Central European Forum For Migration Research

Środkowoeuropejskie Forum Badań Migracyjnych

CEFMR Working Paper 6/2005

BAYESIAN METHODS IN INTERNATIONAL MIGRATION FORECASTING

Jakub Bijak

ul. Twarda 51/55, 00-818 Warsaw, Poland
tel. +48 22 697 88 34, fax +48 22 697 88 43
e-mail: cefmr@cefmr.pan.pl
Internet: www.cefmr.pan.pl

Central European Forum for Migration Research (CEFMR) is a research partnership of the Foundation for Population, Migration and Environment, Institute of Geography and Spatial Organization of the Polish Academy of Sciences and the International Organization for Migration

International Organization For Migration

Foundation for Population, Migration and Environment

Institute of Geography and Spatial Organisation, Polish Academy of Sciences

CEFMR Working Paper
6/2005

BAYESIAN METHODS IN INTERNATIONAL MIGRATION FORECASTING

Jakub Bijak[*]

[*] Central European Forum for Migration Research in Warsaw

Abstract: The paper addresses selected methodological aspects of international migration forecasting. The objective is to present an overview of the existing methods and to propose an alternative based on the Bayesian statistics, combining the formality of inference with the subjective expert opinion. Uncertainty and judgement in migration forecasting are discussed, followed by an introduction to Bayesian statistics, and an overview of the existing forecasting methods. As an illustration, long-term migration between Germany and Poland is predicted for 2004–2010 using three Bayesian models. It is argued that due to the explicit incorporation of subjective elements in order to address the uncertainty issue, and due to the properties of the estimates and predictions, the Bayesian methods are a valuable tool for migration forecasting.

Keywords: migration forecasting, Bayesian methods

Paper prepared for the *Workshop on the Estimation of International Migration in Europe: Issues, Models, and Assessment*, Southampton, United Kingdom, 28–30 September 2005

Acknowledgements:
The paper has been 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, grant 03-34. The author also benefited from the Stipend for Young Scientists of the Foundation for Polish Science. Many thanks go to Prof. Jacek Osiewalski for very helpful comments and discussion.

Editor
ul. Twarda 51/55, 00-818 Warsaw, Poland
tel. +48 22 697 88 34, fax +48 22 697 88 43
e-mail: cefmr@cefmr.pan.pl
Internet: www.cefmr.pan.pl

ISSN 1732-0631
ISBN 83-921915-5-2

Contents

1. Introduction **3**

2. Uncertainty and subjectivity in migration forecasting and in the Bayesian statistics **4**

2.1. Uncertainty and subjectivity in migration forecasting 4

2.2. Bayesian statistics: introductory notes 5

3. Overview of existing methods of forecasting international migration **7**

3.1. Survey-based and Delphi-style migration scenarios 7

3.2. Mathematical models of population flows 8

3.3. Econometric forecasts of international migration 9

3.4. Stochastic forecasts of migration time series 11

3.5. Existing Bayesian models and forecasts of population flows 12

4. Examples of simple Bayesian models for forecasting international migration **13**

4.1. Data, specification and estimation of the models 13

4.2. Forecasts of international migration between Poland and Germany, 2004–2010 15

5. Conclusions **21**

References **22**

Appendix: WinBUGS code for Bayesian AR(1) and VAR(1) forecasting models **27**

1. Introduction

Forecasting international migration is an important, yet difficult research task, characterised by the highest errors among the forecasts of all components of the demographic change (NRC, 2000). Reasons for this include a lack of a comprehensive migration theory (Willekens, 1995), difficulties in operationalisation of the theoretical framework of migration, uncertainty of potential explanatory variables, ignoring forced migration and policy elements in the forecasts, as well as poor data quality (Kupiszewski, 2002).

In order to improve accuracy of the international migration forecasts, attempts should be also made to improve the forecasting methodology. The paper is devoted to selected methodological aspects of international migration forecasting. The objective is to present an overview of the existing forecasting methods and to propose an alternative approach based on the paradigm of Bayesian statistics. The underlying idea is to suggest a methodology that would combine the advantages of the existing ones, including both the formality of the applied statistical tools, and including subjective expert judgement in the forecasting model.

Apart from the introduction, the paper comprises of four sections. In Section 2, the issues of uncertainty, subjectivity and expert judgement in migration forecasting are addressed, followed by brief introductory remarks about the Bayesian statistical inference. The Bayesian philosophy is presented as a formal way to transform the prior beliefs by including sample observations, in order to obtain the posterior knowledge – the outcome of the analysis. This methodology allows for construction of forecasting models combining the formal methods with the subjective expertise.

In Section 3, an overview of the existing methods and practice in international migration forecasting is offered, focusing on (1) survey-based and Delphi-style migration scenarios, (2) mathematical models of population flows, (3) econometric forecasts of international migration, (4) stochastic forecasts of time series, and (5) existing Bayesian forecasts. The models are briefly evaluated according to their methodological features, allowing for identification of their major strengths and weaknesses.

In Section 4, examples of construction and estimation of Bayesian forecasting models are presented for international migration between Poland and Germany, followed by the mid-range forecasts for seven years 2004–2010. Finally, Section 5 contains the main conclusions of the study, and an evaluation of potential usefulness of the Bayesian methodology in forecasting international migration.

2. Uncertainty and subjectivity in migration forecasting and in the Bayesian statistics

2.1. Uncertainty and subjectivity in migration forecasting

Uncertainty about the future values of the forecasted phenomena is an immanent feature of every forecast. 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, and to the improper model specification. Rees and Turton (1998), and Keilman (2001) 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, the United Nations (2005) and the Eurostat (2006, forthcoming);
3. Accommodated within a stochastic approach, which quantifies uncertainty in terms of probabilities of future events. Keilman (2001), and Wilson and Bell (2004) distinguish three types of stochastic forecasts: extrapolation of time series (de Beer, 1990; Lee and Tuljapurkar, 1994; Keilman et al., 2001), propagation of historical forecast errors (Keyfitz, 1981; Stoto, 1983; Alho and Spencer, 1985; Alho, 1990; NRC, 2000), and probabilistic projections based on the expert judgement (Lutz et al., 1996, 1998, 2004).

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

From the probabilistic point of view the deterministic and the scenario 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 scenarios (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 regions under study. The aggregate effects are thus based on the assumption of a perfect correlation between the variables or regions, which is not formally examined, and very often not true (NRC, 2000: 191–192). Unlike in the former two cases, in the stochastic approach uncertainty is quantified in terms of probability. Thanks to this, 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 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 forecasting model selection and assumption-making on the 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), Alho (1990), and Lutz et al. (1996, 1998, 2004).

International migration is a very complex and multi-dimensional 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, political science 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 promising framework of forecasting international migration.

2.2. Bayesian statistics: introductory notes

In the Bayesian paradigm in statistical inference, based on the Bayes Theorem (Bayes, 1763; Laplace, 1812), the sample information is used to transform the *prior knowledge* of the researcher with respect to the phenomenon under study, into the *posterior knowledge*. 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. The Bayesian statistics infers on the unknown parameters of the model describing the phenomenon (θ), conditionally on the statistical information (x), unlike in the traditional *sampling-theory* statistical methods, e.g. in the Neyman-Pearson theory of hypothesis testing. 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). The scheme of Bayesian inference and the Bayes Theorem are presented in Box 1.

Box 1. The Bayes Theorem and the Bayesian statistical inference

posterior knowledge	=	prior knowledge	·	likelihood of the data
$p(\theta \mid x)$	=	$p(\theta)$	·	$p(x \mid \theta) / p(x)$

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 the 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 regions*, analogous to confidence regions in the sampling-theory statistics, but without the problems and inconsistencies regarding the interpretation of the latter (Jaynes, 1976).

A key concept in the Bayesian statistics, distinguishing it from the sampling-theory approach, is subjective probability, independent from the frequency of events under study (Ramsey, 1926; De Finetti, 1937). Statistical inference can be seen as a decision problem, with strong relations between the concepts of probability and utility. 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. However, 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).

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, θ. The predictive probability distribution of $\mathbf{x}^{\mathbf{P}}$ can be calculated according to the formula presented in Box 2 and interpreted as an average from the conditional predictive distribution $p(\mathbf{x}^{\mathbf{P}}|\theta,\mathbf{x})$, weighted with the posterior probabilities of the parameters (Zellner, 1971: 29). Natural results of Bayesian forecasting are predictive credible regions, which formally reflect the uncertainty of the phenomenon under study.

