European Population Conference 2006. Liverpool, United Kingdom, 21–24 June 2006

Session 904: Projections

Bayesian Model Averaging in Forecasting International Migration[*]

Jakub Bijak

Central European Forum for Migration Research (CEFMR)[◊]

Abstract

The paper is devoted to the development of Bayesian methodology for international migration forecasting. Its aim is to present a statistical framework for (1) a formal selection of a forecasting model, and (2) combining knowledge from various competing models in order to obtain averaged forecasts. The main benefit of pooling forecasts yielded by different models is the reduction of forecast error resulting from the uncertainty of model specification. The Bayesian framework ensures the formal status of applied statistical tools in addressing the uncertainty issues, allowing at the same time for an explicit incorporation of subjective expert opinion in the model selection process, as well as in the models themselves. In the paper, a theoretical discussion is supported by an empirical application of the presented methodology to a forecast of migration flows between Germany and three European countries: Poland, Italy, and Switzerland, for the period 2005–2010. The analysis employs macro-level data from the population registers of the countries under study, and is restricted to forecasting models based on simple stochastic processes.

[*] Paper prepared within the framework of the project "Forecasting international migration in Europe: a comparison of traditional methods and the Bayesian approach", financed by the Foundation for Population, Migration and Environment (BMU-PME) from Zurich. The author also benefited from the Stipend for Young Scientists of the Foundation for Polish Science (FNP). Credits go to Prof. Jacek Osiewalski for the original idea underlying the paper and for helpful remarks, as well as to Prof. Marek Kupiszewski for his comments to the earlier draft. All remaining errors are mine.

[◊] Mailing address: CEFMR, ul. Twarda 51/55, 00-818 Warsaw, Poland, tel. +48 22 697 88 34, fax +48 22 697 88 43, e-mail: j.bijak@cefmr.pan.pl, web: www.cefmr.pan.pl.

1. Introduction

Forecasting[1] international migration is an important, yet very difficult research task, characterised by highest errors among the predictions of all components of demographic change (National Research Council, 2000: 156–157). Reasons for this state of affairs include: (1) impossibility of creating a comprehensive migration theory given the multi-dimensionality and complexity of the phenomenon, (2) difficulties in operationalising the existing theoretical framework of migration, (3) uncertainty of potential explanatory variables, (4) ignoring forced migration and policy elements in the forecasts, and (5) poor data quality (Kupiszewski, 2002).

Nevertheless, a need for reliable forecasts of international migration is becoming increasingly significant despite high levels of error. Population movements are currently gaining in importance given the decreasing impact of natural change on population dynamics in the developed countries (cf. van der Gaag and van Wissen, 1999). Hence, migration forecasts constitute an essential part of population predictions, which are crucial for any aspect of socio-economic planning in a longer term. Moreover, migration forecasts themselves are of interest for the policy makers from a purely operational point of view, with respect to the numbers of migrants, or the short-term impact of migration for example on labour markets. Improving forecasts can contribute to better policy decisions, which in turn may have significant consequences for the societies, not limited to their economic aspects (Ahlburg *et al.*, 1998: 192).

Some level of migration forecasting error is always inevitable, as any inference about the future is made under uncertainty, which is further discussed in Section 2.1. Nevertheless, there have been several propositions about the ways to improve the accuracy of forecasts. Among them, the study of Rees *et al.* (2001) suggests that methodological advancements in migration forecasting coincides with visible declines in the *ex-post* prediction errors. This constitutes a rationale for the current research, which aims at contributing to the development of forecasting methodology in order to improve the accuracy of international migration predictions. In particular, I follow the idea of Ahlburg (1995), and Ahlburg and Lutz (1998), who advocated the benefits from combining (averaging) forecasts in order to reduce their errors[2]. It has to be noted that averaging of various deterministic and model-based forecasts of net international migration

[1] With due respect to the demographers' dispute on using the terms *forecast* and *projection*, I apply the former one throughout the paper. Keilman (1990: 7) noted that a forecast is an unconditional result of the process, in which "based on current scientific insights, a forecaster gives his best guess of what the future […] will be". In the probabilistic context, such a 'best guess' is produced by the assumed stochastic model. Although it is sometimes useful to further distinguish *forecasts* and *predictions* (Keilman, 1990: 8), I use these terms interchangeably.
[2] The idea originates from the work of Armstrong (1985).

has been done for the 15 countries of the 'old' European Union in the recent population predictions of the Eurostat (2005), following the methodology of Lanzieri (2004).

The objective of the current paper is to present a Bayesian framework for formal selection of a forecasting model, and for producing averaged forecasts based on various models. The Bayesian approach ensures the formal status of applied statistical tools in addressing the uncertainty issue, allowing at the same time for an explicit incorporation of subjective expert opinion in the models. The main benefit of averaging forecasts is the reduction of forecast error resulting from the uncertainty of model specification.

Apart from the Introduction, the paper comprises of four sections. Section 2 presents a short discussion on the uncertainty and subjectivity issues in migration forecasting. This background aims to justify the application of the Bayesian approach, a brief introduction into which is also offered. The discussion is supplemented by an overview of existing Bayesian forecasts of international migration. The possibilities offered by the Bayesian methodology are further corroborated in Section 3, containing a description of the foundations of the Bayesian model selection and averaging techniques.

The theoretical discussion is supported in Section 4 by an empirical application of the presented methodology to a forecast of long-term migration between Germany and three European countries: Italy, Poland, and Switzerland, for the period 2005–2010. The analysis is based on aggregate data from population registers of the countries under study, and is restricted to forecasting models derived from simple stochastic processes. Finally, Section 5 contains the main conclusions of the study, as well as some suggestions for further research in this area.

2. Uncertainty and subjectivity in migration forecasting and in Bayesian statistics[3]

2.1. Uncertainty and subjectivity in migration forecasting

Uncertainty about the future values of the forecasted phenomena is an immanent feature of every prediction. With respect to the sources of uncertainty in population (and thus also migration) forecasting, Keilman (1990: 19–20) distinguishes seven types of possible errors. Three of them are related to the measurement issues (errors in observed trends, in jump-off data, and rounding errors), one to the randomness of the parameters of the forecasting model, and further three to the errors in the forecasts of exogenous variables, possible future discontinuity in trends,

[3] Section 2 draws on the earlier paper of the author (Bijak, 2005).

and to the improper model specification. Rees and Turton (1998) observed that uncertainty in population forecasting is usually dealt with in a number of ways:

1. Ignored, by constructing single-variant deterministic forecasts;
2. Included, but not quantified in terms of probability, by developing multi-variant scenarios (conventionally: baseline, high, and low), which is often done by the national statistical offices, United Nations (2005) and Eurostat (2005);
3. Accommodated within a stochastic approach, which quantifies uncertainty in terms of probabilities of future events. Keilman (2001) distinguishes three types of such forecasts: extrapolation of time series (de Beer, 1990; Lee and Tuljapurkar, 1994; Keilman *et al.*, 2001), propagation of historical forecast errors (Alho, 1990; National Research Council, 2000), and probabilistic projections based on the expert judgement (Lutz *et al.*, 1996, 2004).

In international migration forecasting, all three possibilities are explored. Deterministic forecasts are often the outcome of various surveys or Delphi analyses. The multi-variant scenarios are most frequently used within the framework of cohort-component or multiregional models of demographic dynamics. Stochastic forecasts of international migration are usually either the outcome of econometric models, or time series extrapolation, and there are only a few examples of forecasts applying the Bayesian approach.

From the probabilistic point of view the deterministic and the scenario-based approaches are methodologically inconsistent. A deterministic forecast formally has a probability of occurrence equal zero under any continuous distribution reflecting uncertainty. The scenario approach is criticised for not providing the information, what are the expected *ex-ante* chances that the phenomena under study will be actually observed between the low and high variants (Lutz *et al.*, 2004: 19). Moreover, the scenario selection (baseline, high or low) often implicitly assumes the presence of a single common underlying factor for all variables (fertility, mortality, migration), and all regions under study. The aggregate effects are thus based on the assumption of a perfect correlation between the variables and regions, which is not formally examined, and very often not true (National Research Council, 2000: 191–192).

Unlike in the former two cases, in the stochastic approach uncertainty is quantified in terms of probability. Due to this fact, as well as to the methodological consistency of the approach, many authors argue that the probabilistic forecasting in demography will become increasingly more in use in the future (Lutz and Goldstein, 2004: 3–4).

As it has been noted by Pittenger (1978), all population forecasts and projections rely heavily on the expert judgement. Uncertainty inherent in the forecasts requires making use of many subjective elements, including the choice of the forecasting model, its assumptions, forecasts of the future changes of exogenous variables and other components of population dynamics, etc. This subjectivity can be either explicitly stated in the forecast, or concealed among the assumptions applied. In either case, it is an inherent element of the selection of a forecasting model, as well as of making assumptions about the future scenarios of demographic change (Gjaltema, 2001). The incorporation of expert judgement in population forecasting is usually not explicitly addressed by the forecasters, with the notable exceptions of, for example, the studies by Alho and Spencer (1985; 2005), Alho (1990), and Lutz *et al.* (1996, 2004).

International migration is a particularly complex and multi-dimensional demographic phenomenon, characterised by a large dose of uncertainty, which ideally should be properly addressed and quantified. The existing methods of migration forecasting include various approaches originating from demography, economics, sociology, geography, etc. Improvement of the forecasting methodology may require combining expertise from various disciplines. However, the subjective and judgemental elements, inevitable in any forecast, should be explicitly visible in the formulation of the model and its assumptions. This is the basic rationale for selecting Bayesian statistics as a framework for forecasting international migration.

2.2. Preliminaries of Bayesian statistics

The Bayesian approach in statistical inference, based on the Bayes Theorem (Bayes, 1763; Laplace, 1812), uses the sample information to transform the *prior knowledge* of the researcher on the phenomenon under study, into the *posterior knowledge*[4]. The former reflects the subjective opinion (belief, intuition) on the subject, without taking observations into account, while the latter is conditional on the sample data. Formally, Bayesian statistics infers on the unknown parameters of the model describing the phenomenon (θ), conditionally on the statistical information (x), both treated as random quantities. This approach is therefore contrary to the traditional sampling-theory statistical methods, for example the Neyman-Pearson theory of hypothesis testing. The latter one is based on the concept of likelihood, i.e. the probability of obtaining a random data sample x, given fixed, yet unknown, model parameters θ.

In terms of probabilities, the Bayes Theorem can be written as:

(1) $$p(\theta \mid x) = \frac{p(\theta) \cdot p(x \mid \theta)}{p(x)} .$$

[4] The Bayesian statistics as a complete inference paradigm originates from the works of Jeffreys (1939), Barnard (1947, 1949) and Savage (1954); its complete theoretical overview is given for example in Bernardo and Smith (2000).

From (1) it follows that the posterior probability $p(\theta | x)$ is proportional to the product of the prior probability $p(\theta)$ and the likelihood of the sample $p(x | \theta)$, where the proportionality factor is the marginal density of the data, $p(x)$.

