Central European Forum For Migration Research

Środkowoeuropejskie Forum Badań Migracyjnych

CEFMR Working Paper 4/2006

FORECASTING INTERNATIONAL MIGRATION: SELECTED THEORIES, MODELS, AND METHODS

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
4/2006

FORECASTING INTERNATIONAL MIGRATION: SELECTED THEORIES, MODELS, AND METHODS

Jakub Bijak[*]

[*] Central European Forum for Migration Research in Warsaw

Abstract: The paper attempts to outline selected theoretical fundamentals of international migration forecasting. In general, socio-economic predictions can be based either on general laws and theories, or on models designed to suit specific research questions. The discussion offers a brief overview of selected migration theories, as well as a short evaluation of their applicability in forecasting international population flows. Subsequently, a survey of particular models and methods used in migration predictions up to date is presented, with a distinction between deterministic and probabilistic ones. The presented models and methods are finally compared and evaluated from the point of view of their usefulness for the purpose of migration studies.

Keywords: migration forecasting, migration theories, models of migration, deterministic methods, probabilistic methods

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. Marek Kupiszewski and Prof. Jacek Osiewalski for their very helpful comments and discussions.

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 83-60462-03-8

Contents

1. Introduction

Contemporarily, the need for high-quality forecasts of international migration is becoming increasingly significant. Population movements are gaining in importance, given the diminishing impact of natural change on population dynamics. This is especially vital in the developed regions of the world, such as Europe, which are already facing zero or negative natural population growth (e.g., van der Gaag and van Wissen, 1999; Kupiszewski, 2002b). Moreover, the role of migration is not limited to demography, as it also affects many other areas of social life, including economy, labour relations, politics, and culture. On the other hand, in a globalising world, the migratory processes are becoming more and more dynamic and complex (Koryś and Okólski, 2004). In consequence, migration forecasts are associated with a high level of uncertainty and usually bear very high *ex-post* errors (NRC, 2000). Nevertheless, as noted by Dawid (1984: 278), "one of the major purposes of statistical analysis is to make forecasts about the future [and] to offer suitable measures of uncertainty associated with unknown events or quantities". What follows, is a need for developing a migration forecasting methodology that would address the mentioned issues and postulates.

In order to propose a new methodology for migration forecasting, a thorough overview and critical analysis of the existing methods should be made beforehand, especially as there is a wide variety of approaches in the prediction-making practice. Since international migration is a very complex and multi-dimensional phenomenon, its modelling and forecasting involves methods emerging from different disciplines of science: demography, geography, economics, statistics, sociology, political science, or even advanced methods of theoretical physics.

The aim of the current paper is to provide a discussion of various methods, together with an overview of particular theories of international migration and their possible role in setting the forecast assumptions. The rationale of the study is thus to provide background information for the development of an alternative methodology for migration predictions, with focus on the probabilistic approach, which acknowledges the uncertainty issue in a coherent manner. The paper is a continuation and an extension of the one of Bijak (2005), where the issues related to uncertainty in migration predictions are discussed in a relatively more detail.

The structure of the current paper reflects the notion that socio-economic predictions can be based either on general, well-grounded laws and theories, or on various descriptive models designed to suit specific research questions. Therefore, the discussion firstly offers a brief overview of selected migration theories, as well as a short evaluation of their applicability in forecasting international population flows. Subsequently, a survey of particular models and methods used in migration predictions up to date is presented, with a distinction between deterministic and probabilistic ones, depending on the way the uncertainty issue is addressed. The presented models and methods are finally compared and evaluated from the point of view of their usefulness for the purpose of migration studies.