Box 2. Forecasting in the Bayesian approach

predictive distribution

$$p(\mathbf{x}^{\mathbf{P}}|\mathbf{x}) = \int_{\Theta} p(\mathbf{x}^{\mathbf{P}},\theta|\mathbf{x})d\theta = \int_{\Theta} p(\mathbf{x}^{\mathbf{P}}|\theta,\mathbf{x}) \cdot p(\theta|\mathbf{x})\, d\theta.$$

Bayesian methodology can reduce the estimation and prediction errors, in case 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.

Additionally, the Bayesian methodology allows for a formal model selection in order to maximally utilise information from the sample, by comparing the posterior odds of different models given the data. Some authors suggest that such criteria favour more straightforward explanations of the phenomena under study, according to the principle of the *Ockham's razor* (Jeffreys and Berger, 1992). Another possibility is *Bayesian inference pooling*, currently known as *Bayesian model averaging* (Hoeting et al., 1999), which allows for combining the features of various predictive models in order to reduce the uncertainty of model specification.

In demography, the Bayesian approach has been successfully applied to forecast population (Daponte et al., 1997), or the main components of population change: fertility (Tuljapurkar and Boe, 1999), mortality (Girosi and King, 2004), and migration (Gorbey et al., 1999; Brücker and Siliverstovs, 2005).

3. Overview of existing methods of forecasting international migration

3.1. Survey-based and Delphi-style migration scenarios

One group of research studies that are used to assess the future international migration flows are sociological survey studies of 'migration potential' (Fassmann and Hintermann, 1997; IOM, 1998). In such studies, the respondents are usually asked, whether they consider undertaking migration, and, if yes, when, for how long, etc. There are a few problems with such studies (Kupiszewski, 2002). Firstly, the definitions used in such studied are very vague, as in the case of the 'migration potential' itself. Secondly, the way the questions are formulated may heavily influence the results obtained. Thirdly, the declarations people make to the survey-takers are not directly transformable into the actual migration behaviour. Finally, in many studies of this type, the sample sizes are to small to obtain meaningful results considering breakdowns by sex, age groups, regions, etc.

Another group of methods for forecasting migration explicitly refers to the expert judgement. They include obtaining migration scenarios from Delphi or quasi-Delphi analyses (Drbohlav, 1995), as well as from surveys conducted among the experts in the field (Bauer and Zimmermann, 1999). The outcomes of such studies can serve as a valuable input into formal statistical forecasts, especially as informative prior distributions in the Bayesian models.

3.2. Mathematical models of population flows*

According to Kupiszewski (2002), mathematical models of migration emerge predominantly from two different approaches: geographic and demographic. The former one focuses more on the spatial outcomes of the redistribution of migrants, while the latter – on the population distributions by sex and age, and on their impact on the overall demographic dynamics. Both of them apply mathematical tools to model and forecast migratory flows.

The geographic approach focuses mainly on the applications of the methodology of Markov chains, as well as various models of spatial interactions. The usage of Markov chains in modelling population movements between regions (states of the chain) evolved from the models with homogeneous transition matrices (Prais, 1955; Brown, 1970). The important modifications included models with unobserved heterogeneity of the populations, in the *mover-stayer* model of Blumen et al. (1955), or in a model with different transition matrices for various subpopulations (Goodman, 1961). Other methodological variations are models with heterogeneous transition matrices (Rogers, 1966; Joseph, 1974), and non-stationary chains. The latter include semi-Markov processes (Ginsberg, 1971), as well as models with the 'cumulative inertia' property, assuming that the longer a person does not migrate, the less propensity (s)he will have to do it in the future (McGinnis et al., 1963).

The examples of models of spatial interactions include the 'intervening opportunities' concept of Stouffer (1940), where the number of migrations is proportional to the number of open 'opportunities' for the migrants at the destination, and inversely proportional to the number of similar 'opportunities' located closer to the region of origin. Stewart (1941) and Isard (1960) developed gravity models of migration, in analogue to Newton's law, with distance as a discounting factor. Other examples of mathematical models of spatial interactions applied to migration are: entropy, catastrophe theory, and bifurcations (Wilson, 1970, 1981).

In the demographic approach, the traditionally-used cohort-component model for demographic projections, originating from Leslie (1945), has evolved to include migration in the population accounting models (Rees and Wilson, 1973), multiregional models (Rogers, 1975), and multi-state models (Keyfitz, 1980). The migration component has been also incorporated in the micro-level event-history analysis, with migration as one of the possible

* Section partially based on the overview of migration models made by Kupiszewski (2002).

demographic events that may happen to an individual (Ginsberg, 1979; Courgeau, 1985). In the latter case, the analysis is performed using the Monte Carlo micro-simulations.

In addition to the geographic and demographic approaches, there has been also an attempt to model social processes, including migration, using the tools of theoretical physics within the framework of the 'sociodynamic' approach (Weidlich and Haag, 1988). However, the complexity of the model rendered it unexploited in practice.

The main drawback of a majority the presented mathematical models of migration, apart from the event-history analysis, is that they themselves do not explicitly address the issue of uncertainty, important for preparing any forecast on their basis. Although some of the models apply stochastic tools (e.g., Markov chains), and can be therefore used to assess uncertainty using simulations, this possibility has not been explored up to date. However, the assessments of uncertainty may be also included in a majority of demographic models (cohort-component, multi-regional, or multi-state) by feeding them at input with stochastic forecasts of particular components of demographic change. The latter may involve econometric forecasts and time series models, both in the sample-theory and the Bayesian frameworks.

3.3. Econometric forecasts of international migration

Since the 1990s, many studies have been published, predominantly in Austria and Germany, focusing on forecasting migration in Europe after the anticipated enlargement of the European Union. An overview of such studies has been recently presented in Alvarez-Plata et al. (2003), and in Brücker and Siliverstovs (2005). In the current paper only a selection of the models is presented, covering different modelling approaches. The notation follows the original studies, with the country subscripts always denoted i and j, respectively for the origin and destination countries, and the normal distributions presented as N(μ, τ), with mean μ, and precision parameter τ being a reciprocal of the variance, $\tau = \sigma^{-2}$.

Franzmeyer and Brücker (1997) built a gravity model of net migration between the regions i and j, based on the logarithm of the difference of the GDP per capita (PPP-adjusted), $\ln(Y_i / Y_j)$. The model has been calibrated on the basis of the empirical analysis of the elasticity of migration on income differentials in Europe presented in Barro and Sala-i-Martin (1995), where an income gap of 10% was found to drive 0.064% of the population of the less wealthy region out from their place of origin. Due to this assumption, the model yielded extremely high migration forecasts, with migration from the 10 Central-Eastern European countries to the EU-15 in the height of 590–1,180 thousands a year, depending on the pace of convergence of income levels between the origin and the destination countries.

Fertig and Schmidt (2000) modelled rates of migration to Germany, m, from the four then-candidate countries (Poland, the Czech Republic, Hungary and Estonia), with the error term decomposed into the country-specific, time-specific, and cross-sectional components:

$$m_{i,t} = \mu + \varepsilon_i + \varepsilon_t + \varepsilon_{i,t}, \tag{3.1}$$

where i denotes country of origin, $\varepsilon_i \sim N(0, \tau_i)$, $\varepsilon_{i,t} \sim N(0, \tau_{i,t})$, and ε_t is a Gaussian autoregressive process AR(1). This forecast yielded an average population inflow to Germany totalling between 15 and 57 thousand migrants a year in the period 1998–2017.

Sinn et al. (2001) forecasted stocks of foreign population in Germany (B) originating from the largest then-candidate countries: Poland, Romania, the Czech Republic, Hungary and the Slovak Republic. The authors used a partial adjustments model:

$$B_t = \lambda\,[\alpha_0 + \alpha_1\, YV_t + \alpha_2\, G_t + \alpha_3\, EU_t + \alpha_4\, FR_t + (1/\lambda - 1 + \alpha_5)\, B_{t-1}] + \varepsilon_t, \tag{3.2}$$

where λ satisfies the condition: $B_t = B_{t-1} + \lambda\,(B_t^* - B_{t-1})$, and B_t^* is a 'long-term equilibrium' of the foreign population stocks under study. The remaining explanatory variables are: YV – a fraction of GDP *per capita* (PPP-adjusted), G – an output gap, EU – a dummy concerning the EU membership, and FR – a dummy concerning freedom of movement of the labour force. The model has been partially calibrated on the historical data for German population stocks originating from Greece, Italy, Portugal, Spain, and Turkey. The B_t was forecasted to increase from the initial 459 thousand to 3.2–4.1 million people by 2015. A similar model has been applied by Brücker and Siliverstovs (2005), with the error term decomposed into the country-specific effect and the white noise.

