Central European Forum for Migration Research

CEFMR ŚFBM

Środkowoeuropejskie Forum Badań Migracyjnych

Applicability of Bayesian methods in international migration forecasting

Jakub Bijak

Central European Forum for Migration Research

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

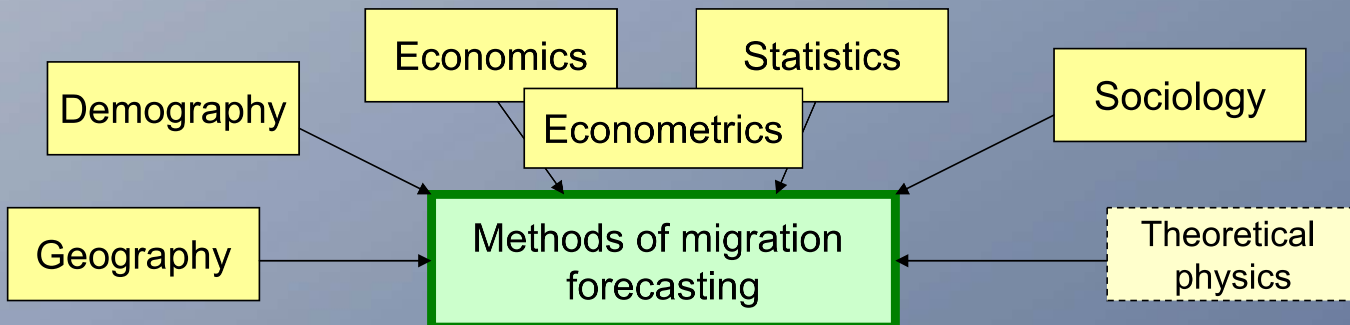
Project financed from the research grant of the Foundation for Population, Migration and Environment (BMU-PME) from Zurich

Plan of the presentation

1. **Introduction**
2. **Uncertainty and subjectivity in migration forecasting**
3. **Bayesian statistics: introductory notes**
4. **Existing migration models and forecasts**
   - Survey-based and Delphi migration scenarios
   - Mathematical models of population flows
   - Econometric forecasts
   - Time series models
   - Existing Bayesian solutions
5. **Bayesian forecasts of Polish-German migration to 2010**
6. **Conclusions**

1. Introduction

Methodological approaches in migration forecasting



International migration is a very complex and multi-dimensional phenomenon, therefore there are various approaches in the analysis, modelling, and forecasting of migration.

2. Uncertainty and subjectivity in migration forecasting

- Uncertainty is an immanent feature of every forecast
- Ways of dealing with uncertainty:
  - Ignoring (deterministic models, surveys, etc.)
  - Variant projections (uncertainty is not quantified)
  - Stochastic forecasts (uncertainty measured by probabilities of realisation within certain intervals)
- Subjectivity of the forecasts – including expert knowledge and judgement, often not explicitly, in:
  - Selection of the forecasting model and its assumptions
  - Construction of the scenarios of future migration patterns

3. Bayesian statistics: introductory notes

Basic terminology

– ***Prior knowledge*** reflects subjective beliefs (knowledge, intuition) of the researcher on the phenomena under study, unconditional on the observations.

– ***Posterior knowledge*** is a transformation of the prior knowledge, conditional on the observations (a sample from a statistical experiment).

– ***Subjective probability*** is a measure of uncertainty, reflecting subjective beliefs of the researcher, and independent from the frequency of the phenomena under study. The measure is probabilistic (fulfilling three axioms of Kolmogorov), or at least in the case of *improper distributions*, σ-finite.

3. Bayesian statistics: introductory notes

The Bayes Theorem

Let θ denote unknown model parameters, and y – data (observations).
Then (Bayes, 1763; Laplace, 1812):

$$p(\theta \mid y) = \frac{p(y \mid \theta) \cdot p(\theta)}{p(y)}$$

Posterior distribution → $p(\theta \mid y)$

Prior distribution → $p(\theta)$

Likelihood of the data, given θ (‘traditional’) → $p(y \mid \theta)$

Marginal likelihood of y (independent from θ) → $p(y)$

Central European Forum for Migration Research

Środkowoeuropejskie Forum Badań Migracyjnych

3. Bayesian statistics: introductory notes

Bayesian forecasting

Let y denote observed (past) values, and y^P – forecasted (future) values.

