B., D. Kupiszewska & M. Poulain (2006), 'Statistics on international migration flows', in M. Poulain, N. Perrin & A. Singleton (eds.), *THESIM: Towards Harmonised European Statistics on International Migration*, 203-231. Louvain-la-Neuve: Presses Universitaires de Louvain.
- Öberg, S. & A.B. Wils (1992), 'East-West migration in Europe: Can migration theories help estimate the numbers', *Popnet* 22: 1-7.
- Pijpers, R. (2008), 'Problematising the "orderly" aesthetic assumptions of forecasts of East-West migration in the European Union', *Environment and Planning A* 40 (1): 174-188.
- Rees, P.H. & I. Turton (1998), *Investigation of the effects of input uncertainty on population forecasting*. Paper for the 3rd International GeoComputation Conference, Bristol, UK, 17-19 September.
- Rowe, G. & G. Wright (2001), 'Expert opinions in forecasting. Role of the Delphi Technique', in J.S. Armstrong (ed.), *Principles of forecasting: A handbook of researchers and practitioners*, 125-144. Boston: Kluwer Academic Publishers.
- Willekens, F. (1994), 'Monitoring international migration flows in Europe: Towards a statistical data base combining data from different sources', *European Journal of Population* 10 (1): 1-42.