Alvarez-Plata et al. (2003) prepared a forecast of the post-enlargement migration to the EU-15 from the ten countries of Central and Eastern Europe. Dependent variables are the rates of migration from county i to j, relative to the population size of the origin country (ms):

$$\begin{aligned} ms_{i,j,t} = {} & \alpha + (1 - \delta)\, ms_{i,j,t-1} + \beta_1 \ln(w_{j,t} / w_{i,t}) + \beta_2 \ln(w_{i,t}) + \beta_3 \ln(e_{i,t}) + \beta_4 \ln(e_{j,t}) + \\ & + \beta_5 \ln(P_{i,t}) + \gamma Z_{i,j} + u_{i,j,t}, \end{aligned} \tag{3.3}$$

where $u_{i,j,t} = \mu_{i,j} + v_{i,j,t}$, and $v_{i,j,t}$ is the white noise. The other explanatory variables are: w – real income levels, e – employment rates, P – population sizes, and Z – cross-country dummy variables denoting the geographic and cultural proximity of particular countries. Under the assumption of a long-term convergence of the economic explanatory variables to their EU-15 levels, the migration from the ten countries of Central and Eastern Europe is forecasted to decline exponentially from 367 thousand a year, shortly after the introducing the freedom of labour force movement in the enlarged EU, to the levels below zero in 2030.

Although the econometric models incorporate the analysis of uncertainty, this issue is not given proper attention in many studies devoted to migration forecasting. In some studies there are also problems with the model specification (Kupiszewski, 2002). Firstly, if numbers of migrants are forecasted, as in Franzmeyer and Brücker (1997), instead of rates denoting the migration risk, there is a lack of control on the demographic characteristics of the populations under study (size and age structure), which may lead to extreme results, as shown above. Even when forecasting rates, if the population size is one of the explanatory variables, as in Alvarez-Plata et al. (2003), the population movements occur *de facto* outside the model, being another source of bias. Finally, many independent variables used as predictors (GDP, unemployment, etc.) can be equally or more difficult to forecast than the dependent variable, that is migration.

3.4. Stochastic forecasts of migration time series

Another important class of stochastic models used in migration forecasting are the models based on the analysis and extrapolation of the time series. Most frequently it is done by applying various ARIMA models (Box and Jenkins, 1976), mainly within the framework of the sampling-theory statistics. For example, de Beer (1997) modelled the total volume of emigration from, as well as immigration to the Netherlands using the AR(1) autoregressive models, $x_t = c + \phi\, x_{t-1} + \varepsilon_t$, where ε_t denotes white noise, while the moving average process MA(1), $x_t = c + \varepsilon_t - \theta\, \varepsilon_{t-1}$, has been found suitable for net migration. For Finland, Alho (1998) used the ARIMA(0,1,1) models, $x_t = c + x_{t-1} + \varepsilon_t - \theta\, \varepsilon_{t-1}$, for the logarithms of immigration, as well as for emigration volumes. Keilman et al. (2001) made a probabilistic population forecast for Norway with the ARMA(1,1) model for the log of immigration, $x_t = c + \phi\, x_{t-1} + \varepsilon_t - \theta\, \varepsilon_{t-1}$, and the ARIMA(0,1,0), i.e. random walk with drift, for the log of emigration, $x_t = c + x_{t-1} + \varepsilon_t$.

Multivariate generalisations of the time series models can be used to include other explanatory variables, as in the VAR(4) vector autoregressive models of Gorbey et al. (1999), used for forecasting migration between Australia and New Zealand. The model, based on quarterly data, has the form:

$$\mathbf{X}_t = \mathbf{C_0} + (\mathbf{C_1}\, \mathrm{L} + \mathbf{C_2}\, \mathrm{L}^2 + \mathbf{C_3}\, \mathrm{L}^3 + \mathbf{C_4}\, \mathrm{L}^4)\, \mathbf{X}_t + \varepsilon_t, \tag{3.4}$$

where $\mathbf{C}_i$ are coefficient matrices, L is the lag operator ($\mathrm{L}(\mathbf{X}_t) = \mathbf{X}_{t-1}$), and ε_t is a multi-dimensional white noise. The authors tested four models based on different vectors of interdependent variables, $\mathbf{X}_t$, 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. Although the authors remarked that the movements between Australia and New Zealand are largely visa-free and they resemble internal migration, similar models can be used also for typically international population flows, after adjusting them to include additional policy-related variables, etc.

An example of a partial departure from the sampling-theory statistical paradigm in demographic forecasting is the concept of the 'expert-based probabilistic population projections' developed by Lutz et al. (1996, 1998, 2004). The method applies subjective expert judgement to set the framework for the stochastic forecasts. In formal terms, with v_t denoting the phenomenon under study (here: migration), the forecasting model is $v_t = \bar{v}_t + \varepsilon_t$, where $\bar{v}_t$ is the average trajectory of the process, assumed *a priori* by the experts, and ε_t is a random process, e.g., AR(p) or MA(q). Lutz et al. (2004) applied $\varepsilon_t \sim$ MA(30), assuming additionally that the standard deviation of ε_t, $\sigma(\varepsilon_t)$, is equal to a pre-defined value $\sigma^*(\varepsilon_t)$, also set on the basis of the expert judgement. For migration, it was assumed that $\bar{v}_t = \bar{v}$ (the mean of the process is time-invariant), while $\sigma^*(\varepsilon_t)$ has been selected so that 80% of the density of the probability distribution of v_t is concentrated between zero and the arbitrarily-chosen value $v_{\max}$. Due to the explicitly expressed subjectivity, this approach can be seen as a hybrid between the traditional and Bayesian methods.

The key advantage of all time series models used in migration forecasting is that they explicitly include the analysis of uncertainty. Using the VAR models additionally allows for including additional predictors of migration, without having to forecast them separately.

3.5. Existing Bayesian models and forecasts of population flows

The existing examples of international migration models and forecasts based on the Bayesian framework are scarce. Gorbey et al. (1999) extended their aforementioned VAR(4) analysis of migration between Australia and New Zealand (3.4) to the Bayesian case. They used the *Minnesota priors* for model coefficients, with parameters on the first lags of the same variables following a priori a normal distribution N(1, $\tau_{i,i,1}$) and the remaining parameters N(0, $\tau_{i,j,k}$) for interrelations between the i-th variable and the k-th lag of the j-th variable, $i \neq j$ or $k > 1$. This reflects an assumptions that the time series of each variable is most likely generated independently by a random walk process. Further, they assumed that $s_{i,j,k} = \tau_{i,j,k}^{-0.5} =$ $= \gamma \cdot g(k) \cdot f(i,j) \cdot s_i / s_j$, where: $\gamma = 0.4$, $g(k) = k^{-1}$, $f(i,i) = 1$, $f(i,j) < 1$ for $i \neq j$, and s_i denotes a standard error in the autoregressive model for the i-th variable. Note, that as the s_i values have to be estimated from the observations and thus the priors are data-based, this is not a fully Bayesian approach, where prior distributions are specified independently from the data. In the *ex-post* analysis, the authors found that the best Bayesian model (with $\mathbf{X}_t$ comprised of net migration rates, the growth of the real GDP ratio for the two countries, and quarterly unemployment growth in Australia) performed slightly worse than the corresponding traditional VAR, likely due to the disagreement between the priors and the data.