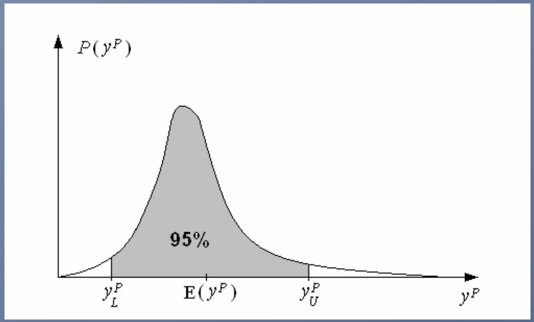
$$p(y^P \mid y) = \int_{\theta \in \Theta} p(y^P \mid y, \theta) \cdot p(\theta \mid y) d\theta$$

Result: predictive distribution

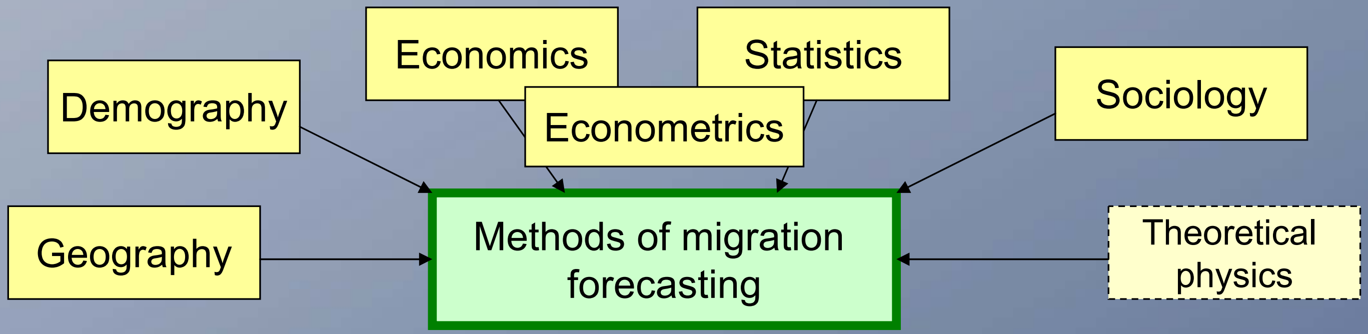
Sample-based predictions

Posterior distribution of θ

The outcome of a Bayesian forecast is the whole predictive distribution, and not a single value.



4. Existing migration models and forecasts



4. Existing migration models and forecasts

- 'Soft' approaches: sociological surveys, Delphi studies
- Mathematical models of population flows
  - Geographic approach (spatial structures): Markov chains, modes of spatial interactions
  - Demographic approach (age/sex structures): Cohort-component model, event history analysis
  - A synthesis of the demographic and geographic approaches: Multi-regional and multi-state models
- Economic approach (labour markets): econometric models and forecasts
- Stochastic approach: time series analysis
- Bayesian models and forecasts: rare examples

4. Existing migration models and forecasts

‘Soft’ approaches

- “Migration potential”
  - Methodology: various sociological survey studies (Fassmann & Hintermann, 1997; IOM, 1999)
  - Problems: definitions, formulation of questions, sample size, translation of declarations into actual migratory behaviour…
- Methods explicitly referring to the expert knowledge
  - Delphi studies (Drbohlav, 1995)
  - Surveys among experts (Bauer & Zimmermann, 1999)

4. Existing migration models and forecasts

Mathematical models of population flows

- Markov chains in migration modelling and forecasting
  - Homogeneous transition matrix: Prais (1955), Brown (1970)
  - Heterogeneous transition matrix: Rogers (1966), Joseph (1974)
  - Models with unobserved population heterogeneity: *mover-stayer* (Blumen *et al.*, 1955), different transition matrices for various subpopulations (Goodman, 1962)
  - Non-stationary transition matrices: „cumulative inertia” (McGinnis *et al.*, 1963), semi-Markov models (Ginsberg, 1971)
  - Evolution into population accounting models (Rees & Wilson, 1973), multi-regional (Rogers, 1975) / multi-state (Keyfitz, 1980)

4. Existing migration models and forecasts

Mathematical models of population flows

- Models of spatial interactions
  - Gravity models (Stewart, 1941; Isard, 1960)
  - Model of "intervening opportunities" (Stouffer, 1940): number of migrants is proportional to the "opportunities" at destination and inversely to the ones available at a smaller distance
  - Entropy, catastrophe theory, bifurcations (Wilson, 1967 & 1981)
- "Sociodynamics" (Weidlich & Haag, 1988)
  - A model of social processes using tools of theoretical physics
- Event history analysis (Courgeau, 1985)
  - Migration as a demographic event in an individual's life history
  - Methodology: Monte Carlo micro-simulations