2. Explaining migration: a brief overview of selected theories

2.1. General Remarks

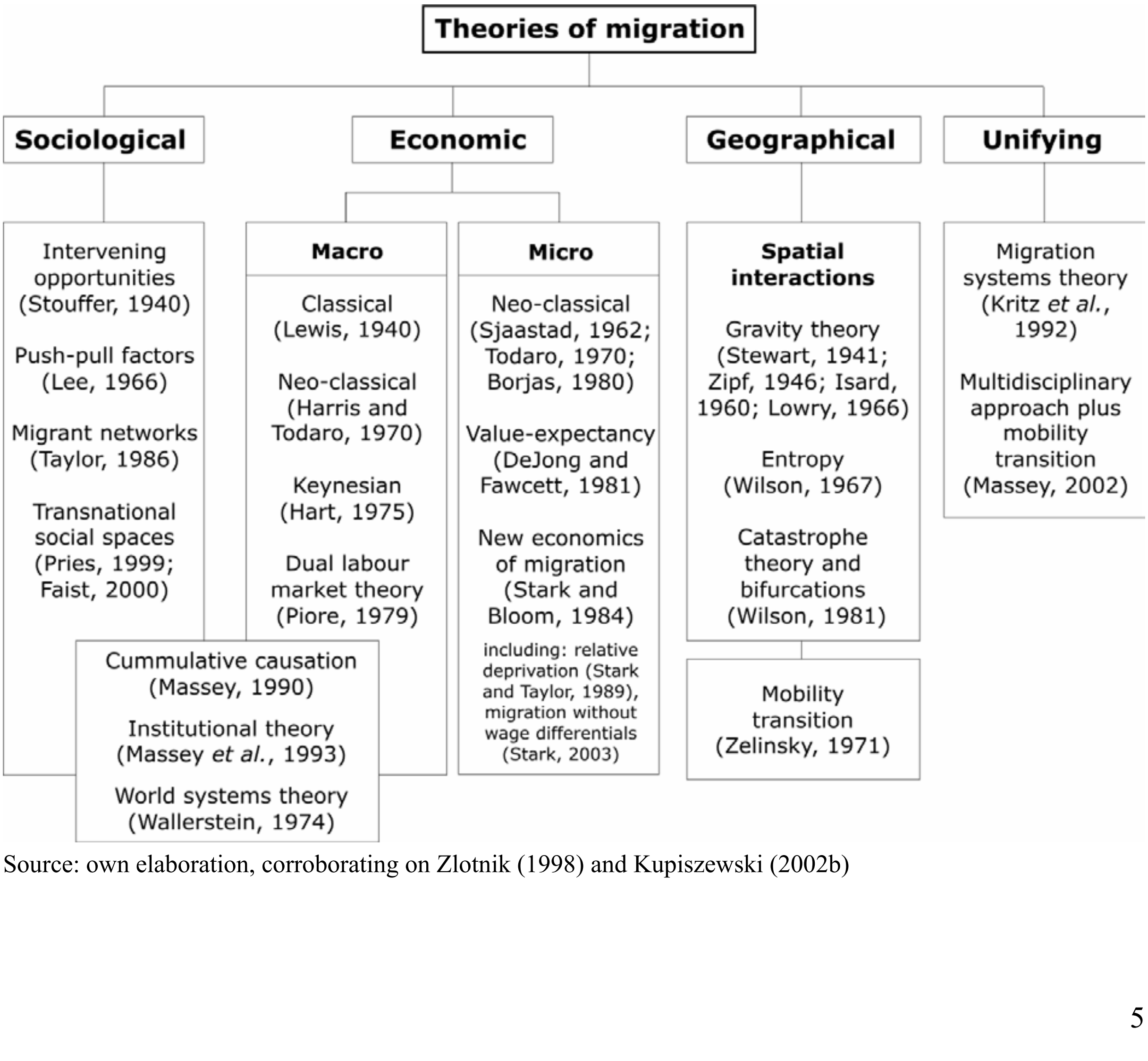
According to Chojnicki (1977), one of the approaches in socio-economic forecasting, referred to as *nomological*, is to derive predictions about the future directly on the basis of theories or laws governing the phenomena under study. However, such theories should be universal and robust enough in order to be straightforwardly concretised in the forecasts. Nevertheless, even if a direct application of the existing theories is not possible, they can provide suggestions for the construction of more specific forecasting models. For these reasons, the current section presents a selective discussion of theories that could be most relevant for that purpose, followed by a brief assessment of their applicability in the macro-level prediction-making in Section 3. The aim is therefore not to offer a comprehensive overview of migration theories, but rather to indicate and evaluate approaches that can be potentially useful for the construction of forecasting models.

A brief insight into the structure of the discussion, reflecting a distinction between sociological, economic, geographical, and unifying (cross-disciplinary) theoretical perspectives of human population flows is shown in Figure 1. It has to be noted that the presented survey of migration theories is by no means complete, and that much more comprehensive discussion is offered in Massey *et al.* (1993), Zlotnik (1998), and Jennissen (2004: 31–57). For example, the distinction between theories explaining the initiation of international migration, and the ones related to the perpetuation of population movements, introduced by Massey *et al.* (1993), is not considered in the current study. Such a division might be potentially useful, if a forecaster would make an attempt to distinguish migration continuing under unchanging external circumstances, from the ones related to system shocks (political or economic), the latter explained by the theories of initiation of population flows.

There is a clear distinction between *internal* and *international* migration. Although the current study focuses on the latter, involving crossing the countries' borders, many theories attempting to explain population movements refer to internal migration. Nevertheless, they are included in the discussion, because contemporarily, given the globalisation processes, the complexity of migratory phenomena increases (Koryś and Okólski, 2004), while the diversity between the mechanisms driving internal and international migration becomes less and less obvious. The Polish case study prepared by Korcelli (1994) suggests that in the present-day world both types of population flows may become more and more substituting, and that prospective migrants would migrate either internally or externally, making decisions on the basis of a cost-benefit analysis. This hypothesis has been recently cautiously confirmed in a study of Polish migration prepared by Kupiszewski (2005). However, in an analysis of 132 European regions from seven countries, prepared by Stillwell *et al.* (1999), very little statistical evidence has been found for the existence of direct linkages between regional international migration and interregional migration.

Generally speaking, in the contemporary Europe the differentiation between internal and international migration seems to be rendered less relevant by the process of European integration with respect to the freedom of movement of persons and labour force, which directly applies to the four countries under study: Germany, Italy, Poland and Switzerland . This notion has been supported by Janicki (2005: 13–17), who also observed that in many border regions within the European Union, international migration flows dominate over the internal ones. This is despite the continuing presence of various socio-cultural barriers, which remain in place even after the administrative ones are gradually being removed. On the other hand, it may be also argued that especially in the situation of Poland and other Central European countries after the EU accession in 2004, the freedom of movement intensifies international migratory processes (see for example Kupiszewska, 2006), rather than the internal ones. This line of reasoning is consistent with the suggestions presented before: given the partial dissolution of institutional barriers, prospective migrants rationally choose their destinations, and reveal their 'true migration preferences', rather than limit themselves to what is available inland.