A key concept in Bayesian statistics, distinguishing it from the sampling-theory paradigm, is subjective probability, independent from the frequency of events under study (Ramsey, 1926; De Finetti, 1937; after Bernardo and Smith, 2000: 3, 15). In this approach, statistical inference can be seen as a decision problem, with strong relations between the concepts of probability and utility (Bernardo and Smith, 2000: 67–81). Interpretation of probability as a measure of belief on the phenomena under study, altered by the observations according to the Bayes theorem, has an advantage in social sciences, where the samples are by nature unrepeatable[5].

Due to the explicitly expressed subjectivism, the Bayesian approach developed in the opposition to the traditional, sampling-theory mathematical statistics. Contemporarily, the attempts to reconcile the two paradigms include the 'objective Bayesianism', assuming no prior information (Bayarri and Berger, 2004), and the pragmatic approach, allowing for choosing the methodology, depending on the nature of the research (Chatfield, 2002).

In the Bayesian approach, an important issue is the selection of the *prior probability distribution* of the estimated parameters, $p(\theta)$, reflecting the knowledge of the researcher, or lack thereof in the case of *non-informative* distributions introduced by Jeffreys (1939). Selection of an informative prior distribution is usually supported by the expert judgement. An analysis of robustness of the results against changes in the prior distribution is an important element of Bayesian inference. A natural outcome of the analysis is the *posterior distribution* $p(\theta | x)$, which can be summarised by its point characteristics (mean, median, etc.), or *credible intervals*, analogous to confidence intervals in the sampling-theory statistics, but without problems and inconsistencies regarding the interpretation of the latter (Jaynes, 1976).

The construction of such an interval for θ, with the posterior probability that θ is actually located within this interval set to equal $1-\gamma$, can be for example performed by determining such k_γ that would satisfy the condition (Silvey, 1978: 203):

(2) $$\int_{\{\theta: p(\theta|x) > k_\gamma\}} p(\theta \mid x) d\theta = 1 - \gamma \, .$$

The formula (2) may concern not only the univariate θ, but also its multivariate generalisation, $\boldsymbol{\theta}$. The interval (or, in multivariate cases, region) (2) contains all *a posteriori* most

[5] A more detailed discussion on the advantages of Bayesian inference in social sciences under non-repeatability of samples is presented by Withers (2002), with focus on applications in human geography.

likely values of θ, so it is called a *highest posterior density* (HPD) interval (region). Moreover, each value of θ outside (2) is less likely than any value from this interval (region). In certain cases, when θ is univariate and the posterior distribution $p(\theta|x)$ is unimodal, the quantiles of rank $\gamma/2$ and $1-\gamma/2$ from $p(\theta|x)$ can be simply taken as the respective lower and upper limits of the Bayesian interval (Bernardo and Smith, 2000: 259–262).

Forecasting in the Bayesian approach is based on the construction of a probability distribution of the vector of future values of the variable under study, $\mathbf{x}^{\mathbf{F}}$, conditional on the vector of past (observed) values, $\mathbf{x}$, and taking into account the posterior knowledge on the parameters of the forecasting model, $\boldsymbol{\theta}$. The predictive probability distribution of $\mathbf{x}^{\mathbf{P}}$ can be calculated according to the following formula:

(3) $$p(\mathbf{x}^{\mathbf{F}} \mid \mathbf{x}) = \int_{\Theta} p(\mathbf{x}^{\mathbf{F}}, \boldsymbol{\theta} \mid \mathbf{x}) d\boldsymbol{\theta} = \int_{\Theta} p(\mathbf{x}^{\mathbf{F}} \mid \boldsymbol{\theta}, \mathbf{x}) \cdot p(\boldsymbol{\theta} \mid \mathbf{x}) d\boldsymbol{\theta}.$$

The predictive density (3) can be interpreted as an average from the conditional predictive distribution $p(\mathbf{x}^{\mathbf{P}} | \boldsymbol{\theta}, \mathbf{x})$, weighted with the posterior probabilities of the parameters (Zellner, 1971: 29). A natural outcome of Bayesian forecasts are the predictive credible regions, which formally reflect the uncertainty of the phenomenon under study.

Bayesian methodology can reduce estimation and prediction errors in such cases, when the prior distribution is informative and consistent with the observations. For non-informative priors, the *ex-ante* errors in one-dimensional problems are often the same as in the traditional maximum likelihood estimation (Bernardo and Smith, 2000: 359). This is important in the small-sample studies (e.g., with population disaggregated by sex, age, regions, etc.), where the prior information has relatively more weight in the posterior result than the observations, unlike in large datasets. The extreme estimates obtained from small-sample data are in this way corrected towards the prior expectations. The same applies to forecasting models based on short time series, where the Bayesian approach is a way to reduce uncertainty, due to the explicitly included prior knowledge – a subjective element of the analysis.

2.3. Existing Bayesian forecasts of international migration flows

The existing examples of international migration models and forecasts based on the Bayesian framework are scarce[6]. Gorbey *et al.* (1999) used vector autoregression models VAR(4) to forecast migration between Australia and New Zealand, both within the sampling-theory and Bayesian paradigms. The authors tested four models based on different vectors of interdependent variables, including net migration rates, growth of the real GDP ratio (or the real income ratio) for the two countries, differences in unemployment rates, country-specific unemployment growth indices, etc. For model coefficients, they used the *Minnesota priors*, with parameters on the first lags of the same variables following *a priori* a normal distribution with the mean equal one, and the parameters of the remaining interrelations – a normal distribution with the mean equal zero. This reflects an *a priori* assumption that the time series of each variable is most likely generated independently by a random walk process. Further, the authors assumed that the variances of the normal priors are based on standard errors estimated from the observations for the variable-specific univariate autoregressive models. As such priors are data-based, this is not a fully Bayesian approach, where prior distributions are specified independently from the data[7].

The econometric study of Brücker and Siliverstovs (2005) also contains some elements of a Bayesian analysis. The authors proposed a partial adjustments model with shares of migrants in particular destination countries changing stepwise towards their long-term equilibrium levels. The model is also based on wage levels and differentials, employment rates, proximity indicators and institutional dummies, while the error term is decomposed into the country-specific effects and the white noise. The Bayesian framework is, however, considered by the authors only as an alternative methodology of estimation, without any mention of the prior distributions used in the analysis, and without the *a posteriori* uncertainty assessment, which elements are both inherent in the Bayesian approach.

An example of a partial departure from the sampling-theory statistical paradigm in demographic forecasting is the concept of 'expert-based probabilistic population projections' developed by Lutz *et al.* (1996, 2004). The method applies subjective expert judgement to set the framework for the stochastic forecasts. In formal terms, the forecasting model assumes random

[6] In general, the Bayesian approach had limited applications in demographic forecasting up to date. One example is a forecast of the Iraqi Kurdish population under limited statistical information (Daponte *et al.*, 1997). Other Bayesian predictions of particular components of population change include fertility (Tuljapurkar and Boe, 1999), and mortality (Girosi and King, 2005). Population flows other than migration have been modeled and forecasted also by Congdon (2000) in his study of moves of patients to hospitals in England. Recently, Alho and Spencer (2005: 244–248, 269–277) have discussed prediction intervals based on the subjective probability, as well as Bayesian forecasting methods for selected time series models, both in the context of population predictions.
[7] I am very grateful to Prof. Jacek Osiewalski for drawing my attention to this problem.

deviations from the mean trajectory of the process under study, formulated *a priori* on the basis of the subjective expert opinion. For global forecasts of international migration, Lutz *et al.* (2004) applied a time-invariant mean trajectory, and assumed that the random component follows the moving average process MA(30). Although this approach is not fully Bayesian, the explicitly expressed subjectivity makes it a hybrid of the traditional and Bayesian methods.

In the presented examples, the forecasting models of international migration either are not fully Bayesian in terms of independence of prior distributions from the data sample, or do not discuss the results in probabilistic terms. Moreover, none of them seems to have so far explored the potential offered by the Bayesian statistical framework in relation to formal model selection and averaging. The current paper aims therefore at filling these gaps.

3. Methodological foundations of Bayesian model selection and averaging

3.1. Bayesian model selection

The Bayesian methodology allows for a formal model selection by comparing the *posterior odds* of different models, given the data, in order to maximally utilise information from the sample of observations[8]. After Osiewalski (2001: 21–22), let M_1, …, M_m be mutually exclusive (not nested) models of the phenomenon under study, adding up to the whole (finite) space of possible models, **M**. Assuming prior probabilities $p(M_1)$, …, $p(M_m)$ for the respective models, the Bayes theorem yields their posterior probabilities, given the data x:

(4)
$$p(M_i \mid x) = \frac{p(M_i) \cdot p(x \mid M_i)}{\sum_{k \in \mathbf{M}} p(M_k) \cdot p(x \mid M_k)} .$$

In (4), $p(x|M_i)$ denotes the marginal density of data in the i-th model, corresponding to $p(x)$ in equation (1).

The posterior probabilities of alternative models i and j can be used for their direct comparison, based on the calculation of the *posterior odds ratio*, R_{ij}:

(5)
$$R_{ij} = \frac{p(M_i \mid x)}{p(M_j \mid x)} = \frac{p(M_i) \cdot p(x \mid M_i)}{p(M_j) \cdot p(x \mid M_j)} .$$

Values of $R_{ij} > 1$ indicate that given the data x and priors $p(M_i)$ and $p(M_j)$, the model i is more likely to accurately describe the phenomenon under study than the model j, while $R_{ij} \in [0, 1)$ gives preference to the model j over i. Some authors argue that it is useful to set threshold values

[8] Technical details are offered in Zellner (1971: 291–317), and Osiewalski (2001: 20–24).

for R_{ij}, denoted O_L and O_R ($O_L < O_R$), so that $R_{ij} < O_L$ provides 'strong evidence' in favour of the model *j*, $R_{ij} > O_R$ for the model *i*, while $R_{ij} \in [O_L, O_R]$ is inconclusive (Hoeting *et al.*, 1999: 385). This issue is also discussed in the next subsection, devoted to the Bayesian model averaging.

Osiewalski (2001: 21–22) noted that the posterior odds ratio may heavily depend on the priors selected for particular models. It is often assumed that all models have equal prior probabilities, especially when there are no arguments for the opposite. In such a case, (5) is reduced to the *likelihood ratio* test, which is also used in the sampling-theory statistics. On the other hand, it may be argued that the formal selection criteria should favour simpler (more straightforward) explanations of the phenomena under study, according to the *Occam's razor* principle[9]. In such a case, the number of the parameters in the *i*-th model, l_i, can be taken as the measure of complexity, and the prior probabilities set in such a way that $p(M_i)$ is proportional to 2^{-l_i} (Osiewalski, 2001: 21). Naturally, there many other ways of assuming the priors on the basis of the expert judgement, which for some reasons may favour one or a group of models.

3.2. Bayesian model averaging (inference pooling)

With respect to demographic forecasting, Ahlburg (1995) noted that there is no firm evidence, whether simple models perform better (or worse) than the more complex ones. He also criticised the quest for the single 'best' forecasting model and suggested that the accuracy of the outcome can be improved by combining various forecasts. In the Bayesian approach, this possibility has been explored within the framework of *inference pooling*, currently known as *Bayesian model averaging*, which allows for merging the features of various predictive models in order to reduce the uncertainty of model specification (Hoeting *et al.*, 1999)[10].