Examples of applications of Bayesian gravity models of population flows are in Congdon (2000, 2001). Although in these studies the flows of patients to hospitals are modelled rather than migration as such, the models can be generalised to cover also the other types of spatial movements. Assuming that a number of patients from region i to the hospital j has a Poisson distribution with a mean ($\mu_{i,j}$), the models can be defined as (Congdon, 2001):

$$\ln(\mu_{i,j}) = k + \alpha_0 \ln(P_i) + \delta \ln(R_{i,j}) + \xi_1 E_{1,j} + \xi_2 E_{2,j} + \phi S_{i,j}, \qquad (3.5a)$$

or, alternatively:

$$\ln(\mu_{i,j}) = k + \alpha_0 \ln(P_i) + \alpha_1 YAN_i + \alpha_2 Aged_i + \delta_i \ln(R_{i,j}) + \xi_1 E_{1,j} + \xi_2 E_{2,j} + \phi S_{i,j}. \qquad (3.5b)$$

The main explanatory variables are: P_i – population size of the i-th region, and $R_{i,j}$ – 'supply' of medical services – number of beds in the j-th hospital, B_j, weighted by an average distance (crow-fly or car-time), from the i-th region to the j-th hospital, $d_{i,j}$. The other predictors include: YAN_i – an index of demand for the health-care services, $Aged_i$ – a fraction of population aged 65+, $E_{1,j}$, $E_{2,j}$ – dummy variables for two selected hospitals, and $S_{i,j}$ – a dummy indicator, whether the j-th hospital is located in the i-th region. The prior distributions are normal, diffuse for the constant, $k \sim N(1, 0.0001)$, and more informative for the remaining parameters: α_i, β, $\delta \sim N(1, 0.1)$, $\gamma \sim N(2, 0.1)$, and ξ_1, ξ_2, $\phi \sim N(0, 0.1)$. The second model, in its variant based on the car-time distance, has been found best fitting the data.

The study of Brücker and Siliverstovs (2005) also contains some elements of a Bayesian analysis. The latter, however, is 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. Their *ex-post* comparison of various estimation methods for a partial-adjustment model showed that the hierarchical Bayes estimator (likely the mean in the appropriate posterior distribution) and the sampling-theory fixed effects estimator performed best.

As in the case of the sampling-theory time series models, the Bayesian ones address the issue of uncertainty in an explicit way and allow for including additional explanatory variables in the VAR models. Furthermore, the subjective expert knowledge on the characteristics of the processes and on interactions between variables can be incorporated in the models in the form of the prior distributions of the parameters.

4. Examples of simple Bayesian models for forecasting international migration

4.1. Data, specification and estimation of the models

The presented numerical example aims at producing Bayesian forecasts of long-term international migration flows between Poland and Germany for the years 2004–2010, prior to the opening of the German labour markets for Polish citizens, expected for 2011. Data on migration flows and population stocks for 1985–2003 predominantly come from the Eurostat. The time series of the economic explanatory variables: the GDP and unemployment rates are respectively from the databases of the United Nations Economic Commission for Europe and of the World Bank. The German data prior to 1991 concern West Germany. As in Poland before 1990 there was officially no unemployment[†], in the current study the respective rates have been assumed to equal 0.1, an arbitrarily chosen small positive number, in order to avoid problems with the logarithmic transformation of the variable.

[†] Hidden unemployment, the rate of which was estimated for Poland for the late 1980s about 25% (Rutkowski, 1990), has not been considered here, as this category cannot be seen as a push factor of migration. Hidden unemployment occurs, when a reduction of an excess employment in an economy would not lead to a decline in the output. This is not visibly related to the migration decisions of the individuals, as it is the case of the real unemployed looking for a job and better life perspectives.

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). The numbers of registered long-term migrants have been transformed into crude occurrence-exposure emigration rates, which are subject to forecasting, both for migration from Poland to Germany, and vice versa. For population at risk, the denominator of crude migration rates, the mid-year population has been used. With respect to discontinuity in the trends of Polish population stocks for 1988–2002, caused by the underestimation of international emigration, a correction has been applied on the basis of the results of the population census from 2002. The post-census adjustment has been distributed throughout the period 1988–2002 proportionally to the registered balance of migration between Poland and Germany, as registered by German official statistics.

Let us denote emigration rate from Poland to Germany per 1,000 inhabitants of the sending country by $MR_{P\text{-}D}$, and from Germany to Poland by $MR_{D\text{-}P}$. Further, let the GDP per capita according to the purchasing power parity (PPP, in 2003 international dollars) in Germany be noted as GDP_D, in Poland as GDP_P, and the respective unemployment rates as UR_D and UR_P. Additionally, a dummy variable Z, equalling 1 for the years 1988–1990 and 0 otherwise, is introduced to account for the shock related to the economic transition in the late 1980s. For $MR_{P\text{-}D}$, we consider the following three types of models:

1. An autoregressive process AR(1): $\ln(MR_{P\text{-}D}(t)) = c + \alpha \cdot \ln(MR_{P\text{-}D}(t-1)) + \beta \cdot Z(t) + \varepsilon(t)$, where $\varepsilon(t) \sim \mathrm{N}(0, \tau)$, τ being the precision parameter, a reciprocal of the variance. The logarithm transformation has been used, as $MR_{P\text{-}D}$ is by definition positive. For c and α, diffuse (rather non-informative) prior distributions N(0, 0.001) are assumed. Parameter β is expected to follow a relatively informative prior distribution N(1, 0.1), assuming that the system transformation in the years 1988–1990 contributed to the magnification of the $\ln(MR_{P\text{-}D})$ by 1 unit on average. Finally, τ is *a priori* expected to follow a chi-squared distribution with one degree of freedom, reflecting a belief in a low precision.

2. A vector autoregressive process VAR(1): $\mathbf{X_1}(t) = \mathbf{c} + \mathbf{A} \cdot \mathbf{X_1}(t-1) + \boldsymbol{\beta} \cdot Z(t) + \boldsymbol{\varepsilon}(t)$, where $\mathbf{X_1}(t) = [\ln(MR_{P\text{-}D}(t)), \ln(GDP_D(t)/GDP_P(t))]^{\mathrm{T}}$ refers to a hypothesis of a role of income differentials in explaining international migration, $\mathbf{c} = [c_1, c_2]^{\mathrm{T}}$, $\mathbf{A} = [\alpha_{i,j}]_{2\mathrm{x}2}$, $\boldsymbol{\beta} = [\beta, 0]^{\mathrm{T}}$ and $\boldsymbol{\varepsilon}(t) \sim \mathbf{N}(\mathbf{0}, \mathbf{T})$. The logarithms have been used for the same reason as in model 1. Analogously to the one-dimensional case, $\mathbf{N}(\mathbf{0}, \mathbf{T})$ denotes the two-dimensional normal distribution with mean $\mathbf{0} = [0, 0]^{\mathrm{T}}$ and precision matrix $\mathbf{T}$. The prior distributions for c_1 and c_2 are both diffuse, following N(0, 0.001). Parameters by own-variable lags ($\alpha_{1,1}$ and $\alpha_{2,2}$) follow N(1, 1), assuming a likely random-walk character of each of the variables separately. For parameters by cross-variable lags, $\alpha_{1,2} \sim$ N(0.5, 1) reflects the initial hypothesis of a positive impact of the income difference between destination and origin countries on migration, and $\alpha_{2,1} \sim$ N(0, 100) depicts firm prior beliefs of non-existence of an inverse relationship. The prior distribution for β is assumed to follow

N(1, 0.1) for the reasons mentioned before. **T** is assumed to follow a two-dimensional Wishart distribution with 2 degrees of freedom and the scale matrix $\mathbf{P} = [p_{i,j}]_{2x2}$, where $p_{1,1} = p_{2,2} = 0.1$ and $p_{1,2} = p_{2,1} = 0.005$, reflecting beliefs in a relatively high precision for both variables independently, and a relatively low precision for their interrelations.

3. Another VAR(1) process, $\mathbf{X_2}(t) = \mathbf{c} + \mathbf{A} \cdot \mathbf{X_2}(t–1) + \beta \cdot Z(t) + \varepsilon(t)$, reflecting a hypothesis that unemployment in the sending country is an important push factor of migration, and thus with $\mathbf{x_2}(t) = [\ln(MR_{P\text{-}D}(t)), \ln(UR_P(t))]^T$. The prior distributions are assumed the same as in model 2, for the similar reasons to the ones discussed before.

The second set of models, for $MR_{D\text{-}P}$, is analogous, with the same prior distributions of parameters assumed for simplicity. These examples of migration forecasting models do not aim at covering a comprehensive set of all possible explanatory variables, but to serve merely as an illustration of an application of the Bayesian methodology in practice.

In computations, Markov chain Monte Carlo (MCMC) simulations have been used, implemented in the WinBUGS 1.4 software (Spiegelhalter et al., 2003). The WinBUGS code, drawing heavily on the examples presented in Congdon (2003: 172–175, 189–191, Programs 5.1 and 5.5), is listed in the Appendix.