4. Existing migration models and forecasts

Econometric forecasts

- Mainly forecasts of post-enlargement migration in the EU (popular in Germany and Austria since the 1990s)
- Several types of the models:
  - Gravity models, based mainly on the GDP per capita differences (e.g., Franzmeyer & Brücker, 1997), or additionally including other predictors, like employment, population size, dummies for 'proximity' between countries, etc. (Alvarez-Plata *et al.*, 2003)
  - Mixed-effects models (e.g., Fertig & Schmidt, 2000) with emigration rates modelled for as $m_{i,t} = \mu + \varepsilon_i + \varepsilon_t + \varepsilon_{i,t}$
  - Partial adjustment models (Sinn *et al.*, 2001; Brücker & Siliverstovs, 2005), with additional explanatory variables

4. Existing migration models and forecasts

Econometric forecasts - evaluation

- Some forecasts concentrate on the numbers of migrants – there is a lack of control of demographic variables, mainly population size and structure, leading to extreme results (e.g., Franzmeyer & Brücker, 1997)
- If population size is one of the explanatory variables, population movements occur outside of the model (e.g., Alvarez-Plata *et al.*, 2003)
- In general, certain socio-economic explanatory variables (e.g., unemployment) may be even more difficult to predict than migration itself

4. Existing migration models and forecasts

Time series models

- The Netherlands (de Beer, 1997)
  IM_t, EM_t ~ AR(1), NM_t ~ MA(1)
- Finland (Alho, 1998)
  Ln(IM_t), Ln(EM_t) ~ ARIMA(0,1,1)
- Australia-New Zealand (Gorbey *et al.*, 1999)
  various VAR(4) models including net migration rates
- Norway (Keilman *et al.*, 2001)
  Ln(IM_t) ~ ARMA(1,1), Ln(EM_t) ~ ARIMA(0,1,0)

IM – immigration, EM – emigration, NM – net migration

4. Existing migration models and forecasts

Time series models

- "Expert-based probabilistic population projections" of Lutz *et al.* (1996–2004).
  - Model: $v_t = v^*_t + \varepsilon_t$, where v_t is the demographic phenomenon under study (here: migration), v^*_t – its average future trajectory assumed *a priori* by the experts, and ε_t is a random process. Lutz *et al.* (2004) assumed $\varepsilon_t \sim MA(30)$ (long memory).
  - Standard deviation of ε_t, $\sigma(\varepsilon_t)$, is assumed to equal a pre-defined value $\sigma^*(\varepsilon_t)$.
  - For migration, Lutz et al. (2004) assumed that $v^*_t = v^*$ (v^*_t is time-invariant), and $\sigma^*(\varepsilon_t)$ has been selected in such a way that 80% of the probability density of v_t falls between zero and a judgementally chosen v_{max}.

4. Existing migration models and forecasts

Existing Bayesian forecasts

- Migration between Australia and New Zealand (Gorbey *et al.*, 1999) – various Bayesian VAR(4) models tested, applying data-based (not fully Bayesian!) *Minnesota priors*
- Gravity models applied for patient flows in the UK (Congdon, 2000, 2001), informative normal priors
- Various econometric models of migration to Germany prepared also using Bayesian estimation by Brücker & Siliverstovs (2005), although presented in a non-Bayesian way (no mention of prior selection, no discussion of predictive / posterior distributions)

5. Bayesian forecasts of Polish-German flows

Data

- Aim: forecast of long-term migration flows between Poland and Germany until 2010
- Forecasted variable: logarithms of emigration rates per 1,000 population of the sending country
- Data series for 1985-2003. Source of population data: Eurostat, migration data: *Statistisches Bundesamt*
- Population stocks include post-census adjustments
- Other variables: GDP per capita (PPP-adjusted) and unemployment rates (*UR*). Data: UNECE, World Bank