Figure 1. Selected migration theories offered by various disciplines of science



Source: own elaboration, corroborating on Zlotnik (1998) and Kupiszewski (2002b)

Attempts to synthesize knowledge on spatial population movements in a form of a coherent theory date back to the migration laws of Ravenstein (1885; 1889). He generalised observations on internal migration in Great Britain and found that the intensity of the process was influenced by several factors: distance, population size of the origin and destination regions, absorption capacities of the latter, sex of migrants, etc. Although he found the rural-urban flows to be dominant, he also acknowledged the presence of return migration.

In the 20th century, attempts to assign a theoretical framework to migratory phenomena have been made by the representatives of various disciplines of science, including sociology, economics, and human geography. The systematics of the theories presented in the current section largely follows, with some modifications, the one proposed by Zlotnik (1998), and adopted by Kupiszewski (2002b: 118–122).

2.2. Sociological theories

The sociological theories of migration date back to the *intervening opportunities* concept of Stouffer (1940, 1960). In his approach, the number of migratory events is proportional to the number of attracting opportunities (e.g., jobs) available for migrants at the destination, and inversely proportional to the number of such opportunities existing closer to the place of origin. It is worth stressing that this concept does not relate migration directly to spatial distance (and can not be thus classified as a purely geographical theory), but to the interplay between distance and opportunities available for prospective migrants in various locations.

The notion of opportunities gave grounds to the synthesising *push-and-pull factors* approach of Lee (1966). His theory explains that migration is determined by the presence of attracting (*pull*) factors at destination, and repelling (*push*) factors at origin. For international migration these factors can be further divided into *hard* and *soft* ones (Öberg, 1996). The former group includes dramatic circumstances like humanitarian crises, armed conflicts, environmental catastrophes, etc., while the latter – less critical problems, such as poverty, social exclusion, or unemployment. The dominance of particular factors determines to some extent the characteristics of the migrating population: the favourable pull factors at destination tend to attract migrants who are positively selected in terms of human capital or motivation. This is not the case, when the unfavourable push factors at origin play a crucial role in instigating the migration process.

Taylor (1986) noted that a very important pull factor is the presence of interpersonal *migrant networks* in the receiving country, which are comprised of people sharing kinship, friendship or origin (after: Massey *et al.*, 1993: 448). Such networks of people interconnected by family or acquaintance ties assist subsequent migrants in many aspects of everyday life. The main role of networks is to diminish various costs (not only monetary, but also psychological, and other) and risks associated with migration, and to facilitate the flow of migrants between the origin and destination countries. Population flows are thus characterised by a large degree of inertia: once started, they are difficult to control by the authorities of the receiving country, and become more and more independent from the factors that originally caused them. The network hypothesis has been verified for example for historical migratory outflows from Europe in the period 1850–1914 by Hatton and Williamson (1998), who used foreign population stocks as proxy variables for the size of particular networks.

More recently, the idea of networks has been generalised within the theory of *transnational social spaces* (Pries, 1999; Faist, 2000), providing a meso-level analytical framework for international migration studies[1]. The theory acknowledges the presence of various migration-related cross-border linkages between individuals and groups. As defined by Faist (2000: 199), "transnational social spaces consist of combinations of social and symbolic ties, their contents, positions in networks and organizations, and networks of organizations that can be found in multiple states. These spaces denote dynamic processes, not static notions of ties and positions". Apart from the networks theory, the presented approach builds on the concept of *social capital*, seen as the "resources that help people or groups to achieve their goals in ties and the assets inherent in patterned social and symbolic ties that allow actors to cooperate in networks and organizations, serving as a mechanism to integrate groups and symbolic communities" (*idem*: 102).

The social capital, embodied in social and symbolic ties between individuals and groups, such as the norms of reciprocity, solidarity, and mutual obligations, is a set of local assets. Hence, it may not be easily transferable across borders, what provides justification for the relative immobility of certain social groups, the explanation of which is lacking in the networks theory. On the other hand, once migration starts, the social capital facilitates both the adaptation to the host society, as well as the maintenance of ties with the country of origin. The transnational social spaces that are formed in this way can be thus seen as bridges between the source and destination countries of migrants (Faist, 2000: 195–241). The transnational social spaces theory is sufficiently general to explain the phenomena of chain migration, return migration, perpetuation of migratory processes, and the 'saturation' of population flows at a certain level. However, despite the potential attractiveness of this theory, its development is relatively recent, and it is still very far from a possible operationalisation for a practical use in migration forecasting.

The important role of institutions in facilitating the migration process is acknowledged in the *institutional theory* of migration sketched by Massey *et al.* (1993), concerning both regular (e.g., active job recruitment, counselling), as well as irregular (e.g., human smuggling and trafficking) aspects of population flows. This notion is complementary to the network theory, extending the set of subjects facilitating migration to various institutions, from for-profit entities and enterprises (legal or illegal), to humanitarian organisations, NGOs, etc. Also in

[1] I am very grateful to Iza Koryś for highlighting the recent paradigm shift from 'networks' to 'transnational social spaces' in an attempt to create a comprehensive sociological migration theory.

the light of this theory, the process of institutionalisation of migratory flows seems to be to some extent self-perpetuating, independent from the initially-dominant migration factors, and increasingly difficult to regulate (Massey *et al.*, 1993: 451). With respect to the acknowledgement of the importance of institutions, these observation can be also seen as linked with the institutional theory in economics, thus *de facto* forming a hybrid of sociological and economic perspectives.