4.2. Forecasts of international migration between Poland and Germany, 2004–2010

For each of the flows, from Poland to Germany and vice versa, three aforementioned models 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, except for model 3 for migration from Germany to Poland, where the first 100,000 iterations were discarded.

As each of the models was based on a different set of data, their formal comparison using for example the Deviance Information Criterion (DIC) incorporated in WinBUGS 1.4 was not possible. Instead, the goodness-of-fit of the three models has been compared simplistically, on the basis of the sum of squares (*SS*) of the residuals between the observed and estimated values of the respective ln(*MR*). Results of estimation of the models are presented in Table 1, containing summaries of posterior distributions of the parameters of all three forecasting models for each of the migratory flows. Table 2 presents summaries of empirical distributions of *SS* obtained from 100,000 samples of the MCMC algorithm.

Summaries of predictive distributions obtained for 2004, 2007, and 2010 in all three models are presented in Table 3, for both $\ln(MR_{P\text{-}D})$ and $\ln(MR_{D\text{-}P})$.

Table 1. Summaries of posterior densities of the parameters of three forecasting models, estimated by the MCMC

Model	Migration from Poland to Germany, ln(MR_{P-D})						Migration from Germany to Poland, ln(MR_{D-P})					
	Parameter	Mean	St. Dev.	0.025	Median	0.975	Parameter	Mean	St. Dev.	0.025	Median	0.975
Model 1	α	0.28	0.23	–0.17	0.28	0.73	α	0.51	0.24	0.04	0.51	0.98
	β	0.96	0.31	0.35	0.96	1.57	β	0.53	0.21	0.11	0.53	0.94
	c	0.70	0.25	0.21	0.70	1.19	c	–0.02	0.08	–0.17	–0.02	0.14
	τ	10.16	3.61	4.35	9.74	18.34	τ	13.63	4.84	5.83	13.07	24.61
Model 2	$\alpha_{1,1}$	0.36	0.14	0.09	0.36	0.63	$\alpha_{1,1}$	0.56	0.13	0.31	0.56	0.81
	$\alpha_{1,2}$	–0.53	0.29	–1.09	–0.54	0.04	$\alpha_{1,2}$	0.19	0.22	–0.24	0.19	0.65
	$\alpha_{2,1}$	0.10	0.04	0.02	0.10	0.17	$\alpha_{2,1}$	–0.13	0.05	–0.23	–0.13	–0.02
	$\alpha_{2,2}$	0.88	0.12	0.64	0.88	1.11	$\alpha_{2,2}$	0.82	0.12	0.58	0.81	1.06
	β	0.78	0.19	0.39	0.78	1.16	β	0.47	0.12	0.23	0.47	0.70
	c_1	1.17	0.29	0.58	1.18	1.74	c_1	0.18	0.23	–0.27	0.18	0.65
	c_2	0.02	0.14	–0.24	0.01	0.29	c_2	–0.18	0.12	–0.42	–0.18	0.07
	$t_{1,1}$	31.62	11.09	13.79	30.31	56.73	$t_{1,1}$	60.64	21.39	26.24	58.13	109.20
	$t_{2,2}$	158.00	54.12	70.32	152.00	280.30	$t_{2,2}$	156.40	54.25	68.86	150.10	278.70
	$t_{1,2} = t_{2,1}$	–8.81	17.30	–44.54	–8.23	24.16	$t_{1,2} = t_{2,1}$	–3.21	23.95	–51.09	–3.09	44.01
Model 3	$\alpha_{1,1}$	0.42	0.13	0.17	0.42	0.68	$\alpha_{1,1}$	0.33	0.15	0.03	0.32	0.64
	$\alpha_{1,2}$	–0.04	0.03	–0.10	–0.04	0.02	$\alpha_{1,2}$	–0.35	0.21	–0.76	–0.35	0.10
	$\alpha_{2,1}$	0.05	0.10	–0.14	0.05	0.25	$\alpha_{2,1}$	–0.08	0.09	–0.25	–0.08	0.10
	$\alpha_{2,2}$	0.86	0.10	0.65	0.86	1.06	$\alpha_{2,2}$	0.78	0.20	0.40	0.78	1.17
	β	0.72	0.18	0.36	0.73	1.08	β	0.54	0.09	0.35	0.54	0.72
	c_1	0.62	0.17	0.28	0.62	0.94	c_1	0.70	0.44	–0.22	0.70	1.53
	c_2	0.40	0.28	–0.14	0.40	0.98	c_2	0.46	0.40	–0.33	0.47	1.22
	$t_{1,1}$	42.11	14.87	18.23	40.36	75.71	$t_{1,1}$	75.30	26.54	32.61	72.21	135.60
	$t_{2,2}$	2.15	0.80	0.91	2.04	4.02	$t_{2,2}$	47.61	16.33	21.47	45.70	84.64
	$t_{1,2} = t_{2,1}$	5.48	2.95	0.53	5.19	12.04	$t_{1,2} = t_{2,1}$	–3.30	15.60	–34.57	–3.21	27.29

Source: Own calculations in WinBUGS

Table 2. Precision of estimation: summaries of distributions of sums of squares (*SS*) obtained from 100,000 MCMC samples

Model	Migration from Poland to Germany, ln(MR_{P-D}) Parameter	Mean	St. Dev.	0.025	Median	0.975	Migration from Germany to Poland, ln(MR_{D-P}) Parameter	Mean	St. Dev.	0.025	Median	0.975
Model 1	*SS*	0.91	0.33	0.60	0.82	1.79	*SS*	0.43	0.25	0.19	0.36	1.08
Model 2	*SS*	0.57	0.13	0.43	0.53	0.91	*SS*	0.24	0.07	0.17	0.23	0.42
Model 3	*SS*	0.66	0.20	0.44	0.61	1.17	*SS*	0.18	0.06	0.12	0.16	0.34

Source: Own calculations in WinBUGS

Table 3. Summaries of predictive distributions of the ln(*MR*) values forecasted for 2004, 2007, and 2010, estimated by the MCMC

Model	Migration from Poland to Germany, ln($MR_{P–D}$) Year	Mean	St. Dev.	0.025	Median	0.975	Migration from Germany to Poland, ln($MR_{D–P}$) Year	Mean	St. Dev.	0.025	Median	0.975
Model 1	2004	0.97	0.37	0.23	0.98	1.71	2004	–0.04	0.35	–0.76	–0.04	0.64
	2007	0.97	0.40	0.18	0.97	1.76	2007	–0.05	0.44	–0.95	–0.04	0.75
	2010	0.97	0.42	0.17	0.97	1.75	2010	–0.07	0.60	–1.08	–0.04	0.78
Model 2	2004	1.07	0.23	0.60	1.07	1.53	2004	–0.01	0.18	–0.36	–0.01	0.36
	2007	1.07	0.27	0.55	1.06	1.62	2007	0.00	0.22	–0.42	0.00	0.44
	2010	1.05	0.29	0.49	1.05	1.65	2010	–0.01	0.24	–0.46	–0.01	0.46
Model 3	2004	0.88	0.27	0.34	0.88	1.40	2004	–0.08	0.16	–0.40	–0.08	0.24
	2007	0.84	0.34	0.12	0.86	1.46	2007	–0.08	0.23	–0.55	–0.08	0.33
	2010	0.81	0.40	–0.05	0.85	1.48	2010	–0.09	0.41	–0.72	–0.07	0.40