5. Bayesian forecasts of Polish-German flows

Models

- Model 1 – an autoregressive process AR(1):
  $\ln(MR_t) = c + \alpha \cdot \ln(MR_{t-1}) + \varepsilon_t$, where $\varepsilon_t \sim N(0, \tau)$
- Model 2 – a vector autoregressive process VAR(1):
  $\mathbf{x_t} = \mathbf{c} + \mathbf{A} \cdot \mathbf{x_{t-1}} + \varepsilon_\mathbf{t}$, where $\varepsilon_\mathbf{t} \sim \mathbf{N}(\mathbf{0}, \mathbf{T})$
  $\mathbf{x_t} = [\ln(MR_t); \ln(GDP^{Rec}{}_t / GDP^{Sen}{}_t)]'$ reflects a hypothesis of a role of income differentials as a migration pull factor
- Model 3 – another VAR(1), $\mathbf{x_t} = [\ln(MR_t); \ln(UR^{Sen}{}_t)]'$ reflects a hypothesis that unemployment in the sending country is an important migration push factor

5. Bayesian forecasts of Polish-German flows

Prior distributions

- Model 1: diffuse $c \sim N(0, 0.001)$, $\alpha \sim N(0, 0.001)$, $\tau \sim \chi^2_{(1)}$
- Models 2 and 3: diffuse $c_1, c_2 \sim N(0, 0.001)$

Elements of coefficient matrix **A**:

$\alpha_{11}, \alpha_{22} \sim N(1, 1)$ – assumption that each of the variables alone is a random walk process

$\alpha_{12} \sim N(0.5, 1)$ – assumption of a positive impact of the lagged migration factor on *MR*

$\alpha_{21} \sim N(0, 100)$ – firm assumption of no inverse impact

T ~ Wishart (2, [[0.1 0.005] [0.005 0.1]])

Bayesian forecast of Polish-German migration

Other remarks

- Sample distribution – normal
- Estimation: numerical simulation using Markov chain Monte Carlo (MCMC), with 10,000 iterations in the burn-in phase and further 100,000 used in the estimation
- Software: WinBUGS 1.4 (Spiegelhalter *et al.*, 2003), code drawing on examples from Congdon (2003)
- Convergence assessment: visual inspection of quantiles
- Goodness-of-fit: sum of squares (SS) for *MR* (DIC has not been used due to different data in Models 1–3)