Another example of a cross-cutting, interdisciplinary theoretical approach is the notion of *cumulative causation* (Massey, 1990), corroborating on the economic theories of Veblen (1898) and Myrdal (1957). The theoretical construction is based on the proposition that migration is an evolutionary process that contributes to institutional and socio-economic change both at origin and destination, through various feedback mechanisms. Examples of the latter include the redistribution of income of households involved in migration, both in absolute and relative terms, as well as the redistribution of land and capital. These processes ultimately result in changes in the social hierarchy (Massey *et al.*, 1993: 451–454). As one of the consequences, due to the fact that return migrants usually possess more human capital than the immobile population, the non-migrant group increasingly aspires to a privileged position in the community.

The factors listed above instigate a ‘migrant culture’, in which migration is perceived as an activity with expected positive gains, what in turn contributes to changes in the human capital distribution of the source regions. In such a way, the ‘brain drain’ of a sending country can actually turn into the ‘brain gain’. This theoretical possibility has been further corroborated by Stark and Wang (2001) within the microeconomic analytical framework. Their analysis provides arguments that expected positive returns from migration are a source of externalities in the form of an overall increase in the human capital in the region of origin. It is also worth noting that the principle of cumulative causation is an important element of the transnational social spaces theory (Faist, 2000: 129–132).

2.3. Macroeconomic theories

The economic theories of labour migration, as all economic theories in general, follow either the macro- or micro-level perspective, and reflect a diversity of paradigms of theoretical economics that either dominated, or co-existed in various periods. To begin with, the neoclassical macroeconomic migration theory (Lewis, 1954; after: Massey *et al.*, 1993: 433) explains that given wage differentials between two capitalist economies, one characterised by a surplus of labour (unemployment) and the other by a surplus of capital, migration and capital movements occur. The flows of both production factors in opposite directions, and the related convergence of wage levels, are a way of return to the economic equilibrium, where the wage differential reflects only the costs of undertaking migration. It is worth noting that in the whole neoclassical approach, migration is an disequilibrium phenomenon, which ceases as soon as the equilibrium is reached (cf. Harris and Todaro, 1970: 129).

Nevertheless, the presented approach is very simplistic and does not explain for example the phenomenon of return migration, or population flows in the absence of wage differentials, as indicated by Stark (2003). Additionally, Jennissen (2004: 46) points out to the fact that there exists an alternative Keynesian view on migration-induced labour market adjustments towards the economic equilibrium, through the elimination of differences in unemployment, not in wages (cf. Hart, 1975). What follows is the pro-cyclical nature of migration, as observed for example for Canada by Milne (1993), who showed that net migration rates of particular regions go along with the business cycle, although with a time lag.

A similar framework is explored in the neoclassical macroeconomic migration theory of Harris and Todaro (1970). They formulated their conceptual model in terms of *expected* income (wages), taking into account the probability that labour migrants from the 'agricultural' sector (A) in the 'rural' region r, find jobs in the 'manufacturing' sector (M) of the 'urban' economy (u), the latter characterised by minimum wages (W^*_M) and unemployment. The equilibrium condition, to which the system should optimally converge, is defined as (Harris and Todaro, 1970: 129):

(1) $$\frac{dN_u}{dt} = \psi\left(\frac{W_M^* \cdot N_M}{N_u} - P \cdot \frac{dX_A}{dN_A}\right),$$

where N_u is the total urban labour supply (original population and migrants), N_M is the number of actually employed in the manufacturing sector, and N_A – in the agricultural sector. Further, $X_A=X_A(N_A)$ is the monotonously increasing and concave production function of the agricultural sector, and P denotes the price of the agricultural output defined in terms of the goods produced by the manufacturing sector (the terms of trade). The expected wages in the urban region are equal to the minimum wages adjusted for the chances of being employed, $W^*_M{\cdot}N_M/N_u$. The function $\psi(x)$ is such that ψ '>0 and $\psi(\alpha)$=0 for some real number α, which under (1) implies that migration stops, when the expected wage differential equals α (in particular, $\alpha = 0$ can be assumed).

Harris and Todaro additionally tested the impact of various policy instruments (among others, minimum wages and migration restrictions) on the economic equilibrium. They found that in the absence of full wage flexibility "either a limited wage-subsidy or a migration-restriction policy will lead to a welfare improvement [at the destination]" (*idem*: 137). On the other hand, the outcome of both of these policies will be sub-optimal from the point of view of the economy as a whole system of the sending and receiving regions.

Apart from the purely neoclassical and Keynesian approaches, there have been many further attempts to explore the economic motives of migrants and the nature of migratory processes.

The *dual labour markets theory* (Piore, 1979) is based on an observation that migration flows are to a large extent determined by labour demand characteristics at the destination. This notion puts forward that immigrant labour is necessary for the economies of the developed countries. An important explanation is that wages are not only the price of labour, but also a proxy measure of the employee's position in the occupational and social hierarchy. If there are labour shortages at the bottom of the hierarchy, the entrepreneurs would prefer to hire immigrant workers without aspirations to a higher social status, than to raise wages in order to attract local labour force. The latter option would require proportional wage increases in the whole sector to preserve the whole job ladder, and would ultimately lead to a structural inflation.