Source: Own calculations in WinBUGS

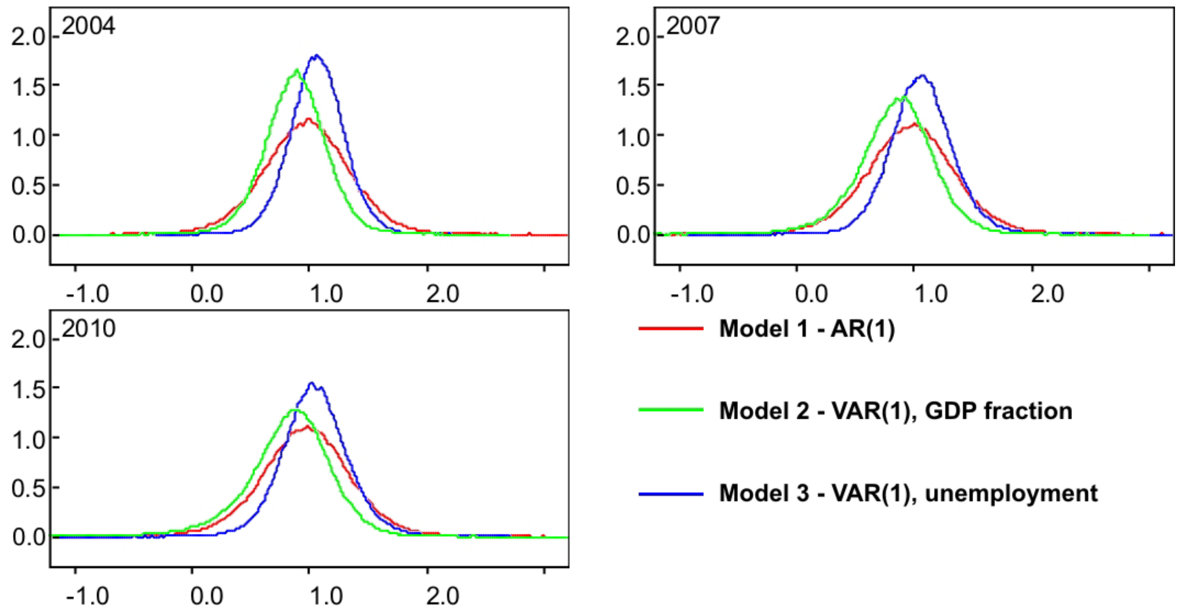
The best fit, according to the *SS*, has been obtained for model 2 for migration from Poland to Germany, reflecting the role of income differentials, and for model 3 for migration from Germany to Poland, corresponding to the unemployment hypothesis. Results of the forecasts are presented in Figures 1 and 2 for flows from Poland to Germany, and in Figures 3 and 4 for flows from Germany to Poland. Figures 1 and 3 show the estimated kernel densities of predictive distributions of the respective ln(*MR*) forecasts for 2004, 2007, and 2010. Figures 2 and 4 show the respective *MR* data series for 1985–2003, as well as the out-of-sample forecasts for 2004–2010 for the three models defined in Section 4.1, transformed back from logarithms to crude rates. For all models, the posterior quantiles of rank 0.025, 0.5 and 0.975 are presented.

For many parameters, with the exception of most own-variable autoregression coefficients (α, $\alpha_{1,1}$, and $\alpha_{2,2}$), and the dummy-related coefficients β, the intervals between the quantiles of rank 0.025 and 0.975 are wide and often cover 0 (Table 1). For migration from Germany to Poland all constants are thus 'insignificant', as are most of the interaction parameters $\alpha_{1,2}$ and $\alpha_{2,1}$ in models 2 and 3. For migration from Poland to Germany, this is the case of the autoregression coefficient α in model 1, constants c_2 and coefficients $\alpha_{1,2}$ in models 2 and 3, as well as of $\alpha_{1,2}$ in model 3 only. For this reason, the inference on constants and on interactions between ln(*MR*) and other variables in models 2 and 3, apart from the dummies, is in many cases very vague.

The estimated posterior precision (parameters τ or **T**) usually appeared to be high, contrary to the prior beliefs. In that respect, the data have had more weight in the posterior distributions than the priors, an exception being ln($UR_P(t)$) in model 3 for flows from Poland to Germany. Knowing that, an additional analysis might be additionally performed on the robustness on the results, taking into consideration prior beliefs in more precision, not discussed in this paper.

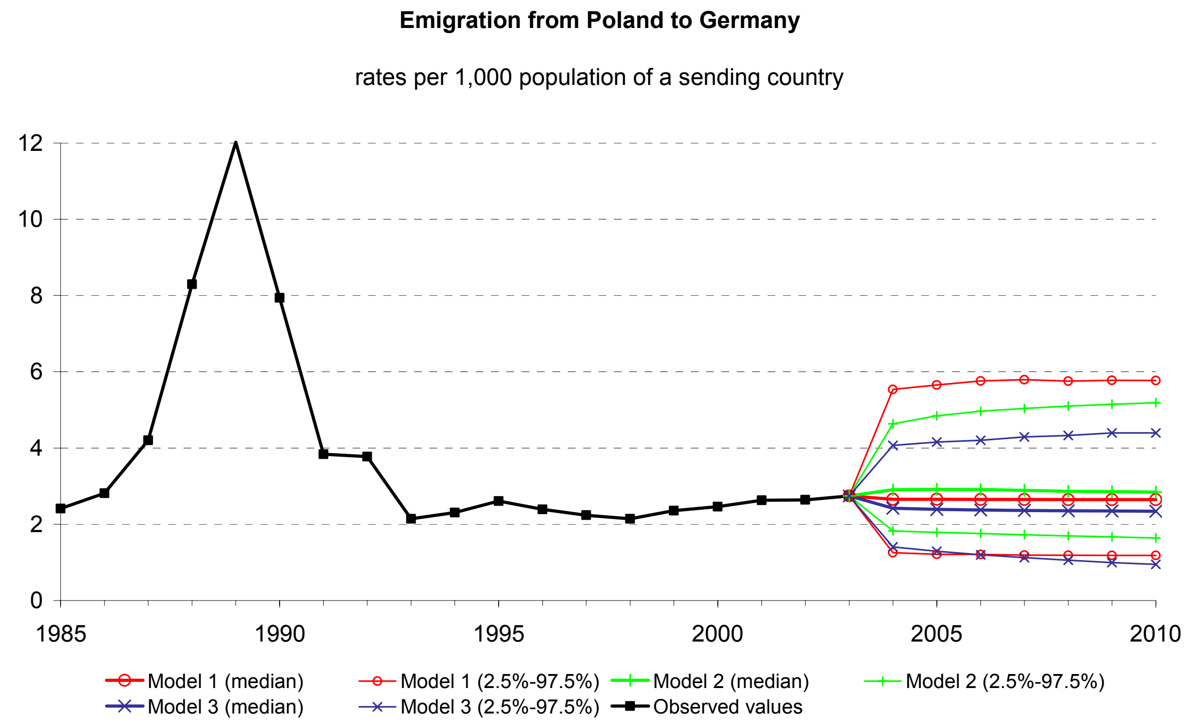
Almost all own-variable autoregression coefficients, with the exception of α in model 1 for migration from Poland to Germany, are highly likely positive and smaller than one. This indicates a long-term stationarity of the logarithmic transformations of the variables under study (*MR*, *GDP* fraction, and *UR*). In the case of models 2 and 3 for migration from Poland to Germany, and of model 3 for migration from Germany to Poland, interpretation of the model-specific coefficients $\alpha_{1,2}$ is counter-intuitive. With a decreasing income gap or unemployment in the sending country, one would expect a decreasing $MR_{P\text{-}D}$, thus a positive $\alpha_{1,2}$ rather than a negative one, whereas for the mentioned coefficients most of the probability mass is concentrated below zero. This conclusion, however, needs to be treated with caution, due to the fact that the 95% credible intervals cover zero in all mentioned cases.

Figure 1. Predictive densities of $\ln(MR_{P\text{-}D})$ for 2004, 2007, and 2010, estimated by the MCMC



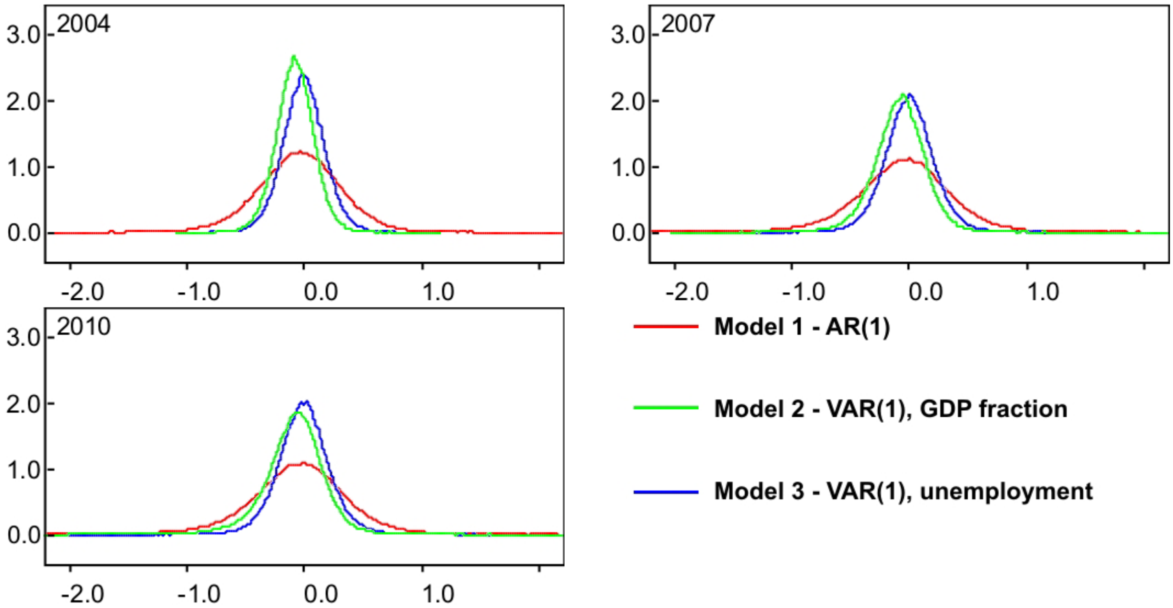
Source: Own calculations in WinBUGS

Figure 2. $MR_{P\text{-}D}$: observed for 1985–2003, forecasted for 2004–2010 (predictive quantiles)