Following the notation used in (3), let $\mathbf{x^F}$ denote the vector of future values of the variable under study. Then, the conditional density of the averaged predictive distribution of $\mathbf{x^F}$ over the model space $\mathbf{M}$, given the data $\mathbf{x}$, denoted by $\bar{p}(\mathbf{x^F} \mid \mathbf{x})$, can be calculated as follows (Hoeting *et al.*, 1999: 383; Osiewalski, 2001: 24):

(6) $$\bar{p}(\mathbf{x^F} \mid \mathbf{x}) = \sum_{i \in \mathbf{M}} p(M_i \mid \mathbf{x}) \cdot p(\mathbf{x^F} \mid \mathbf{x}, M_i).$$

---

[9] According to the Merriam-Webster Online Dictionary, 'Occam's razor' [or 'Ockham's razor'] is "a scientific and philosophic rule that entities should not be multiplied unnecessarily, which is interpreted as requiring that the simplest of competing theories be preferred to the more complex, or that explanations of unknown phenomena be sought first in terms of known quantities" («http://www.m-w.com», accessed on 31 March 2006). A thorough discussion of the 'Occam's razor' notion in the Bayesian context is given by Jeffreys and Berger (1992).

[10] For a discussion of various options on acknowledging the model error in demographic forecasting, see Alho and Spencer (2005: 240–242).

The element $p(\mathbf{x}^{\mathbf{F}} | \mathbf{x}, M_i)$ corresponds to the predictive distribution (3) in the i-th model, while $p(M_i | \mathbf{x})$ is the posterior probability of the i-th model, defined in (4).

The pre-selection of models for the purpose of their averaging can be also done on the basis of posterior odds, according to the 'Occam's razor' principle. In this context, Hoeting *et al.* (1999: 385) corroborate on the idea of the threshold values O_L and O_R for the posterior odds ratio, R_{ij}, as defined in the previous subsection. They label the interval $[O_L, O_R]$ as *Occam's window* and suggest that models that have strong evidence against them (such models i, for which $R_{ij} < O_L$) be disregarded in the pre-selection process. In order to apply the Occam's principle, the values of O_L and O_R should be set in favour of simpler models. For example, if the i-th model has a simpler structure (less parameters) than the j-th model, then $O_L = 1/20$ and $O_R = 1$, as advocated by Madigan and Raftery (1994; after: Hoeting *et al.*, 1999: 385), strongly support the former one in the pre-selection procedure. Alternatively, Raftery *et al.* (1996; after: Hoeting *et al.*, *idem*) proposed setting $O_L = 1/20$ and $O_R = 20$. They found that these values lead to the improved performance of the averaged predictions made on the basis of the ultimately selected models, yet without an explicit reference to the Occam's principle.

4. Empirical application: migration between selected European countries, 2005–2010

4.1. Data: sources, quality, and adjustments

The current study aims at producing Bayesian forecasts of long-term international migration flows between Germany and three European countries: Poland, Italy, and Switzerland, for the period 2005–2010. The forecast horizon ends prior to the expected for 2011 opening of the German labour market for Polish citizens. The analysis employs macro-level data on migration flows and population stocks for 1985–2004 (in the case of flows from/to Poland for 1991–2004, thus after the system transformation). The data used originate from the population registers of the countries under study, and are taken either from the Eurostat (NewCronos database), or from the websites of respective statistical offices. The German data prior to 1991 concern West Germany. With regard to migration between Poland and Germany, it has to be noted that the official statistics of neither country cover the short-term Polish seasonal labour migration undertaken under a bilateral agreement between the two countries. A significant source of migration, only in 2002 amounting to 272 thousand persons (Okólski, 2004: 206), is therefore excluded from the analysis.

As the numbers of migrants reported by the origin and destination countries usually differ, the greater of the two values has been taken as the estimate of the real magnitude of each of the flows, following Kupiszewski (2002: 111–112). In practice it means that, as Germany applies one of the broadest definitions of migratory events in Europe (Nowok *et al.*, 2006: 218), the German data on migration have been taken almost universally. The only exception are flows from Germany to Switzerland in the period 2001–2004, where the Swiss figures on incoming German citizens have been used, being slightly larger than the German ones.

Subsequently, the numbers of registered long-term migrants have been transformed into crude occurrence-exposure emigration rates, the logarithms of which are subject of forecasting. As the size of the population at risk (denominator of the rate), the mid-year population of the sending country has been used. As in the data on Polish population stocks in 2002 there is a discontinuity in trend caused by underestimated international emigration prior to the recent population census, a simple correction has been applied based on the census results. The 'statistical adjustment' of –390,300 persons has been distributed over the period 1988–2002, proportionally to the size of net migration from Poland registered in Germany. As for 1993 the German sources themselves include an administrative adjustment of the Polish migrant stock by –23,000 persons, this quantity has been distributed uniformly over the period 1988–1992, in order to adjust the German population stocks, as well as the magnitude of inflows from Poland.

4.2. Specification of the forecasting models

Let $m_{x\text{-}y}(t)$ denote logarithms of emigration rates from country x to y per 1,000 population of the country of origin, in year t, where $x, y \in$ {CH, DE, IT, PL} for Switzerland, Germany, Italy, and Poland, respectively. The logarithmic transformation has been used, as emigration rates are by definition positive numbers.

For each of the variables $m_{x\text{-}y}(t)$, let the following five simple types of models, belonging to the wide class of the autoregressive integrated moving average (ARIMA) models (Box and Jenkins, 1976), be defined (the indices x and y are skipped for the transparency of presentation):

M_1: $m(t) = c_1 + \varepsilon_1(t)$ / oscillations around a constant

M_2: $m(t) = c_2 + m(t{-}1) + \varepsilon_2(t)$ / random walk with drift

M_3: $m(t) = c_3 + \phi_3\, m(t{-}1) + \varepsilon_3(t);\ \phi_3 \neq 0 \wedge \phi_3 \neq 1$ / autoregressive process AR(1)

M_4: $m(t) = c_4 + \varepsilon_4(t) - \theta_4\, \varepsilon_4(t{-}1);\ \theta_4 \neq 0$ / moving average process MA(1)

M_5: $m(t) = c_5 + \phi_5\, m(t{-}1) + \varepsilon_5(t) - \theta_5\, \varepsilon_5(t{-}1);\ \phi_5 \neq 0,\ \theta_5 \neq 0$ / ARMA(1,1) process

In all models $M_1, \ldots, M_5$, $\varepsilon_i(t)$ denotes the Gaussian white noise, $\varepsilon_i(t) \sim N(0, \sigma_i^2)$.

The whole class **M** covers all autoregressive moving average ARMA(1,1) models, in general terms specified as $m(t) = c + \phi\, m(t-1) + \varepsilon(t) + \theta\, \varepsilon(t-1)$. The mathematical formulation of particular models $M_1, \ldots, M_5$ depends on the values of parameters ϕ and θ, as shown in Table 1.

Table 1. Specification of forecasting models in the ARMA(1,1) class

Parameter ϕ	Parameter θ	
	$\theta = 0$	$\theta \neq 0$
$\phi = 0$	M_1	M_4
$\phi = 1$	M_2	} M_5
$\phi \neq 0 \wedge \phi \neq 1$	M_3	

Source: own elaboration

Such classification satisfies the conditions to apply the Bayesian model selection and averaging techniques, as $M_1, \ldots, M_5$ are mutually exclusive and add up to the whole model space, **M**.

As the analysis is primarily designed to illustrate the methods rather than to obtain the best-possible migration predictions, which should consider a wider class of models and include other explanatory variables, the forecasting models are limited to simple stochastic processes. Nevertheless, some of these models have already been used in forecasting international migration on the basis of time series within the sampling-theory approach. The examples include a study of de Beer (1997), who modelled migration to and from the Netherlands by an AR(1) process (M_3), and net migration by an MA(1) process (M_4). For Norway, Keilman *et al.* (2001) modelled immigration by means of an ARMA(1,1) process (M_5), and emigration using a random walk (M_2). Alho (1998) predicted migration for Finland using an ARIMA(0,1,1) model, combining a random walk with a moving average component.

With respect to the prior distributions of the model parameters, it is assumed that they are the same for the models $M_1, \ldots, M_5$, wherever applicable. Let the constants c_i follow a diffuse (rather non-informative) normal prior distribution $N(0, 100^2)$. Further, I assume that the priors for the autoregression parameters ϕ_i are also normal, yet more informative, following $N(0.5, 1^2)$. This assumption reflects my belief that the autoregressive parts of appropriate models are likely stationary. For θ_i, I also take normal priors $N(0.5, 1^2)$, which indicate some level of uncertainty as to the degree of smoothing in the moving average parts of the respective models. Finally, the precision parameters τ_i of the error terms $\varepsilon_i(t)$, being reciprocal of the variances, $\tau_i = 1/\sigma_i^2$, are assumed to follow a Gamma distribution with scale parameter $\alpha = 0.5$ and shape parameter $\nu =$

= 0.5. Such a distribution has expected value $\nu/\alpha = 1$ and variance $\nu/\alpha^2 = 2$, what reflects my *a priori* belief in a relatively low precision of estimation.

The estimation has been performed using the Markov chain Monte Carlo (MCMC) algorithm of numerical integration, implemented in the WinBUGS 1.4 software (Spiegelhalter *et al.*, 2003). For computational simplicity, no constraints for ϕ_i and θ_i have been set in the models M_3, M_4, and M_5, as the values that should be excluded (0 and/or 1) have probability of occurring equal zero. The WinBUGS code used for the computations is listed in the Annex to the paper[11].

4.3. Estimation of the models

For each of the time series of migration flows under study, $m_{x\text{-}y}(t)$, all five aforementioned models M_1, …, M_5 have been estimated. The posterior distributions of the parameters of each model have been calculated on the basis of 100,000 iterations of the MCMC algorithm, obtained after discarding the preceding iterations from the 'burn-in' phase of the procedure. After visual checks of convergence of the simulations following the suggestions of Spiegelhalter *et al.* (2003), the length of the 'burn-in' phase has been established as 10,000.

Results of the estimation of particular models are presented in Table 2, summarising the posterior distributions of parameters with their median values and the quantiles of rank 0.025 and 0.975, denoting the 95 percent credible intervals. All estimates have been obtained from the MCMC simulations. Note that, as the posterior distributions of parameters are unimodal, the intervals derived on the basis of respective quantiles have the HPD property defined in (2).

As shown in Table 2, the credible intervals for c_i, ϕ_i, and θ_i that do not cover zero, what indicates the 'significance' of the estimates, consider a minority rather than a majority of models. This reflects high uncertainty of the processes under study, some of which do not perfectly fit into the proposed framework based only on simple models from the ARMA(1,1) class. This gives a rationale for the future research to further extend the model space **M**, for example to ARIMA models, both with and without constants, as suggested by Alho and Spencer (2005: 217).