The dual labour markets theory justifies the segmentation of labour markets – the local population moves to more attractive professions, while immigrants take up the '3D' (dirty, dangerous and difficult) jobs. On the top of that, this division is intensified by the very nature of various occupations – the attractive ones are usually capital-intensive, while the 3D – labour-intensive. Moreover, immigrant workforce is a more flexible production factor than the local one, protected by various institutions (trade unions, regulations of work conditions, etc.), which is another factor petrifying the labour market dualism. This also gives reasons for the dominance of women and young people among migrants – they are more willing to accept poorly-paid jobs in unfavourable conditions than other groups (Massey *et al.*, 1993: 441–443). Thus, the dual labour markets theory provides useful explanations for many features of migratory phenomena, but nevertheless seems difficult to operationalise to suit macro-level migration predictions.

In an attempt to generalise the macroeconomic perspective, the *world systems theory* (Wallerstein, 1974) assumes that international migration is associated with the advances of the capitalist system and global markets, not only in the world's economic 'core', but also in semi-peripheral and peripheral regions. A flow of goods and capital from the core to the peripheries "in search of land, raw materials, labour, and new consumer markets" (Massey *et al.*, 1993: 445) is counterbalanced by the flow of labour in the opposite direction. There are many driving forces behind this process. On one hand, there is an increasing demand for low-skilled labour in the core regions, where the jobs in the manufacturing sector become less and less desirable for the local population, following the shift towards the service-based economy. On the other hand, the commercialisation of agricultural production in the peripheries caused by the capital flows leads to an increased productivity and the resulting reduction of demand for the local labour. These factors instigate migration of the surplus of agricultural workforce, which is 'uprooted' from the peripheral regions in the search of either formal or informal job opportunities in the low-paid segments of the manufacturing and services sectors in the core (Zlotnik, 1998: 7–8).

The world systems theory also acknowledges the presence of many other links (not only material, but also historical, cultural, linguistic, etc.) between the origin and destination countries, which influence migratory flows. These elements are the reason, why the whole theory cannot be simply classified as 'macroeconomic', but rather offers another cross-cutting perspective, combining economic and sociological explanations. As noted by Massey *et al.* (1993: 448), in the world systems approach, "international migration ultimately has little to do with wage rates or employment differentials between countries; it follows from the dynamics of market creation and the structure of global economy". Special attention is paid to the asymmetrical relationships between the former colonies and their former metropolis, the latter considered to have more favourable terms of trade. In that respect, Jennissen (2004: 53–54) noted that although this theory gives grounds to the philosophy of alter-globalisation movement, its view on international trade can be seen as controversial, as free trade is contemporarily thought to reduce income and employment disparities (and thus also migration). Nonetheless, the world systems theory is articulated only verbally, not in formal mathematical terms, and is too general in order to serve as a direct reference for migration forecasting.

2.4. Microeconomic theories

The individual-level migration decisions are explained for example by the neoclassical microeconomic theory of Sjaastad (1962), which treats migration as an investment in human capital, and a result of a rational cost-benefit analysis. According to this theory, prospective migrants choose the destinations that are maximising the net present value of their expected future income, less various direct and indirect costs of migration. In a more comprehensive framework, this can be formalised as (Massey *et al.*, 1993: 435, referring to Borjas, 1990):

(2) $$ER(0) = \int_0^n \left[p_1(t) p_2(t) Y_d(t) - p_3(t) Y_o(t) \right] \cdot e^{-rt} dt - C(0).$$

In the above equation, $ER(0)$ denotes the expected returns from migration at the moment 0, n is the time horizon of the decision-making process, p_1 is the probability of not being deported (p_1<1 for irregular migrants), Y_o and Y_d are earnings at the origin and destination, while p_2 and p_3 denote the respective probabilities of finding a job. Further, r is the discount rate, and $C(0)$ is the sum of all costs of migration, both economic and psychological (*idem*), although the latter seem to be very difficult to operationalise and measure.

A generalisation of the micro-economic decision framework is the *value-expectancy* concept of DeJong and Fawcett (1981; presented after: Faist, 2000: 36–37). The underlying formal model defines the individual motivation to migrate (*MM*), being subject to maximisation, as:

(3) $$MM = \sum_i P_i \cdot E_i \ .$$

The values of P_i refer to the preferred outcomes of migration, and E_i – to the 'expectations' of their realisation through migration which are hold by a potential migrant, and thus can be interpreted as subjective probabilities (*sic!*). The index i denotes the 'values' or 'desires' of an individual, that is, various dimensions of the decision-making problem. The presented theory is very comprehensive and may cover different aspects of human decisions in the migration context (different values of *i*), not only limited to the economic ones (unemployment, wages), but also potentially taking into account the social and psychological spheres of life. On the other hand, Faist (*idem*) points out to the fact that the value-expectancy theory would be difficult to apply for prediction-making, as people tend to rationalise their actions *ex post* rather than to reason *ex ante*, as it is suggested by the formulation of (3). Moreover, many of the dimensions of the migratory decision problem may be, again, hardly possible to measure.