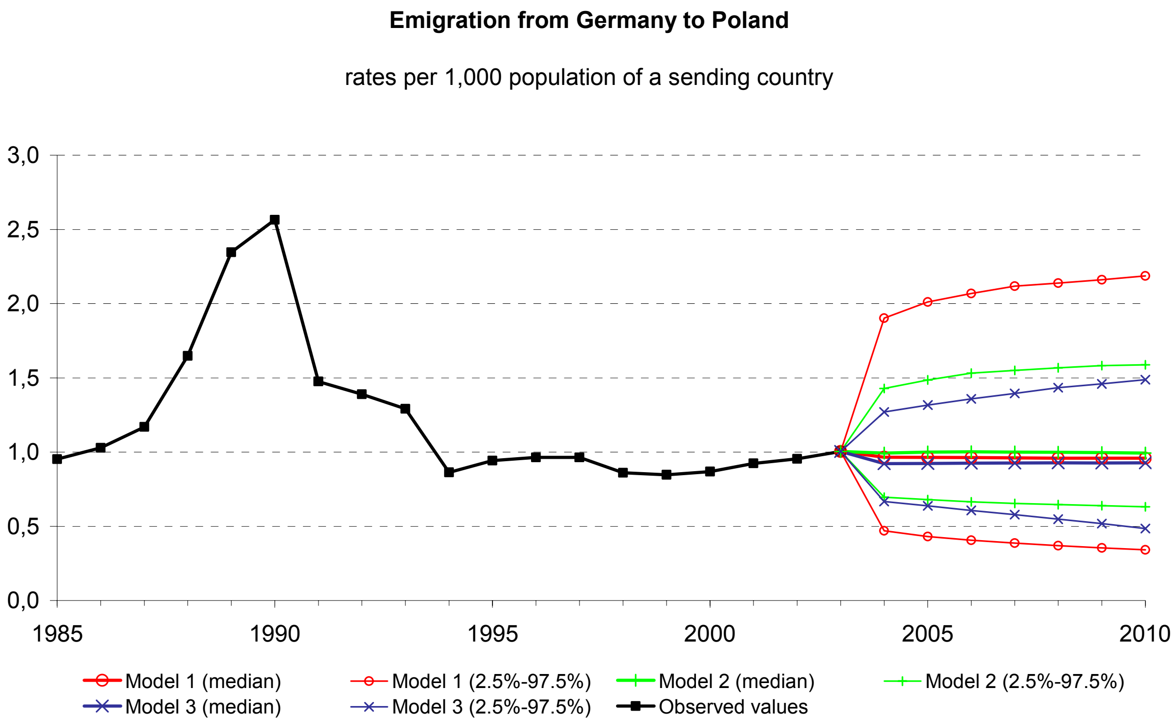
Source: Own calculations in WinBUGS

Figure 3. Predictive densities of $\ln(MR_{D\text{-}P})$ for 2004, 2007, and 2010, estimated by the MCMC



Source: Own calculations in WinBUGS

Figure 4. $MR_{D\text{-}P}$: observed for 1985–2003, forecasted for 2004–2010 (predictive quantiles)



Source: Own calculations in WinBUGS

For all models, the dummy-related coefficients β are most likely positive, equalling between 0.47 and 0.54 for migration from Germany to Poland, and between 0.78 to 0.96 for flows in the opposite direction. This indicates that the system transformation significantly contributed to the increase of migration between the two countries in 1988–1990.

The median forecasts of migration from Poland to Germany predict a stabilisation of $MR_{P\text{-}D}$ around 2.65 in model 1, or a slight declining tendency: from 2.90 to 2.84 throughout the forecast horizon in model 2, or from 2.42 to 2.34 in model 3. The median forecasts yielded by model 1 are thus in-between the ones produced by models 2 and 3. All forecasts, especially produced by model 1, are characterised by relatively wide uncertainty spans, as indicated by the predictive quantiles shown in Figure 2.

The predicted median-forecast rates of emigration from Germany to Poland are expected to be almost stable throughout the forecast horizon in all models, with values either equal, or just below one. Again, model 1 produced the median forecasts that are in-between the ones yielded by models 2 (higher) and 3 (lower) in the whole period 2004–2010. The uncertainty range is also large in all forecasting models, especially in model 1 (Figure 4). In the presented examples, adding explanatory variables to the model visibly reduced the uncertainty of the forecasts.

5. Conclusions

There are three major advantages of using the Bayesian methodology in the context of international migration forecasting. Firstly, its eclectic character allows for combining the positive features of various forecasting methods in a formal way, by the means of Bayesian statistics and the subjective theory of probability. This approach also offers simple tools for model selection and averaging, not exploited in the current example. Secondly, the quantitative analysis of uncertainty with respect to the future developments of phenomena under study is inherent in the Bayesian forecasts, which yield whole predictive distributions. The latter can serve as a natural and straightforward way of obtaining projection variants. Thirdly, with informative prior distributions consistent with the observations, the Bayesian estimates and forecasts are expected to carry smaller prediction errors than the sampling-theory ones, what is especially important in the small-sample cases. Although the major disadvantage of the approach is the computational complexity, it can be overcome by using numerical methods (Markov chain Monte Carlo) included in an easily available free software (e.g., WinBUGS), in order to obtain meaningful forecasts, as it has been shown in Section 4.

Bayesian statistics also allows for eliminating some problems with interpretation of the results, caused by relating probability to the frequency of events, as it is usually done in the sampling-theory approach. This is crucial with respect to the repeatable sample assumption, which is not a natural premise of analysis in the social sciences, including migration research.

Further studies in Bayesian migration forecasting, not covered by the current paper, may additionally include the issues of formal model selection, model averaging, and robustness of the results on different types of the prior information.

References

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

Alvarez-Plata, P., H. Brücker, and B. Siliverstovs (2003). *Potential Migration from Central and Eastern Europe into the EU-15 – An Update*. Report for the EC DG of Employment and Social Affairs. Berlin: Deutsches Institut für Wirtschaftsforschung.

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.

Barro, R. J. and X. Sala-i-Martin (1995). *Economic Growth*. New York: McGraw-Hill.

Bauer, T., and K. F. Zimmermann (1999). *Assessment of Possible Migration Pressure and its Labour Market Impact Following EU Following EU Enlargement to Central and Eastern Europe*. IZA Research Report 3. Bonn: Institut zur Zukunft der Arbeit.

Bayarri, M. J., and J. O. Berger (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., and A. F. M. Smith (2000). *Bayesian Theory*. Chichester: John Wiley.

Blumen, I., M. Kogan, and P. J. McCarthy (1955). *The Industrial Mobility of Labour as a Probability Process*. Ithaca, NY: Cornell University Press.

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

Brown, L. A. (1970). On the use of Markov chains in movement research. *Economic Geography*, 46 (Suppl.), 393–403.

Brücker, H., and B. Siliverstovs (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. (2001). The development of gravity models for hospital patient flows under system change: a Bayesian modelling approach. *Health Care Management Science*, 4 (4), 289–304.

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

Courgeau, D. (1985). Interaction between Spatial Mobility, Family and Career Life Cycle: A French Survey. *European Sociological Review*, 1 (2), 139–162.

Daponte, B. O., J. B. Kadane, and L. J. Wolfson (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.