[11] Examples of WinBUGS programmes of univariate AR(p), MA(q) and ARMA(p,q) models, which have been helpful in preparing the current paper, are provided by Congdon (2003: 172–189, Programmes 5.1–5.3). With respect to the moving average component, Congdon (2003: 187) argues that due to the way WinBUGS operates (based on the centred Normal form), there is a need for assuming an additional error term, $u(t) \sim N(0, \sigma_u^2)$, which for the MA(1) model would yield: $m(t) = c + \varepsilon(t) - \theta\, \varepsilon(t-1) + u(t)$. In my view, this is an unnecessary complication of calculations, which can lead to unreasonable results. Therefore, in the current study I apply the standard moving average model, without the $u(t)$ term, which nevertheless uses the centred Normal form (see Annex).

Table 2. Summaries of posterior distributions of forecasting models' parameters: median, 2.5 and 97.5 percent quantiles (estimated by MCMC)

Model M_i	Constant c_i			Autoregression parameter ϕ_i			Moving average parameter θ_i			Precision $\tau_i = 1/\sigma_i^2$		
	2.5%	median	97.5%	2.5%	median	97.5%	2.5%	median	97.5%	2.5%	median	97.5%
Migration from Italy to Germany, $m_{IT\text{-}DE}$												
M_1	**–0.59**	**–0.43**	**–0.27**	-	-	-	-	-	-	4.20	8.71	15.64
M_2	–0.15	–0.04	0.07	-	-	-	-	-	-	7.86	18.07	35.38
M_3	–0.28	–0.02	0.22	**0.46**	**1.03**	**1.58**	-	-	-	7.09	15.81	32.91
M_4	**–0.66**	**–0.43**	**–0.21**	-	-	-	**–1.09**	**–0.68**	**–0.10**	5.65	11.99	22.12
M_5	–0.36	–0.04	0.22	**0.29**	**0.98**	**1.56**	–0.92	–0.44	0.33	7.73	19.52	38.64
Migration from Germany to Italy, $m_{DE\text{-}IT}$												
M_1	**–0.84**	**–0.69**	**–0.55**	-	-	-	-	-	-	5.45	11.25	20.25
M_2	–0.14	–0.04	0.07	-	-	-	-	-	-	8.46	18.84	40.46
M_3	–0.51	–0.19	0.12	**0.31**	**0.77**	**1.21**	-	-	-	8.51	20.71	44.85
M_4	**–0.89**	**–0.69**	**–0.49**	-	-	-	**–1.05**	**–0.63**	**–0.01**	6.67	14.16	26.90
M_5	–0.62	–0.21	0.15	**0.14**	**0.74**	**1.25**	–0.96	–0.32	0.51	8.49	20.70	44.88
Migration from Poland to Germany, $m_{PL\text{-}DE}$												
M_1	**0.76**	**0.94**	**1.13**	-	-	-	-	-	-	3.83	9.75	20.23
M_2	–0.19	0.00	0.19	-	-	-	-	-	-	3.62	9.07	18.71
M_3	–0.38	0.52	1.42	–0.47	0.45	1.38	-	-	-	3.74	9.95	21.58
M_4	**0.74**	**0.95**	**1.20**	-	-	-	–1.07	–0.28	0.68	3.97	10.96	24.41
M_5	–0.36	0.60	1.58	–0.61	0.37	1.36	–1.04	–0.19	0.73	3.78	10.21	22.88
Migration from Germany to Poland, $m_{DE\text{-}PL}$												
M_1	–0.19	0.00	0.19	-	-	-	-	-	-	3.83	9.58	19.85
M_2	–0.19	–0.01	0.16	-	-	-	-	-	-	4.21	11.03	23.70
M_3	–0.18	–0.01	0.16	–0.24	0.58	1.40	-	-	-	4.22	11.67	26.08
M_4	–0.24	0.00	0.25	-	-	-	–1.06	–0.37	0.62	3.99	10.46	23.35
M_5	–0.22	–0.01	0.22	–0.48	0.50	1.36	–0.95	–0.16	0.75	4.09	11.26	25.34
Migration from Switzerland to Germany, $m_{CH\text{-}DE}$												
M_1	**0.01**	**0.12**	**0.23**	-	-	-	-	-	-	8.52	18.08	39.60
M_2	–0.09	0.01	0.12	-	-	-	-	-	-	9.10	20.71	46.79
M_3	–0.11	0.05	0.21	–0.57	0.63	1.79	-	-	-	8.74	19.46	45.20
M_4	–0.03	0.12	0.26	-	-	-	–0.95	–0.29	0.60	8.69	18.70	42.97
M_5	–0.13	0.05	0.25	–0.72	0.60	1.85	–0.93	–0.24	0.64	8.74	19.23	45.26
Migration from Germany to Switzerland, $m_{DE\text{-}CH}$												
M_1	**–2.19**	**–2.04**	**–1.88**	-	-	-	-	-	-	4.68	9.59	17.10
M_2	–0.06	0.03	0.13	-	-	-	-	-	-	8.86	22.18	42.08
M_3	–1.03	0.21	1.33	**0.49**	**1.09**	**1.62**	-	-	-	7.95	18.68	40.26
M_4	**–2.24**	**–2.03**	**–1.81**	-	-	-	**–1.07**	**–0.64**	**–0.06**	5.90	12.46	22.87
M_5	–1.54	0.13	1.44	**0.24**	**1.05**	**1.68**	–0.90	–0.24	0.48	7.66	17.13	38.04

Notes: **Boldface** denotes 95 percent credible intervals for the estimates of c_i, ϕ_i, and θ_i, which do not cover zero.
Source: own elaboration in WinBUGS

With respect to particular parameters, the constants c_i proved to be much less diffuse than assumed *a priori*. The posterior estimates of autoregressive coefficients ϕ_i, appear to be consistent with the priors in four models (for $m_{DE\text{-}IT}$, $m_{PL\text{-}DE}$, $m_{DE\text{-}PL}$, and $m_{CH\text{-}DE}$), indicating the likely stationarity of the AR(1) component. In the remaining two cases ($m_{IT\text{-}DE}$ and $m_{DE\text{-}CH}$) the values of ϕ_i are likely greater than one, providing arguments for non-stationarity. The posterior estimates of the moving average parameters θ_i, all likely being negative, strongly disagree with my prior beliefs in probably positive values of this parameter.

In all cases, the estimated precision parameters $\tau_i = 1/\sigma_i^2$ have been higher, on average more than ten times, as compared to the values assumed *a priori*. This indicates that the prior assumption on a relatively low precision of estimation was too pessimistic, as the data indicated the opposite. Nevertheless, even small samples of 20 observations (or 14 in the case of migration to and from Poland) were able to correct the improper prior beliefs and give relatively much weight to the data in the posterior distributions of the parameters.

4.4. Results of Bayesian model selection: posterior probabilities

Following the discussion in Section 3.1, two alternative types of prior distributions for the models M_1, …, M_5 have been set: the non-informative uniform one, with $p(M_i) = 0.2$, and the 'Occam's razor' one, with $p(M_i)$ proportional to 2^{-l_i}, l_i denoting the number of parameters of the i-th model. Both types of prior distributions over the model space **M**, and the respective posterior probabilities $p(M_i|x)$, estimated using the MCMC algorithm, are presented in Table 3.

A comparison of the posterior odds ratios R_{ij} defined in (5), on the basis of probabilities shown in Table 3, yields that for half of the forecasted flows ($m_{PL\text{-}DE}$, $m_{CH\text{-}DE}$, and $m_{DE\text{-}CH}$) both types of priors lead to the selection of the same model: in two latter cases – a random walk with drift (M_2), in the former one – a moving average MA(1) process (M_4). In two other cases ($m_{DE\text{-}IT}$, and $m_{DE\text{-}PL}$), assuming the Occam's prior instead of the equiprobable one changed the posterior odds-based decision on the model selection from the autoregressive model M_3 to the simpler random walk M_2. For migration from Italy to Germany ($m_{IT\text{-}DE}$), the decision changed analogously from the moving average process (M_4) to the random walk (M_2). The simplicity of the latter is the reason, why the posterior probabilities under the 'Occam's razor' prior are visibly higher than the ones under the non-informative prior, for all flows under study. Especially in such cases, where the posterior probabilities of particular models under a uniform prior are very similar, the 'Occam's razor' prior penalises models with more parameters.

Table 3. Prior and posterior probabilities for models M_i: non-informative and 'Occam's razor'

Model (M_i)	M_1	M_2	M_3	M_4	M_5	Σ
Prior probabilities						
(A) Non-informative prior, $p(M_i) \propto const.$	0.200	0.200	0.200	0.200	0.200	1
(B) 'Occam's razor' prior, $p(M_i) \propto$ 2^(–l_i)	0.308	0.308	0.154	0.154	0.077	1
Migration from Italy to Germany, $m_{IT\text{-}DE}$						
$p(M_i\|x)$, prior (A)	0.000	0.347	0.205	0.007	**0.441**	1
$p(M_i\|x)$, prior (B)	0.000	**0.616**	0.181	0.007	0.196	1
Migration from Germany to Italy, $m_{DE\text{-}IT}$						
$p(M_i\|x)$, prior (A)	0.000	0.249	**0.367**	0.018	0.366	1
$p(M_i\|x)$, prior (B)	0.000	**0.456**	0.356	0.016	0.171	1
Migration from Poland to Germany, $m_{PL\text{-}DE}$						
$p(M_i\|x)$, prior (A)	0.155	0.092	0.198	**0.313**	0.241	1
$p(M_i\|x)$, prior (B)	0.272	0.168	0.175	**0.275**	0.111	1
Migration from Germany to Poland, $m_{DE\text{-}PL}$						
$p(M_i\|x)$, prior (A)	0.079	0.207	**0.291**	0.171	0.252	1
$p(M_i\|x)$, prior (B)	0.135	**0.361**	0.249	0.147	0.108	1
Migration from Switzerland to Germany, $m_{CH\text{-}DE}$						
$p(M_i\|x)$, prior (A)	0.119	**0.283**	0.224	0.166	0.208	1
$p(M_i\|x)$, prior (B)	0.187	**0.431**	0.173	0.128	0.081	1
Migration from Germany to Switzerland, $m_{DE\text{-}CH}$						
$p(M_i\|x)$, prior (A)	0.000	**0.469**	0.311	0.003	0.217	1
$p(M_i\|x)$, prior (B)	0.000	**0.684**	0.232	0.002	0.081	1

Notes: **Boldface** indicates highest posterior probabilities of models for particular migratory flows. Some numbers may not add up to 1 due to rounding.
Source: own elaboration in WinBUGS

Applying formal criteria for R_{ij} can help identify models with the relatively biggest disagreement with data, for example by setting $O_L = 1/20$ and $O_R = 20$ after Raftery *et al.* (1996). For models of migration from Italy to Germany ($m_{IT\text{-}DE}$) and from Germany to Switzerland ($m_{DE\text{-}CH}$), the ratios that do not fall into the $[O_L, O_R]$ window include all R_{ij} involving models M_1 and M_4. Almost the same applies for models of migration from Germany to Italy ($m_{DE\text{-}IT}$), with two slight exceptions: $R_{42} = 0.07$ under the uniform prior and $R_{54} = 10.7$ under the 'Occam's razor' fit within the $[O_L, O_R]$ interval. This indicates the implausibility of using the oscillations model M_1, as well as the moving average model M_4, for the three mentioned migratory flows.