The *new economic theory of migration* (Stark and Bloom, 1985; Stark, 1991) suggests that migration-related decisions are made by households rather than by individuals. This finding coincides with the observations that migratory processes are characterised by visible family patterns, as noted both by economists (Mincer, 1978), and demographers (Castro and Rogers, 1983). From this point of view, diversified migration strategies of particular household members are an instrument of risk management at the household level, rather than a simplistic maximisation of the expected income. In this approach, wage differentials between origin at destination countries are not a prerequisite for migration. One of the possible extensions of this perspective is the life-cycle theoretical model of Dustmann (1997), who formulated a stochastic framework for the analysis of migration and return migration under the condition of uncertainty, with focus on precautionary savings of the individuals. Savings can be seen as another element of risk management at the household level, related to remittances – monetary transfers from migrants to their families in the country of origin.

Other theoretical examples explaining the presence of migratory flows in spite of the absence of differences in income levels have been summarised by Stark (2003). He stresses an important role of individual preferences (‘taste’) in favour of migration, as well as of the difference in the purchasing power of savings generated by migrants between the regions of origin and destination, which justifies the presence of return migration. Another very important migration factor is the community context, which can be associated with the *relative deprivation* concept (Stark and Taylor, 1989). This approach proposes that migration is fuelled not by absolute, but by relative income differentials towards the reference group of potential migrants. This notion provides a link between economy and psychology, the latter with respect to the sense of resentment towards the reference group. The relative deprivation *RD* of a household with income *y* is defined as (Stark and Taylor, 1989; after: Massey *et al.*, 1993: 439):

(4) $$RD(y) = \int_{y}^{y_{max}} h[1 - F(z)]dz,$$

where $F(y)$ denotes the cumulated income distribution in a given community, y_{max} is the maximum income observed in this group, and $h(\cdot)$ is a monotonously increasing function, measuring the dissatisfaction with respect to the relative position of the household’s income in the reference community.

Recently, Quinn (2006) extended the definition of relative deprivation to other durables, land and housing, in addition to income. In his empirical analysis of migration within Mexico, as well as from Mexico to the United States, the data supported a hypothesis about a combined effect of absolute wage differentials and relative deprivation on migration flows. Such a synthesising approach may provide a partial solution to the problem noticed by Massey *et al.* (1993: 440), that the neoclassical and new economic theories of migration "lead to divergent conclusions about the origins and nature of international migration".

Further attempts to reconcile the two paradigms (neoclassical and 'new economics') may potentially be aimed at combining the two levels of decision-making (individual or household), their distinct goal functions (maximising income or minimising risk), or presence of the social context (e.g., whether income is perceived in absolute terms, or in relation to the income distribution in a reference group). Such a comprehensive micro-level model explaining migratory decisions would have to be based on the multi-criteria programming with many parameters and constraints. Problems of this type, although potentially very complex, may either have relatively simple analytical solutions, or be solved using numerical methods, and thus be able to provide useful input for practical forecasting applications. Nevertheless, the issue of constructing a multi-level, multi-criteria micro-model of migration remains far beyond the scope of a current study.

2.5. Geographical theories

In the tradition of human geography, the theories of migration are focused on the role of distance in explaining spatial movements. Distance is viewed as a factor moderating the *spatial interactions* between regions, which include population flows. For example, the *gravity theory* of migration (Stewart, 1941; Zipf, 1946; Isard, 1960/1965), analogous to Newton's law of gravity, assume that migration between regions i and j, $m_{i,j}$ is proportional to the product of population sizes in the origin and destination regions (P_i and P_j), corresponding to masses in the Newtonian model, and inversely proportional to the b-th power of distance between the two regions, d_{ij}, which is a discounting factor (Isard, 1960/1965: 350):

(5) $$m_{i,j} = G \cdot \frac{P_i \cdot P_j}{d_{ij}^b}.$$

There have been various hypotheses concerning the values of the coefficient b that would optimally model the numbers of spatial interactions. The Stewart's (1941) proposition, assuming b=2, defines (5) as the 'demographic force' in a full analogy to Newton's law, while the hypothesis of Zipf (1946) is based on the assumption of b=1 and on a logarithmic transformation of the right-hand side of (5) (after: Isard, 1960/1965: 352–356).

Isard (*idem*: 357–358) noted that in empirical research on spatial interactions, the notions of mass and distance can be defined in many different ways. Instead of population sizes, such economic measures as employment or income can be used as masses, while distance can be measured according to a different metric: either Euclidean (crow-fly), or taking into account the structure of the existing transport network, time, or cost of transportation. Various mass factors can be also considered jointly, like for example in the model of Lowry (1966), who built a gravity model relating migration to unemployment rates, wage levels, numbers of persons in civilian labour force (non-agriculture), as well as in the armed forces, both at the origin and at the destination (after: Morrison, 1973: 132–133). Alecke at al. (2001) observed that the notion of gravity is also used in many econometric models, where income (GDP) differentials *per capita* are most commonly applies as masses, instead of population sizes.

In addition to the gravity framework, there have been several attempts to utilise more advanced mathematical tools in order to find the patterns of spatial interactions. Wilson (1967, 1970; after: Mazurkiewicz, 1986: 25–34) proposed that interactions between regions i and j, including migration ($m_{i,j}$), maximise the *entropy* of the whole system of regions under study:

(6) $$\Sigma_{ij}\, m_{i,j} \ln(m_{i,j}) \rightarrow \max.$$

The maximisation is constrained by the costs of such interactions ($x_{i,j}$), which are related to covering the distance between i and j ($\Sigma_{ij}\, m_{i,j}\, x_{i,j} = c$).