5. Bayesian forecasts of Polish-German flows

Estimation results

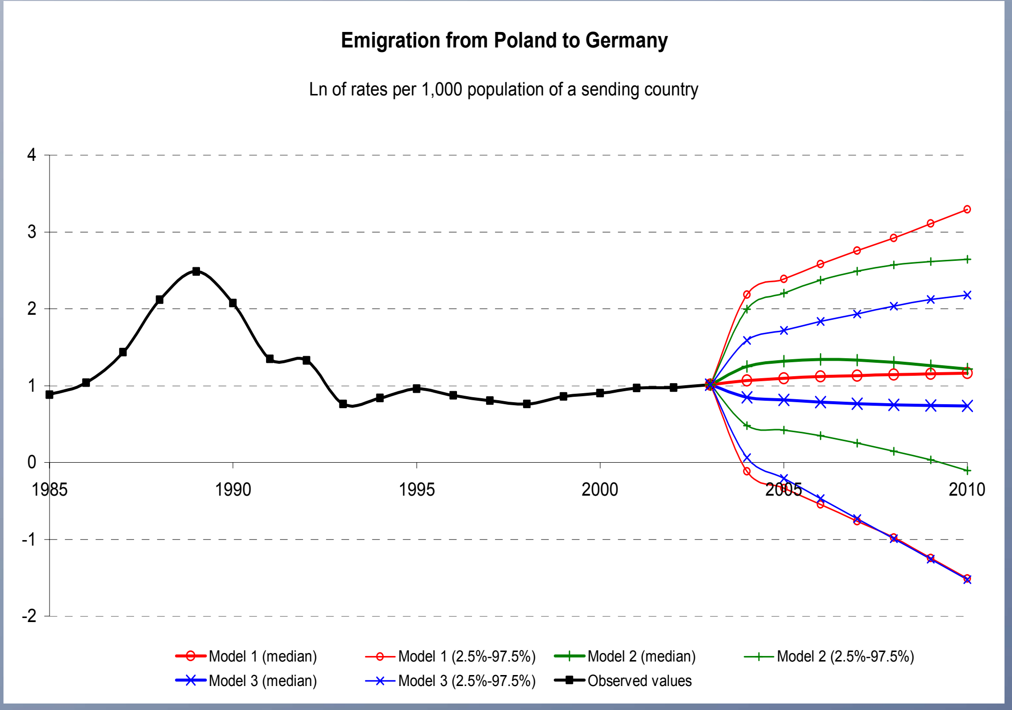
Model	Parameter	Migration from Poland to Germany ln(MR_{P-D}) Mean	St. Dev.	0.025	Median	0.975	Parameter	Migration from Germany to Poland ln(MR_{D-P}) Mean	St. Dev.	0.025	Median	0.975
Model 1	α	0.79	0.20	0.40	0.79	1.18	α	0.80	0.24	0.32	0.80	1.27
	c	0.25	0.25	-0.24	0.25	0.75	c	0.03	0.09	-0.14	0.03	0.20
	τ	6.40	2.18	2.85	6.15	11.31	τ	10.01	3.45	4.45	9.62	17.81
Model 2	$\alpha_{1,1}$	0.76	0.13	0.52	0.76	1.02	$\alpha_{1,1}$	0.85	0.14	0.56	0.85	1.14
	$\alpha_{1,2}$	-0.96	0.35	-1.63	-0.97	-0.23	$\alpha_{1,2}$	0.63	0.27	0.09	0.64	1.15
	$\alpha_{2,1}$	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	0.65	0.88	1.12	$\alpha_{2,2}$	0.82	0.12	0.59	0.82	1.06
	c_1	1.26	0.40	0.43	1.26	2.03	c_1	0.66	0.27	0.11	0.67	1.19
	c_2	0.01	0.13	-0.25	0.01	0.28	c_2	-0.17	0.12	-0.41	-0.17	0.07
	$t_{1,1}$	16.30	5.61	7.22	15.69	29.02	$t_{1,1}$	30.20	10.30	13.59	29.01	53.59
	$t_{2,2}$	156.60	53.35	70.21	150.60	277.10	$t_{2,2}$	156.80	54.17	69.58	150.50	280.30
	$t_{1,2} = t_{2,1}$	-2.64	12.34	-27.61	-2.46	21.39	$t_{1,2} = t_{2,1}$	2.23	16.80	-30.78	2.07	36.07
Model 3	$\alpha_{1,1}$	0.66	0.15	0.37	0.66	0.95	$\alpha_{1,1}$	0.63	0.26	0.13	0.62	1.15
	$\alpha_{1,2}$	-0.04	0.01	-0.07	-0.04	-0.02	$\alpha_{1,2}$	-0.29	0.38	-1.02	-0.30	0.47
	$\alpha_{2,1}$	0.02	0.10	-0.18	0.02	0.21	$\alpha_{2,1}$	-0.08	0.09	-0.25	-0.08	0.09
	$\alpha_{2,2}$	0.83	0.11	0.60	0.83	1.05	$\alpha_{2,2}$	0.76	0.19	0.39	0.75	1.15
	c_1	0.38	0.18	0.02	0.38	0.73	c_1	0.62	0.77	-0.93	0.65	2.11
	c_2	0.53	0.61	-0.66	0.53	1.74	c_2	0.51	0.38	-0.29	0.51	1.24
	$t_{1,1}$	22.05	7.55	9.84	21.20	39.00	$t_{1,1}$	23.53	8.08	10.47	22.59	41.95
	$t_{2,2}$	0.23	0.08	0.10	0.22	0.42	$t_{2,2}$	49.76	17.14	22.37	47.75	88.70
	$t_{1,2} = t_{2,1}$	0.90	0.65	-0.26	0.85	2.31	$t_{1,2} = t_{2,1}$	7.02	8.76	-9.14	6.55	25.54

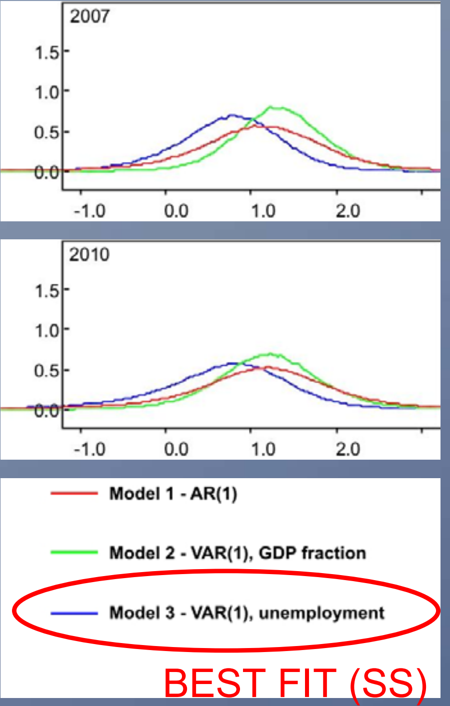
Central European Forum for Migration Research