4.5. International migration predictions: formally-selected and averaged models

Forecasts of migration between the countries under study for the period 2005–2010 have been made within the same WinBUGS programme (see Annex). Predictive distributions of respective emigration rates, showing their median values, and the 80 percent predictive intervals[12] for 2006, 2008, and 2010, are summarised in Table 4. The forecasted values are presented for all

[12] In probabilistic population forecasting, 80 percent predictive intervals are commonly applied, either instead, or alongside the 95 percent intervals traditionally used in statistical inference (see Alho, 1998; Keilman *et al.*, 2001; Lutz *et al.*, 2004). The rationale is to avoid unnecessary amplification of the uncertainty assessment, which is already high in demographic predictions. Lutz *et al.* (2004: 37) argue that they "prefer to use 80 percent intervals […] because the forecast distributions are themselves uncertain at the extremities. The 80 percent intervals are far more robust to the technicalities in the forecasting methodology than the 95 percent intervals".

models M_1, …, M_5, as well as for the averaged ones, applying posterior probabilities listed in Table 2. Due to the unimodality of the predictive distributions, the quantile-based predictive intervals also have the HPD property defined in (2).

Table 4. Summaries of predictive distributions of emigration rates forecasted for 2006, 2008 and 2010: median, 10 and 90 percent quantiles (estimated by MCMC)

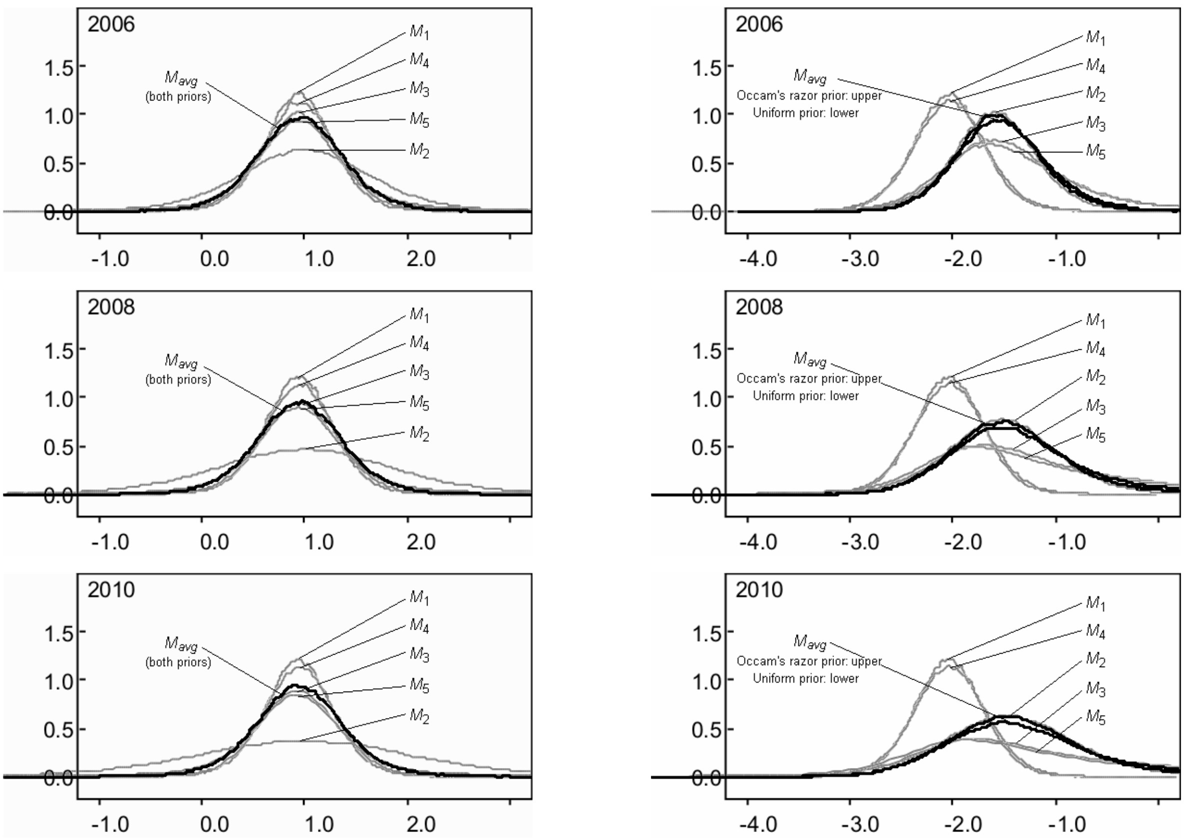
Model	2006			2008			2010		
	10%	median	90%	10%	median	90%	10%	median	90%
Rates of emigration from Italy to Germany, exp(m_{IT-DE})									
M_1*	0.42*	0.65*	1.02*	0.42*	0.65*	1.03*	0.42*	0.65*	1.03*
M_2 (B)	**0.21**	**0.37**	**0.65**	**0.16**	**0.34**	**0.73**	**0.13**	**0.32**	**0.81**
M_3	0.12	0.35	0.69	0.03	0.31	0.77	0.00	0.28	0.83
M_4*	0.40*	0.65*	1.04*	0.40*	0.65*	1.05*	0.40*	0.65*	1.05*
M_5 (A)	**0.14**	**0.37**	**0.72**	**0.05**	**0.35**	**0.85**	**0.01**	**0.33**	**0.93**
M_{avg} (A)	**0.18**	**0.36**	**0.64**	**0.11**	**0.33**	**0.71**	**0.05**	**0.30**	**0.76**
M_{avg} (B)	**0.20**	**0.36**	**0.63**	**0.14**	**0.33**	**0.69**	**0.09**	**0.31**	**0.73**
Rates of emigration from Germany to Italy, exp(m_{DE-IT})									
M_1*	0.33*	0.50*	0.74*	0.34*	0.50*	0.75*	0.34*	0.50*	0.75*
M_2 (B)	**0.21**	**0.37**	**0.64**	**0.16**	**0.34**	**0.72**	**0.13**	**0.32**	**0.78**
M_3 (A)	**0.26**	**0.42**	**0.64**	**0.23**	**0.43**	**0.68**	**0.21**	**0.43**	**0.71**
M_4*	0.33*	0.50*	0.77*	0.33*	0.50*	0.77*	0.33*	0.50*	0.77*
M_5	0.26	0.43	0.70	0.22	0.44	0.75	0.19	0.44	0.78
M_{avg} (A)	**0.26**	**0.42**	**0.64**	**0.23**	**0.41**	**0.68**	**0.21**	**0.41**	**0.70**
M_{avg} (B)	**0.25**	**0.40**	**0.63**	**0.22**	**0.39**	**0.66**	**0.19**	**0.38**	**0.68**
Rates of emigration from Poland to Germany, exp(m_{PL-DE})									
M_1	1.66	2.57	3.98	1.66	2.57	3.97	1.66	2.57	3.99
M_2	1.19	2.73	6.30	0.87	2.73	8.54	0.66	2.72	11.32
M_3	1.50	2.61	4.62	1.40	2.59	4.94	1.34	2.59	5.14
M_4 (A, B)	**1.62**	**2.60**	**4.19**	**1.62**	**2.60**	**4.18**	**1.62**	**2.60**	**4.17**
M_5	1.46	2.62	4.96	1.37	2.61	5.26	1.33	2.60	5.47
M_{avg} (A)	**1.53**	**2.61**	**4.50**	**1.49**	**2.60**	**4.58**	**1.46**	**2.59**	**4.63**
M_{avg} (B)	**1.53**	**2.62**	**4.52**	**1.47**	**2.60**	**4.68**	**1.44**	**2.60**	**4.76**
Rates of emigration from Germany to Poland, exp(m_{DE-PL})									
M_1	0.64	1.00	1.54	0.64	1.00	1.54	0.64	1.00	1.55
M_2 (B)	**0.46**	**0.98**	**2.06**	**0.34**	**0.95**	**2.66**	**0.25**	**0.93**	**3.35**
M_3 (A)	**0.57**	**0.98**	**1.71**	**0.51**	**0.98**	**1.82**	**0.47**	**0.98**	**1.92**
M_4	0.62	1.00	1.62	0.62	1.00	1.62	0.62	1.00	1.62
M_5	0.55	0.98	1.77	0.51	0.98	1.89	0.48	0.97	1.97
M_{avg} (A)	**0.58**	**0.99**	**1.69**	**0.54**	**0.98**	**1.76**	**0.52**	**0.98**	**1.81**
M_{avg} (B)	**0.56**	**0.99**	**1.72**	**0.52**	**0.98**	**1.83**	**0.48**	**0.98**	**1.90**
Rates of emigration from Switzerland to Germany, exp(m_{CH-DE})									
M_1	0.82	1.12	1.53	0.82	1.12	1.53	0.82	1.12	1.53
M_2 (A, B)	**0.72**	**1.21**	**2.05**	**0.62**	**1.25**	**2.52**	**0.54**	**1.29**	**3.04**
M_3	0.76	1.17	1.92	0.70	1.17	2.42	0.65	1.17	3.25
M_4	0.80	1.12	1.58	0.80	1.12	1.57	0.80	1.12	1.58
M_5	0.73	1.16	2.01	0.66	1.17	2.61	0.60	1.17	3.72
M_{avg} (A)	**0.79**	**1.16**	**1.75**	**0.76**	**1.17**	**1.92**	**0.74**	**1.18**	**2.07**
M_{avg} (B)	**0.79**	**1.17**	**1.77**	**0.76**	**1.17**	**1.94**	**0.73**	**1.18**	**2.12**
Rates of emigration from Germany to Switzerland, exp(m_{DE-CH})									
M_1*	0.08*	0.13*	0.20*	0.08*	0.13*	0.20*	0.08*	0.13*	0.20*
M_2 (A, B)	**0.12**	**0.20**	**0.34**	**0.11**	**0.22**	**0.44**	**0.10**	**0.23**	**0.54**
M_3	0.12	0.23	0.56	0.11	0.26	2.07	0.10	0.31	25.08
M_4*	0.08*	0.13*	0.21*	0.08*	0.13*	0.21*	0.08*	0.13*	0.21*
M_5	0.11	0.22	0.59	0.10	0.24	2.35	0.09	0.27	33.95
M_{avg} (A)	**0.13**	**0.22**	**0.40**	**0.12**	**0.24**	**0.66**	**0.12**	**0.27**	**1.49**
M_{avg} (B)	**0.13**	**0.21**	**0.37**	**0.12**	**0.23**	**0.53**	**0.12**	**0.25**	**0.81**

Notes: **Boldface** denotes models selected on the basis of the posterior odds criterion (Table 2), as well as averaged models (M_{avg}). In both cases, (A) and (B) respectively denote the non-informative and 'Occam's razor' model priors. Asterisks (*) indicate models that are hardly probable *a posteriori*, with $p(M_i | x) < 0.05$.
Source: own elaboration in WinBUGS

In Table 4, asterisks indicate forecasts based on models with very low posterior probabilities ($p(M_i|x) < 0.05$). They are presented merely for the sake of comparison with the remaining ones, especially as their results in terms of predictions visibly differ from the outcome of the more probable models. For two selected flows, from Poland to Germany, and from Germany to Switzerland, various predictive distributions of the logarithms of respective emigration rates are also illustrated in Figure 1.