In addition to the entropy concept, the same author also undertook an effort to describe spatial interactions within the framework set by the catastrophe and bifurcation theory (Wilson, 1981), where the dynamic systems under study may undergo substantial qualitative changes as a consequence of very small modifications of some of their parameters.

Another geographical theory, the one of the *mobility transition* (Zelinsky, 1971), attempts to explain changes in spatial mobility by a hypothesis akin to the 'demographic transition' concept. According to Zelinsky (*idem*; after: Gawryszewski, 1989: 11–19), social modernisation caused an increase and a continuous diversification of human mobility patterns. During a transition from a pre-modern to a modern society, in the 18th and 19th centuries, migratory movements were primarily undertaken towards the national borders, as well as to other countries. Along with the industrialisation processes, migration from rural to urban areas was on the rise – it began to decline only in the advanced societies in the second half of the 20th century. In these advanced societies, in turn, migration between and within urban areas have been increasing, as have the short-term circulatory movements (commuting, business trips, tourism, etc.). Recently, circulation have been absorbing more and more mobility of other types, and has itself been substituted by an increasing role of communication systems.

The mobility transition theory has been designed as a comprehensive framework describing human mobility, although Kupiszewski (2002b: 122) noted that it ignores the phenomena of suburbanisation and counter-urbanisation, characteristic for the advanced societies.

Nevertheless, the hypothesis of substitution between spatial mobility and other means of communication seems increasingly important in the contemporary world of Internet and telecommuting. The whole theory, however, would be very difficult to apply directly to migration forecasting, given the limited availability of data on various types of mobility and communication.

2.6. Unifying perspectives

Apart from the discipline-specific theories of migration, there have been also several attempts to propose a unified explanation for population flows. The *migration systems theory* (Kritz *et al.*, 1992; following the pioneering work by Mabogunje, 1970; after: Zlotnik, 1998: 12–13) distinguishes migration systems comprised of various sending and receiving countries characterised by similar migratory patterns. In such a dynamic system, migration is in a continuous interplay with historical, economic, cultural and political linkages between the countries, both on the micro and macro levels. The presence of feedback effects makes population flows both a cause, as well as an outcome of the other interactions. Despite the clear advantages of such a synthesising and multi-perspective approach, it is at the moment too complex to be applied in practice, especially given the problems with availability and quality of internationally-comparable migration statistics (Zlotnik, 1998).

Another attempt to create of a synthesising theoretical framework of international migration has been recently undertaken by Massey (2002). His proposition combines economic, political, sociological and psychological determinants with the notion of migration transition (similar to the one proposed by Zelinsky, 1971), and with the role of duration-of-stay effects. In general, Massey (2002) perceives international migration in the post-industrial countries as an outcome of socio-economic development and integration processes. Nonetheless, at the current stage of development, these ideas are far from constituting an all-inclusive theory of migration, and seem hardly possible to be operationalised in practical applications.

3. From migration theories to model-based forecasting of population flows

3.1. Use of theories in migration predictions

With respect to the evaluation of migration theories, Öberg and Wils (1992) observed that each of the existing ones explains the actual phenomena only partially and therefore has a limited use in the forecasting process. They stressed that the geographical theories are more suitable for internal migration, as they do not include such elements as institutional barriers (state borders, visa requirements, etc.), which are inherent in the case of international flows (cf. Zolberg, 1989). Also Willekens (1994) noted that the existing theories of international migration do not substantially differ from those of internal migration, although contemporarily this may be less relevant due to the reasons discussed before, such as globalisation and integration processes in Europe. In addition, the existing theories ignore forced migration and migration policy factors, which are crucial for the actual magnitudes of observed population flows. Also the economic theories do not adequately explain migration during system shocks, like for example the socio-economic transformation in the post-socialist Europe in the late 1980s and early 1990s.

According to Öberg and Wils (1992: 6–7), although all theories – geographical, economic and sociological – are useful in explaining migration *ex post*, their forecasting potential is very limited. These objections are shared by Kupiszewski (2002b: 122–124), who argues that the existing theories of migration can be hardly used in a direct manner to forecast international population flows. None of the theories is comprehensive and self-contained, while migration is a too complex phenomenon to be explained by a single, narrow theory. In contrast, the wide-ranging theories, as for example, the mobility transition theory of Zelinsky (1971), the world systems theory of Wallerstein (1974), the migration systems theory of Kritz *et al.* (1992), as well as the unifying perspective of Massey (2002) are difficult to operationalise, because they are not sufficiently formal in terms of the mathematical expressions applied. The other theories, although potentially transformable into the forecast input through a range of proxy variables, listed for example by Jennissen (2004: 57), can have a limited explanatory capacity. Due to incompleteness and various deficiencies of macro-level statistics on migration in Europe (cf. Kupiszewska and Nowok, 2005; Nowok *et al.*, 2006), the reliability of any empirical research aimed at verifying particular theories would be only partial and could be easily questioned.