Środkowoeuropejskie Forum Badań Migracyjnych

5. Bayesian forecasts of Polish-German flows

Results: Poland to Germany



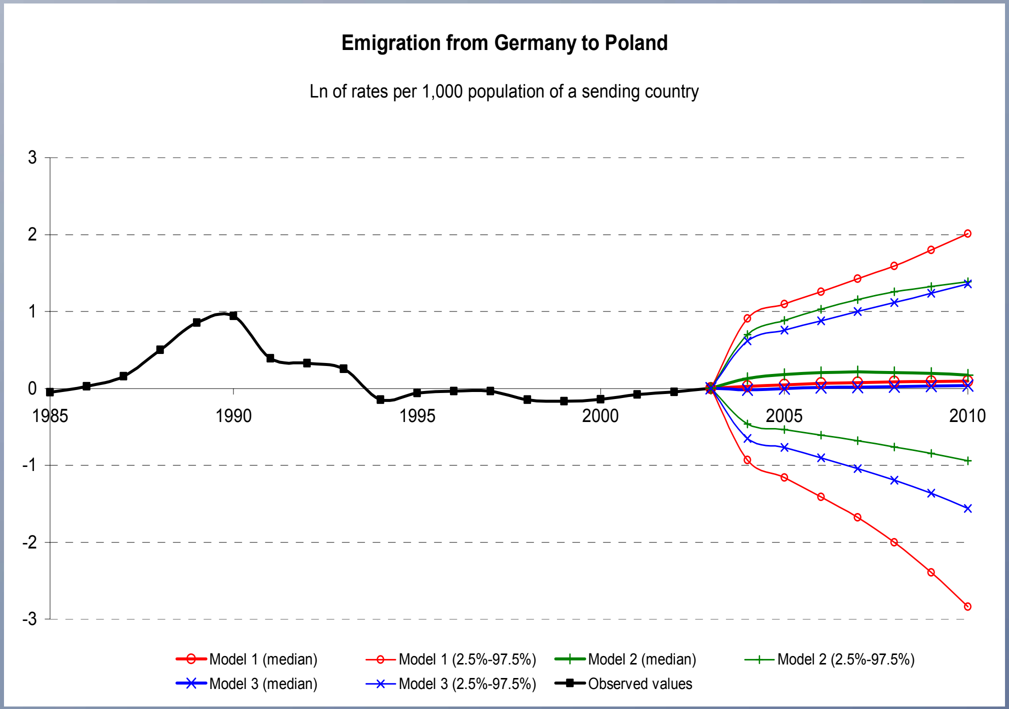


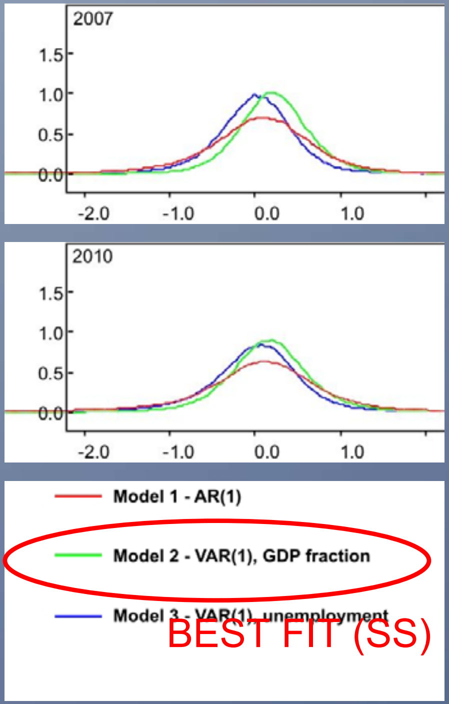
Central European Forum for Migration Research

Środkowoeuropejskie Forum Badań Migracyjnych

5. Bayesian forecasts of Polish-German flows

Results: Germany to Poland





6. Conclusions

- Advantages of Bayesian approach in migration forecasting:
  - Formal combination of various forecasting methods (econometric and time series models), and the expert judgement
  - Inherent analysis of uncertainty: predictive posterior distributions
  - Coherent interpretation of the results (Bayesian intervals), probability not related to the frequency of events
  - With reasonable informative priors, usually smaller errors than in the sampling-theory forecasts (important in the small-sample studies)
- The major disadvantage: computational complexity
  - Solution: numerical methods (MCMC), available in various software
- Further research: formal model selection and averaging, robustness of the results on different prior distributionss

Thank you!