Figure 1. Predictive distributions of $m_{PL\text{-}DE}$ and $m_{DE\text{-}CH}$ for 2006, 2008 and 2010: various models

a) migration Poland-Germany: $m_{PL\text{-}DE}$ *b) migration Germany-Switzerland:* $m_{DE\text{-}CH}$



Note: M_{avg} denotes averaged models (thick black lines).
Source: own elaboration in WinBUGS

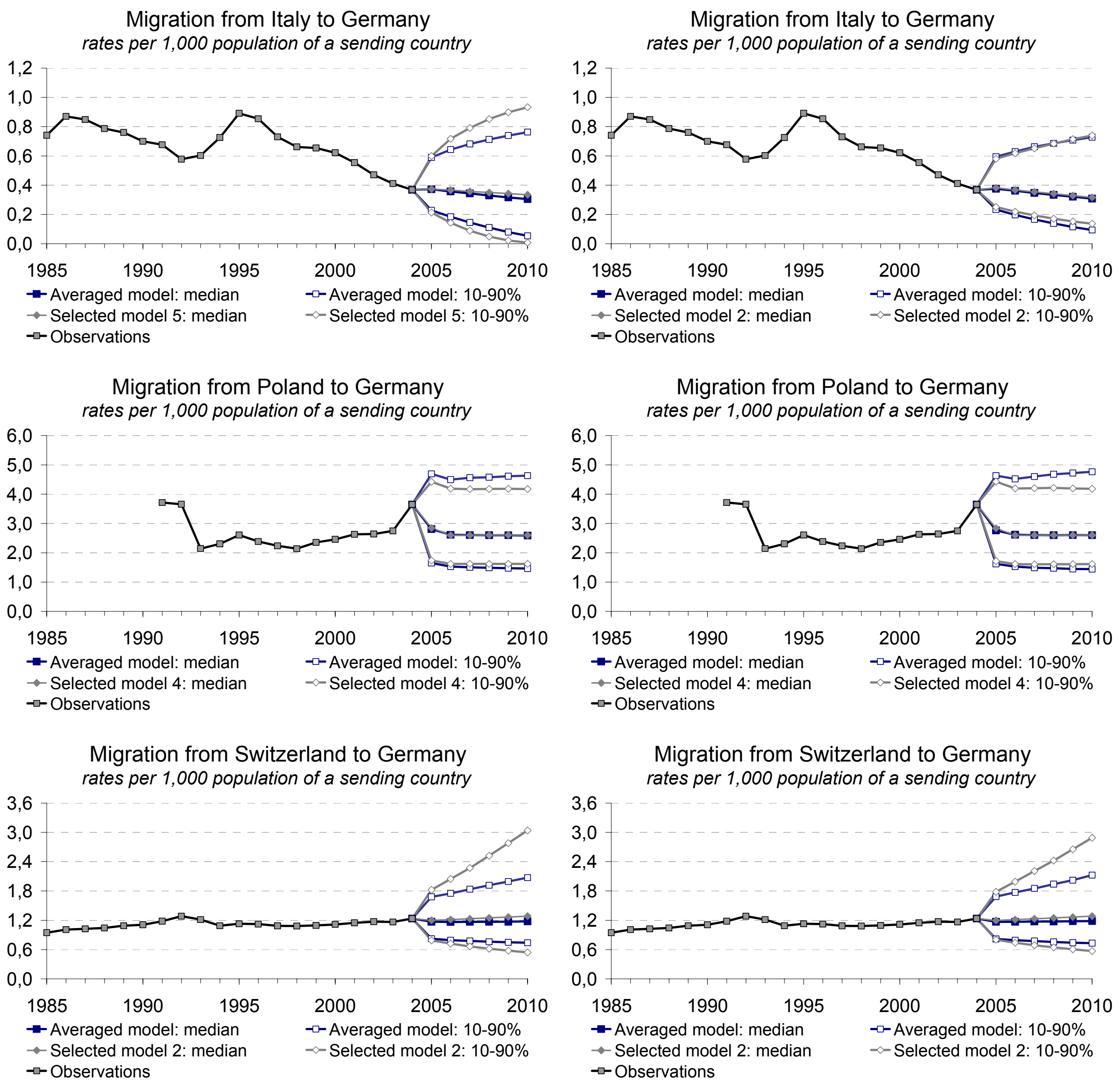
On the example of forecasts of migration from Germany to Switzerland, it is worth noting that models with very low posterior probabilities (M_1 and M_4 in Figure 1) are almost entirely excluded from the calculations of the averaged models. It was even hardly necessary to remove M_1 and M_4 from the modelling framework, as they have effectively been excluded due to their strong disagreement with the data. From the demographic point of view, also the upper limits of the 80 percent credible intervals yielded by models M_3 and M_5 are also implausibly high (cf. Table 4), but these models are included in the averaged forecasts also to a limited extent.

If the models do not differ much from each other with respect to their posterior probabilities, and yield consistent predictions, as for flows from Poland to Germany, then the averaged models take all the single ones into account to a similar degree.

The predicted median values, as well as the limits of the 80 percent predictive intervals for immigration to and emigration from Germany concerning the three countries under study in the period 2005–2010, are illustrated respectively in Figures 2 and 3. The graphs show the historical data series, on the basis of which the estimation was made, and the predictions from the formally-selected and averaged models, both under the uniform and 'Occam's razor' priors.

Figure 2. Forecasted immigration to Germany, 2005–2010: selected and averaged models

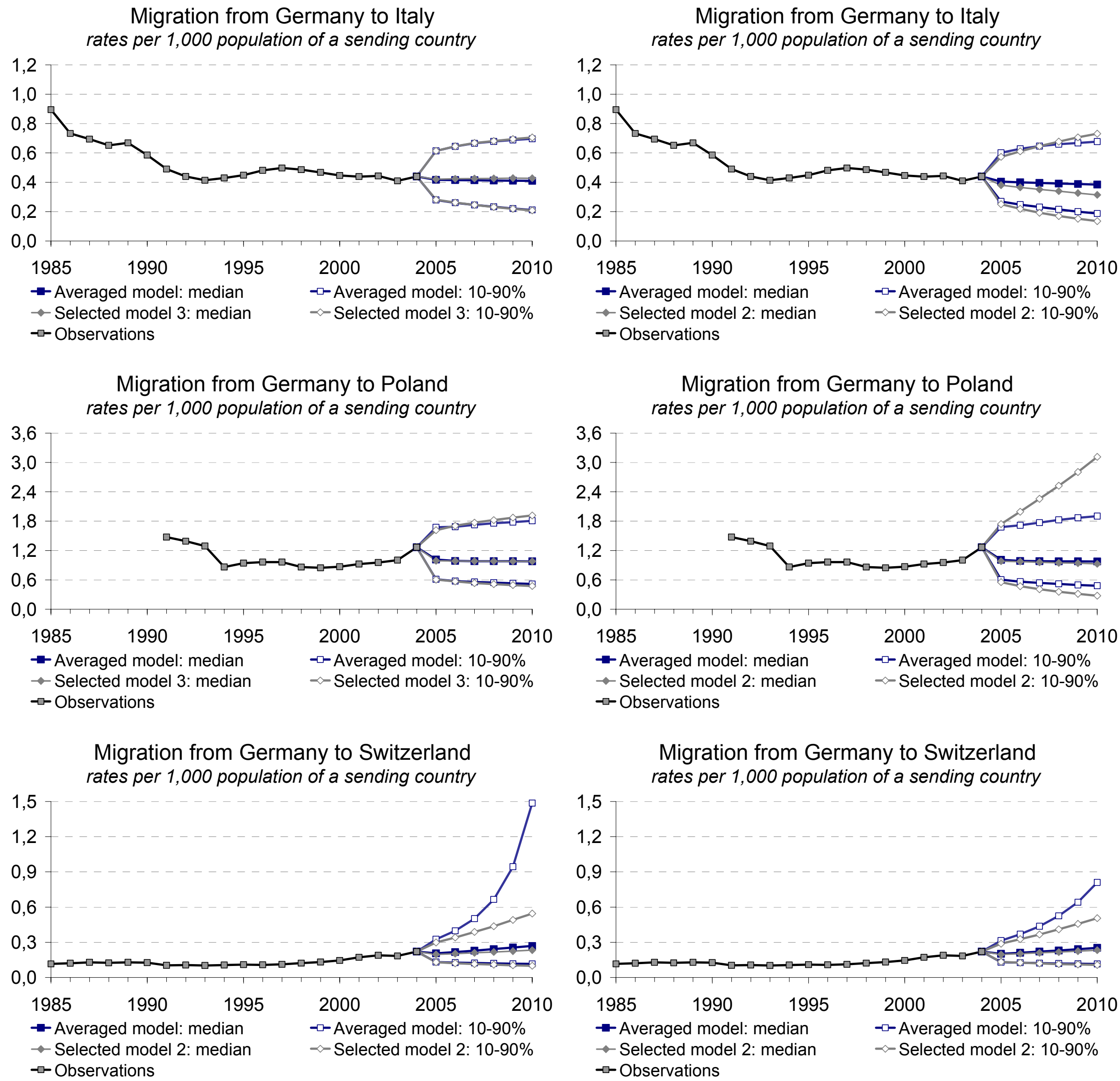
a) non-informative model priors *b) 'Occam's razor' model priors*



Note: Trajectories for alternative priors may slightly differ, as they have been estimated in two separate simulations.
Source: own elaboration in WinBUGS. Data series until 2004: Eurostat/NewCronos; national statistical offices

Figure 3. Forecasted emigration from Germany, 2005–2010: selected and averaged models

a) non-informative model priors *b) 'Occam's razor' model priors*



Note: Trajectories for alternative priors may slightly differ, as they have been estimated in two separate simulations.
Source: own elaboration in WinBUGS. Data series until 2004: Eurostat/NewCronos; national statistical offices

From Figures 2 and 3 it seems that in a majority of cases the median forecast trajectory obtained from the model selected on the basis of the posterior odds criterion is almost the same as the one obtained from the averaged model. This conclusion is valid for both types of model priors. The most significant exception is migration from Germany to Italy under the 'Occam's razor' prior, where the selected random-walk trajectory obtained from M_2 has visibly lower values than the one based on the averaged model.

With respect to the uncertainty of predictions made using particular models, there is no clear rule as to which one of them is higher. In some cases the predictive intervals of averaged models are broader than the ones obtained from the models selected on the basis of posterior odds, which would reflect taking into account the uncertainty of the model selection process. Nevertheless, the opposite can also hold. The latter regards for example situations with relatively similar posterior model probabilities, where the selection criteria (posterior odds) indicate a model with a high variance, as M_2 for migration from Switzerland to Germany under both priors, or from Germany to Poland under the 'Occam's razor' prior.