Drbohlav, D. (1995). Abschätzung der zukünftigen Ost-West-Migration in Europa. Eine Delphi-Studie. In: H. Fassmann, and R. Münz (eds.), *Die Zukunft der Ost-West Wanderung. Quantitative Entwicklungen, ökonomische Konsequenzen und politische Antworten.* Wien: Österreichische Akademie der Wissenschaften, 24–49.

Eurostat (2006). *Population projections for the European Union 2004–2050*. Luxembourg: Eurostat, forthcoming.

Fassmann, H., and C. Hintermann (1997). *Migrationspotential Ostmitteleuropa*. ISR-Forschungsberichte, 15. Wien: Institut für Stadt- und Regionalforschung.

Fertig, M., and C. M. Schmidt (2000). *Aggregate-level migration studies as a tool for forecasting future migration streams*. IZA Discussion paper 183, Bonn: Institut zur Zukunft der Arbeit.

Franzmeyer, F., and H. Brücker (1997). *Europäische Union: Osterweiterung und Arbeitskräftemigration*. DIW Wochenbericht 5/97. Berlin: Deutsches Institut für Wirtschaftsforschung, 89–96.

Ginsberg, R. B. (1971). Semi-Markov Processes and Mobility. *Journal of Mathematical Sociology* 1 (1), 233–262.

Ginsberg, R. B. (1979). Timing and duration effects in residence histories and other longitudinal data: I – stochastic and statistical models. *Regional Science and Urban Economics*, 9 (4), 311–331.

Girosi, F., and G. King (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.

Goodman, L. A. (1961). Statistical methods for the mover-stayer model. *Journal of the American Statistical Association*, 56 (296), 841–868.

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

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

IOM [International Organization for Migration] (1998). *Migration Potential in Central and Eastern Europe*. Geneva: IOM.

Isard, W. (1960). *Methods of regional analysis: an introduction to regional science*. New York: John Wiley.

Jaynes, E. T. (1976). Confidence intervals vs. Bayesian intervals. In: W. L. Harper, and 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., and J. O. Berger (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 and Zeitlinger.

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

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

Keyfitz, N. (1980). *Multidimensionality in Population Analysis*. IIASA Report RR-80-33. Laxenburg: International Institute for Applied System Analysis.

Keyfitz, N. (1981). The limits of population forecasting. *Population and Development Review*, 7 (4), 579–593.

Kupiszewski, M. (2002). *Modelowanie dynamiki przemian ludności w warunkach wzrostu znaczenia migracji międzynarodowych*. Warszawa: Instytut Geografii i Przestrzennego Zagospodarowania PAN.

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

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

Leslie, P. H. (1945). On the Use of Matrices in Certain Population Mathematics. *Biometrika*, 33 (3), 183– 212.

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

Lutz, W., W. C. Sanderson, and 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?* London: Earthscan, 397–428.

Lutz, W., W. C. Sanderson, and S. Scherbov (1998). Expert based probabilistic population projections. *Population and Development Review*, 24 (Suppl.), 139–155.

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

McGinnis, R., G. C. Myers, and J. Pilger (1963). *Internal migration as a stochastic process*. Paper for the 34th session of the International Statistical Institute, Ottawa, 21–29 August 1963.

NRC [National Research Council] (2000). *Beyond Six Billion*. Washington, DC: National Academies Press.

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

Prais, S. J. (1955). Measuring social mobility. *Journal of the Royal Statistical Society A*, 118, 56–66.

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 and Co.; New York: Harcourt, Brace and Co., 156–196.

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

Rees, P. H., and A. G. Wilson (1973). Accounts and models for spatial demographic analysis 1: Aggregate population. *Environment and Planning A*, 5 (1), 61–90.

Rogers, A. (1966). A Markovian Policy Model of Interregional Migration. *Papers of the Regional Science Association*, 17, 205–224.

Rogers, A. (1975). *Introduction to Multiregional Mathematical Demography*. New York: John Wiley.

Rutkowski, M. (1990). *Labour Hoarding and Future Unemployment in Eastern Europe: The Case of Polish Industry*. CEP Discussion Paper No. 6/1990. London: Centre of Economic Performance, London School of Economics.

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

Sinn, H.-W., G. Flaig, M. Werding, S. Munz, N. Duell, and H. Hofmann (2001). *EU-Erweiterung und Arbeitskräftemigration. Wege zu einer schrittweisen Annäherung der Arbeitsmärkte*. München: IFO-Institut für Wirtschaftsforschung.

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

Stewart, J. Q. (1941). An inverse distance variation for certain social influences. *Science*, 93 (2404), 89–90.

Stoto, M. A. (1983). The Accuracy of Population Projections. *Journal of the American Statistical Association*, 78 (381), 13–20.

Stouffer, S. A. (1940). Intervening opportunities: a theory relating mobility and distance. *American Sociological Review*, 5 (6), 845–867.

Tuljapurkar, S., and C. Boe (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.

Weidlich, W., and G. Haag (eds.) (1988). *Interregional migration: dynamic theory and comparative analysis*. Berlin-Heidelberg: Springer.

Willekens, F. (1995). Monitoring international migration flows in Europe. Towards a statistical data base combining data from different sources. *European Journal of Population*, 10 (1), 1–42.

Wilson, A. G. (1970). *Entropy in Urban and Regional Modelling*. London: Pion.

Wilson, A. G. (1981). *Catastrophe Theory and Bifurcation*. London: Croom Helm.

Wilson, T., and M. Bell (2004). *Australia's uncertain demographic future*. Discussion Paper 2003/04. Brisbane: Queensland Centre for Population Research, the University of Queensland.

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

■

Appendix: WinBUGS code for Bayesian AR(1) and VAR(1) forecasting models

A. Model AR(1)

```
model { # AR(1) with a dummy
# Priors
c ~ dnorm(0,0.001)
alpha ~ dnorm (0,0.001)
beta ~ dnorm (1,0.1)
tau ~ dchisqr(1)
# Data definitions
for (t in 1:n) { y[t] <- log(MR[t]) }
for (t in 1:N) { z[t] <- Dummy[t] }
# Model
for (t in 2:n) { mu[t] <- c + alpha * y[t-1] + beta * z[t]
                 y[t] ~ dnorm(mu[t], tau)
                 y.new[t] ~ dnorm(mu[t], tau)
                 sqresid[t] <- pow(mu[t] - y[t],2) }
ss <- sum(sqresid[2:n])
# Forecasts for t = n+1 … N
for (t in n+1:N) { mu.new[t] <- c + alpha * y.new[t-1] + beta * z[t]
                   y.new[t] ~ dnorm (mu.new[t], tau) }
}
```

B. Model VAR(1)

```
model { # VAR(1) with a dummy for migration
# Priors
c[1] ~ dnorm(0,0.001); c[2] ~ dnorm(0,0.001)              # constants
alpha[1,1] ~ dnorm(1,1); alpha[2,2] ~ dnorm(1,1)          # own variable lags
alpha[1,2] ~ dnorm(0.5,1); alpha[2,1] ~ dnorm(0,100)      # cross-variable lags
beta[1] ~ dnorm(1,0.1); beta[2] <- 0                      # dummy
T[1:2,1:2] ~ dwish(P[1:2,1:2],2)                          # scale matrix
P[1,1] <- 0.1; P[2,2] <- 0.1; P[1,2] <-0.005; P[2,1] <- 0.005
# Data definitions
for (t in 1:n) { y[t,1] <- log(VAR1[t]); y[t,2] <- log(VAR2[t]) }
for (t in 1:N) { z[t] <- Dummy[t] }
# Model
for (t in 2:n) { for (i in 1:2) { mu[t,i] <- alpha[i,1] * y[t-1,1] + alpha[i,2] * y[t-1,2] +
                                                            + beta[i] * z[t] + c[i] }
                 y[t,1:2] ~ dmnorm(mu[t,1:2],T[1:2,1:2])
                 y.new[t,1:2] ~ dmnorm(mu[t,1:2],T[1:2,1:2])
                 sqresid[t] <- pow(mu[t,1] - y[t,1],2) }
ss <- sum(sqresid[2:n])
# Forecasts for t = n+1 … N
for (t in n+1:N) { for (i in 1:2) { mu.new[t,i] <- alpha[i,1] * y.new[t-1,1] +
                                    + alpha[i,2] * y.new[t-1,2] + beta[i] * z[t] + c[i] }
                   y.new[t,1:2] ~ dmnorm(mu.new[t,1:2], T[1:2,1:2]) }
}
```

ISSN 1732-0631
ISBN 83-921915-5-2

Printed in Poland