On the basis of the median trajectories derived from the respective predictive distributions (Figures 2 and 3), the obtained migration forecasts can be summarised as follows. In the period 2005–2010 it is expected that the propensity to migrate from Italy to Germany will visibly diminish. Some decline, although very slight, is also envisaged for migration in the opposite direction. In 2005, the both-ways migration between Germany and Poland is predicted to return to the levels observed before the increase in 2004, which was likely related to the accession of Poland to the European Union. From 2005 onwards the intensity of migration between both these countries is expected to remain almost constant until the end of the forecast horizon. Finally, the relative magnitude of flows between Germany and Switzerland in both directions is envisaged to increase, albeit very slightly, throughout the whole period 2005–2010. From the demographic point of view, the results of median forecasts of the intensities of all migratory movements under study seem therefore reasonable.

A vast majority of averaged models also yielded plausible uncertainty ranges, slightly increasing with time and spanning realistic magnitudes of possible future migration flows. The only exception considers migration from Germany to Switzerland, where the upper limits of the 80 percent predictive intervals are very high and increase very fast, as shown in Figure 3. This is due to the fact that the averaged predictions include the outcome of models M_3 and M_5, which themselves produce hardly plausible results (cf. Table 4). Moreover, in such cases applying the logarithmic transformation of the modelled variables additionally contributes to an overestimation of the upper bounds of predictive intervals due to the way the models are formulated. Hence, in the case of the German-Swiss migration, the uncertainty assessment should be made on the basis of the formally-selected model M_2, rather than of the averaged one, as the results of the former are much closer to what can be reasonably believed for the future.

It is also worth noting that in all cases the predicted uncertainty ranges are relatively wide. This is not surprising, given the inevitable problems underlying high errors of international migration forecasts that have been mentioned in Section 1, at the very beginning of the paper. Although median forecasts are merely the continuation of past trends, there is a high probability that the actual migration developments will substantially deviate from the central trajectories.

Therefore, apart from the median projections, the presented forecasts provide the policy makers with an important piece of information: uncertainty related to migration predictions between the countries under study is high and it should be properly acknowledged in the policy decisions.

5. Conclusion

General advantages of the use of Bayesian methods in forecasting international migration have been discussed in Bijak (2005: 21). Firstly, Bayesian forecasting contains an inherent quantitative analysis of uncertainty, embodied in the predictive distributions of the variables under study. Secondly, Bayesian statistics allows for pooling the features of various forecasting models in a formal way, which has been explored in the current paper using the model selection and averaging techniques. Thirdly, methods including subjective prior information in addition to statistical observations allow for a formal inference in small-sample studies, as in the presented forecasts based on relatively short time series. Finally, the independence of the concept of probability from the frequency of events under study allows for avoiding certain problems with interpretation of results, including the assumption of a repeatable sample. The latter is especially important in social sciences, including migration research, where the samples are usually unique.

The Bayesian model selection procedures presented in the current paper allow for incorporating the prior knowledge on the plausibility of particular models in the ultimate decisions, on which one to use for producing forecasts. One option is the *a priori* equiprobability of models, but the analysis can be designed so as that the prior beliefs prefer the simple models over the more complex ones, according to the 'Occam's razor' principle.

The Bayesian model averaging technique provides an alternative, in which the forecasts from various models are combined together. According to Hoeting *et al.* (1999: 398), its advantages include a better (in theory) predictive performance than any single model, no need to justify the use of any particular model, and a possibility to incorporate several competing models proposed by different researchers working on the topic under study. Moreover, as noted by Hoeting *et al.* (1999: 399), "model averaging is more correct because it takes account of a source of uncertainty that analyses based on model selection ignore", and that "[its] conclusions are more robust than those that depend upon the particular model that has been selected".

A general disadvantage of Bayesian model selection and averaging techniques – the seeming complexity of computations – can be overcome by using numerical tools available in easily-accessible software packages like the WinBUGS.

There are several options of additional studies on the presented methodological aspects of Bayesian forecasting, not discussed in the current paper. Hoeting *et al.* (1999: 399) suggested that the issues that need to be further explored in the context of model selection and averaging include: robustness of the results on different prior information, performance of the methods if the 'true model' is not in the class **M** under study, fine-tuning of the 'Occam's window' approach, as well as development of computational algorithms more efficient than the MCMC.

In the field of applications of the presented methodology to forecasting international migration, the further work should concentrate on extending the model space **M** to cover a wider variety of models. Another option is the analysis of models containing additional socio-economic explanatory variables. Further, proper attention should be paid to the issue of disaggregating the forecasted origin-destination migration rates additionally by sex and age, in order to use them in the cohort-component or multiregional framework of population forecasting. Ideally, such disaggregation should be also made within a probabilistic framework consistent with the methodology used for forecasting the origin-destination emigration rates. In this context, a first attempt has been made by Congdon (2005), who presented Bayesian models for age schedules of migration using the Rogers-Castro double-exponential curve, and the non-parametric random-effects approach.

In my view, the Bayesian model selection and averaging techniques provide valuable tools for reducing uncertainty emerging from an important source – the model specification. The issue is especially important in the case of international migration, a phenomenon that can possibly be explained by many competing hypotheses and theories, leading to a wide variety of predictive models. In that respect, the presented approaches precisely address some of the suggestions made in the discussion on the methodological aspects of demographic forecasting, focusing on the analysis of uncertainty. Quoting one of the leading theorists of the Bayesian approach, D. V. Lindley (2000: 294), "statistics is essentially a study of uncertainty", and the demographic applications are here by no means an exception.

References

Ahlburg, D. A. (1995). Simple versus complex models: Evaluation, accuracy and combining. *Mathematical Population Studies*, 5 (3), 281–290.

Ahlburg, D. A., & Lutz, W. (1998). The Need to Rethink Approaches to Population Forecasts. *Population and Development Review*, 24 (Suppl.), 1–14.

Ahlburg, D. A., Lutz, W., & Vaupel, J. W. (1998). Ways to Improve Population Forecasting: What Should Be Done Differently in the Future? *Population and Development Review*, 24 (Suppl.), 191–198.

Alho, J. (1990). Stochastic methods in population forecasting. *International Journal of Forecasting*, 6 (4), 521–530.

Alho, J. (1998). *A stochastic forecast of the population of Finland*. Reviews 1998/4. Helsinki: Statistics Finland.

Alho, J., & Spencer, B. (1985). Uncertain population forecasting. *Journal of the American Statistical Association*, 80 (390), 306–314.

Alho, J., & Spencer, B. (2005). *Statistical demography and forecasting*. New York: Springer.

Armstrong, J. S. (1985). *Long-Range Forecasting*. New York: John Wiley.

Barnard, G. A. (1947). The Meaning of a Significance Level. *Biometrika*, 34 (2), 179–182.

Barnard, G. A. (1949). Statistical inference. *Journal of the Royal Statistical Society B*, 11 (1), 115–149.

Bayarri, M. J., & Berger, J. O. (2004). The Interplay of Bayesian and Frequentist Analysis. *Statistical Science*, 19 (1), 58–80.

Bayes, T. (1763). An Essay towards solving a Problem in the Doctrine of Chances. *Philosophical Transactions of the Royal Society*, 53, 370–418.

Bernardo, J. M., & Smith, A. F. M. (2000). *Bayesian Theory*. Chichester: John Wiley.

Bijak, J. (2005). *Bayesian methods in international migration forecasting*. CEFMR Working Paper 6/2005. Warsaw: Central European Forum for Migration Research.

Box, G. E. P., & Jenkins, G. M. (1976). *Time Series Analysis: Forecasting and Control*. San Francisco: Holden-Day.

Brücker, H., & Siliverstovs, B. (2005). *On the Estimation and Forecasting of International Migration: How Relevant is Heterogeneity Across Countries*. IZA Discussion Paper 170. Bonn: Institut zur Zukunft der Arbeit.

Chatfield, C. (2002). Confessions of a pragmatic statistician. *The Statistician*, 51 (1), 1–20.

Congdon, P. (2000). A Bayesian Approach to Prediction using the Gravity Model, with an Application to Patient Flow Modelling. *Geographical Analysis*, 32 (3), 205–224.

Congdon, P. (2003). *Applied Bayesian Modelling*. Chichester: John Wiley.

Congdon, P. (2005). *Bayesian modelling of migration age schedules between Scotland and England: A comparison of the Rogers-Castro parameter-based approach and a completely nonparametric random effects approach*. Paper for the "Workshop on the Estimation of International Migration in Europe: Issues, Models and Assessment", Southampton, UK, 28–30 September 2005.

Daponte, B. O., Kadane, J. B., & Wolfson, L. J. (1997). Bayesian Demography: Projecting the Iraqi Kurdish Population, 1977–1990. *Journal of the American Statistical Association*, 92 (440), 1256–1267.

De Beer, J. (1990). *Uncertainty of international-migration projections for the 12 EC-countries*. Voorburg: Central Bureau of Statistics.

De Beer, J. (1997). The effect of uncertainty of migration on national population forecasts: the case of the Netherlands. *Journal of Official Statistics*, 13 (3), 227–243.

De Finetti, B. (1937). La prévision: ses lois logiques, ses sources subjectives. *Annales del'Institute Henri Poincaré*, 7, 1–68.

Eurostat (2005). *Population projections for the European Union 2004–2050*. Luxembourg: Eurostat.

Girosi, F., & King, G. (2005). *Demographic Forecasting*. Manuscript. Boston, MA: Center for Basic Research in the Social Sciences, Harvard University; «http://gking.harvard.edu/files/smooth.pdf».

Gjaltema, T. A. (2001). *Judgement in Population Forecasting*. Paper for the European Population Conference, Helsinki, 7–9 June 2001.

Gorbey, S., James, D., & Poot, J. (1999). Population Forecasting with Endogenous Migration: An Application to Trans-Tasman Migration. *International Regional Science Review*, 22 (1), 69–101.

Hoeting, J. A., Madigan, D., Raftery, A. E., & Volinsky, C. T. (1999). Bayesian Model Averaging: A Tutorial. *Statistical Science*, 14 (4), 382–417.

Jaynes, E. T. (1976). Confidence intervals vs. Bayesian intervals. In: W. L. Harper, & C. A. Hooker (eds.), *Foundations of Probability Theory, Statistical Inference and Statistical Theories of Science 2*. Dordrecht: Reidel, 175–257 (with discussion).

Jeffreys, H. (1939). *Theory of Probability*. Oxford: Oxford University Press.

Jeffreys, W. H., & Berger, J. O. (1992). Ockham's razor and Bayesian analysis. *American Scientist*, 80 (1), 64–72.

Keilman, N. (1990). *Uncertainty in national population forecasting: Issues, backgrounds, analyses, recommendations*. Amsterdam: Swets & Zeitlinger.

Keilman, N. (2001). Demography: Uncertain population forecasts. *Nature*, 412 (6846), 490–491.

Keilman, N., Pham, D. Q., & Hetland, A. (2001). *Norway's Uncertain Demographic Future*. Oslo: Statistics Norway.

Kupiszewski, M. (2002). *Modelowanie dynamiki przemian ludności w warunkach wzrostu znaczenia migracji międzynarodowych* [*Modelling population dynamics considering the increasing role of international migration*]. Warsaw: Institute of Geography and Spatial Organization, Polish Academy of Sciences.

Lanzieri, G. (2004). *EUROPOP 2003: Methodology for Drafting International Migration Assumptions in the EU-15 Member States*. Paper for the meeting of the Working Group on Population Projections, 15–16 July 2004. Luxembourg: Eurostat.

Laplace, P. S. (1812). *Théorie Analytique des Probabilités*. Paris: Veuve Courcier.

Lee, R. D., & Tuljapurkar, S. (1994). Stochastic population projections for the United States: Beyond high, medium and low. *Journal of the American Statistical Association*, 89 (419), 1175–1189.

Lindley, D. V. (2000). The philosophy of statistics. *The Statistician*, 49 (3): 293–337.

Lutz, W., & Goldstein, J. R. (2004). Introduction: How to Deal with Uncertainty in Population Forecasting? *International Statistical Review*, 72 (1), 1–4.

Lutz, W., Sanderson, W. C., & Scherbov, S. (1996). Probabilistic population projections based on expert opinion. In: W. Lutz (ed.), *The future population of the World. What can we assume today?* London: Earthscan, 397–428.

Lutz, W., Sanderson, W. C., & Scherbov, S. (2004). The End of World Population Growth. In: W. Lutz, W. C. Sanderson, & S. Scherbov (eds.), *The End of World Population Growth in the 21st Century: New Challenges for Human Capital Formation and Sustainable Development*. London: Earthscan, 17–83.

Madigan, D., & Raftery, A. E. (1994). Model selection and accounting for model uncertainty in graphical models using Occam's window. *Journal of the American Statistical Association*, 89 (428), 1535–1546.

National Research Council (2000). *Beyond Six Billion. Forecasting the World's Population*. Washington, DC: National Academies Press.

Nowok, B., Kupiszewska, D., & Poulain, M. (2006). Statistics on International Migration Flows. In: M. Poulain, N. Perrin, & A. Singleton (eds.), *THESIM: Towards Harmonised European Statistics on International Migration*. Louvain-la-Neuve: Presses universitaires de Louvain, 203–231.

Okólski, M. (2004). Seasonal labour migration in the light of the German-Polish bilateral agreement. In: *Migration for Employment: Bilateral Agreements at a Crossroads*. Paris: Organisation for Economic Co-operation and Development, 203–214.

Osiewalski, J. (2001). *Ekonometria bayesowska w zastosowaniach* [*Bayesian econometrics in applications*]. Cracow: Cracow University of Economics.

Pittenger, D. (1978). The role of judgement, assumptions, techniques and confidence limits in forecasting population. *Socioeconomic Planning Sciences*, 12 (5), 271–276.

Raftery, A. E., Madigan, D., & Volinsky, C. T. (1996). Accounting for model uncertainty in survival analysis improves predictive performance. In: J. Bernardo, J. Berger, A. Dawid, and A. Smith (eds.), *Bayesian Statistics 5*. Oxford: Oxford University Press, 323–349 (with discussion).

Ramsey, F. P. (1926). Truth and Probability. In: R. B. Braithwaite (ed.) (1931), *F. P. Ramsey: The Foundations of Mathematics and other Logical Essays*. London: Kegan, Paul, Trench, Trubner & Co.; New York: Harcourt, Brace and Co., 156–196.

Rees, P. H., Kupiszewski, M., Eyre, H., Wilson, T., & Durham, H. (1999). *The evaluation of regional population projections for the European Union*. Eurostat Working Paper 3/2001/E/nº9. Luxembourg: Eurostat.

Rees, P. H., & Turton, I. (1998). *Investigation of the effects of input uncertainty on population forecasting*. Paper for the 3rd International GeoComputation Conference, Bristol, 17–19 September 1998.

Savage, L. J. (1954). *Foundations of Statistics*. New York: John Wiley.

Silvey, S. D. (1978). *Statistical inference* [Polish edition: *Wnioskowanie statystyczne*]. Warsaw: PWN.

Spiegelhalter, D. J., Thomas, A., Best, N. G., & Lunn, D. (2003). *WinBUGS Version 1.4 Users Manual*. Cambridge: Medical Research Council Biostatistics Unit; «http://www.mrc-bsu.cam.ac.uk/bugs».

Tuljapurkar, S., & Boe, C. (1999). Validation, probability-weighted priors and information in stochastic forecasts. *International Journal of Forecasting*, 15 (3), 259–271.

United Nations (2005). *World Population Prospects: 2004 Revision*. New York: United Nations.

van der Gaag, N., & van Wissen, L. (1999). *Analysis and Forecasting of International Migration by Major Groups (Part II)*. Eurostat Working Paper 3/1999/E/nº9. Luxembourg: Eurostat.

Withers, S. D. (2002). Quantitative methods: Bayesian inference, Bayesian thinking. *Progress in Human Geography*, 26 (4), 553–566.

Zellner, A. (1971). *An Introduction to Bayesian Inference in Econometrics*. New York: John Wiley.

Annex: WinBUGS code for the Bayesian model selection and averaging problems

```
# Specification of the models

# M[1]: m(t) = c[1] + e(t)
# M[2]: m(t) = c[2] + m(t-1) + e(t)
# M[3]: m(t) = c[3] + phi[3] m(t-1) + e(t); phi[3]<>0 & phi[3]<>1
# M[4]: m(t) = c[4] + u(t) - theta[4] e(t-1) + e(t); theta[4]<>0
# M[5]: m(t) = c[5] + phi[5] m(t-1) + e(t) - theta[5] u(t-1) + e(t); phi[5]<>0 & theta[5]<>0
# Mav: m(t) = m(mod) [averaged model]

model {

# Priors for variables
for (k in 1:5) { c[k] ~ dnorm(0,0.0001) } # constants
phi[3] ~ dnorm (0.5,1); phi[5] ~ dnorm (0.5,1) # autoregression coefficients
theta[4] ~ dnorm (0.5,1); theta[5] ~ dnorm (0.5,1) # moving average coefficients
for (k in 1:5) { tau[k] ~ dgamma(0.5,0.5) } # precision for the random terms e(t)
e[1,4] ~ dnorm(0,tau[4]); e[1,5] ~ dnorm(0,tau[5]) # artificial MA error terms for t=1

# Priors for models
mod ~ dcat(p[]) # categorical prior over the model space
for (t in 1:5) { p[t] <- 1/5 } # uniform distribution
# p[1] <- 0.30769; p[2] <- 0.30769; p[3] <- 0.15385; p[4] <- 0.15385; p[5] <- 0.07692
# Occam's razor prior, to be used alternatively, instead of the uniform one

# Definitions of data
for (t in 1:n) { for (k in 1:5) { m[t,k] <- log(MR[t]) }
               mav[t] <- log(MR[t]) }

# Estimation of the models' parameters
for (t in 2:n) { mu[t,1] <- c[1]
               mu[t,2] <- c[2] + m[t-1,2]
               mu[t,3] <- c[3] + phi[3] * m[t-1,3]
               mu[t,4] <- c[4] - theta[4] * e[t-1,4]
               mu[t,5] <- c[5] + phi[5] * m[t-1,5] - theta[5] * e[t-1,5]
               for (k in 1:5) { m[t,k] ~ dnorm(mu[t,k], tau[k]);
                              m.new[t,k] ~ dnorm(mu[t,k], tau[k]) }
               e[t,4] <- m[t,4]-mu[t,4]
               e[t,5] <- m[t,5]-mu[t,5]
               muav[t] <- mu[t,mod]; muav.new[t] <- m.new[t,mod] # averaged model
               mav[t] ~ dnorm(muav[t], tau[mod]); mav.new[t] ~ dnorm(muav.new[t], tau[mod]) }

# Forecasts for t = n+1 … N
for (t in n+1:N) { mu.new[t,1] <- c[1]
               mu.new[t,2] <- c[2] + m.new[t-1,2]
               mu.new[t,3] <- c[3] + phi[3] * m.new[t-1,3]
               mu.new[t,4] <- c[4] - theta[4] * e[t-1,4]
               mu.new[t,5] <- c[5] + phi[5] * m.new[t-1,5] - theta[5] * e[t-1,5]
               for (k in 1:5) { m.new[t,k] ~ dnorm(mu.new[t,k], tau[k]) }
               e[t,4] <- m.new[t,4]-mu.new[t,4]
               e[t,5] <- m.new[t,5]-mu.new[t,5]
               muav.new[t] <- m.new[t,mod] # averaged forecast
               mav.new[t] ~ dnorm(muav.new[t], tau[mod]) }
}

# Data sets

# Data CH-DE
list( n = 20, N = 26, MR = c(0.9442, 1.0071, 1.0237, 1.0424, 1.0874, 1.1061, 1.1804, 1.2833,
1.2131, 1.0884, 1.1282, 1.1225, 1.0856, 1.0812, 1.0932, 1.1149, 1.1458, 1.1714, 1.1646,
1.2346) )

# Data DE-CH
list( n = 20, N = 26, MR = c(0.1156, 0.1213, 0.1290, 0.1246, 0.1295, 0.1266, 0.1036, 0.1060,
0.1024, 0.1067, 0.1098, 0.1081, 0.1119, 0.1220, 0.1314, 0.1449, 0.1719, 0.1889, 0.1834,
0.2208) )

# Data IT-DE
list( n = 20, N = 26, MR = c(0.7412, 0.8701, 0.8481, 0.7867, 0.7598, 0.6996, 0.6761, 0.5775,
0.6024, 0.7256, 0.8909, 0.8531, 0.7304, 0.6617, 0.6537, 0.6213, 0.5542, 0.4703, 0.4115,
0.3682) )
```

```
# Data DE-IT
list( n = 20, N = 26, MR = c(0.8949, 0.7328, 0.6938, 0.6513, 0.6687, 0.5855, 0.4900, 0.4392,
0.4131, 0.4294, 0.4481, 0.4810, 0.4968, 0.4859, 0.4673, 0.4465, 0.4384, 0.4429, 0.4096,
0.4396) )

# Data PL-DE
list( n = 14, N = 20, MR = c(3.7151, 3.6538, 2.1421, 2.3050, 2.6049, 2.3845, 2.2346, 2.1410,
2.3537, 2.4580, 2.6273, 2.6409, 2.7464, 3.6478) )

# Data DE-PL
list( n = 14, N = 20, MR = c(1.4752, 1.3900, 1.2912, 0.8635, 0.9428, 0.9631, 0.9638, 0.8608,
0.8466, 0.8686, 0.9231, 0.9545, 1.0046, 1.2669) )

# Initial values

list( mod = 1, c = c(0, 0, 0, 0, 0), phi = c(NA, NA, 0.5, NA, 0.5), tau = c(1, 1, 1, 1, 1),
theta = c(NA, NA, NA, 0.5, 0.5) )
```

■