For the reasons mentioned above, the nomological forecasting of migration, based directly on the laws or theories of population movements, is not an option, as the existing laws and theories are not universal enough to allow for the practical application of this approach. Therefore, migration forecasting should follow a model-based approach, rather than be based on any particular theory. An exception may be the inclusion of selected push and pull factors as explanatory variables in forecasting models, which would partially refer to the theory of Lee (1966), to the extent it is made possible by the data availability.

3.2. Migration forecasting methods and models: state of the art and typology

The nomological, theory-based, predictions of socio-economic phenomena are contrasted by Chojnicki (1977) with the *heuristic-model* approach, contemporarily prevailing in the forecasting practice. In order to reconstruct the events under study, and analyse their consequences, various descriptive *models*[2] can be built. Unlike theories, which should be by nature general and well-founded, models are hypothetical, relate to a particular reality, and are characterised by a large degree of flexibility, although Chojnicki (*idem*) noted that in some cases it may be difficult to distinguish between theories and models. In addition, there exist heuristic methods, which are not model-based, yet can be applied in forecasting (e.g., surveys, or the Delphi approach).

One group of methods used in preparing forecasts of international migration is based on deterministic mathematical models, or other approaches and techniques, which do not explicitly address the uncertainty of migratory phenomena[3]. It should be noted that the crucial distinction between deterministic and probabilistic methods or models, applied in the current study, is to some extent arbitrary, and can be related to the common practice in the existing research, rather than to the potential of methods in addressing the uncertainty issues.

Unlike the deterministic methods and models, the stochastic (probabilistic) tools of migration forecasting are rooted in the probability theory, and explicitly address, at least potentially, the issue of uncertainty[4]. With respect to the typology of forecasting methods presented by Chojnicki (1977), they are usually model-based, a few of them including additional heuristic elements (e.g., qualitative research in the ethnosurvey). Although some of the models make explicit reference to particular theories of migration, like for example econometric models to economic theories, they do not attempt to construct a general theoretical framework of population flows in probabilistic terms. It should be remarked that the stochastic models discussed in this paper are presented in a relatively more detail than the deterministic ones, in order to present possibilities of addressing the uncertainty issue in migration forecasting.

---

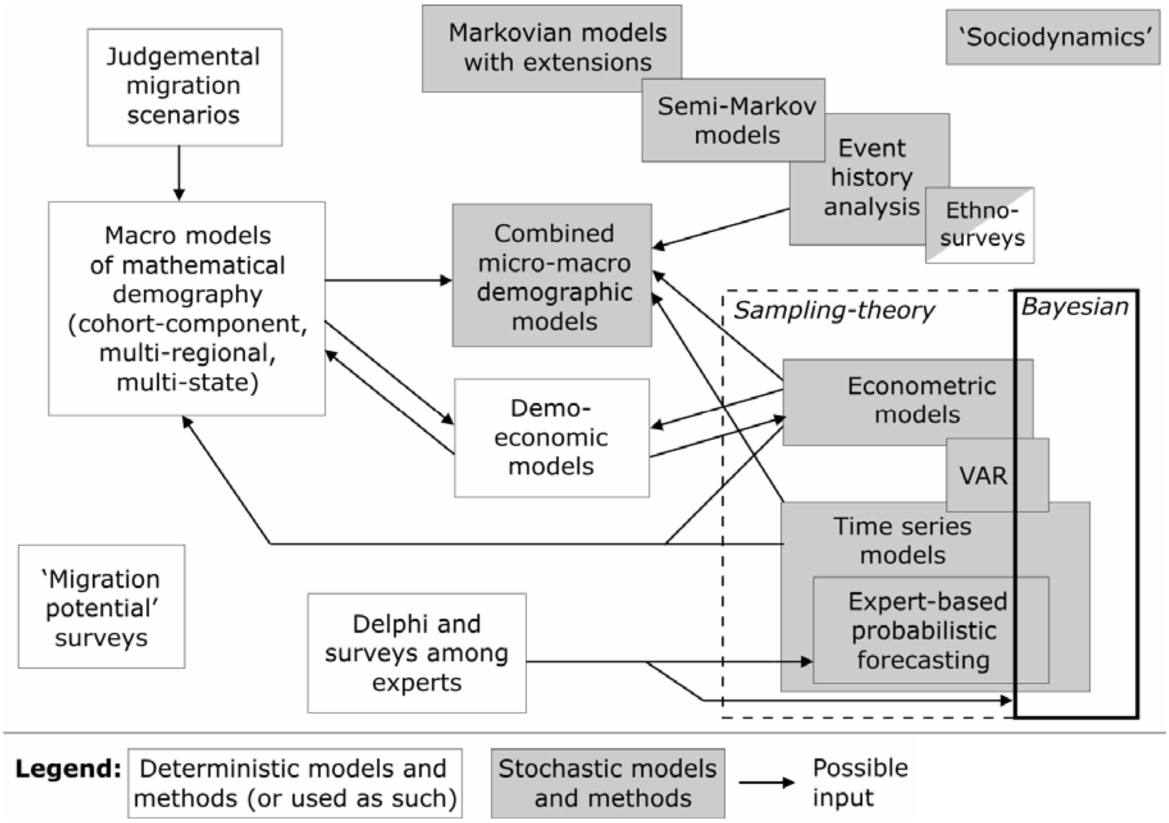
[2] 'Model' can be defined as "a system of postulates, data, and inferences presented as a mathematical description of an entity or state of affairs" (the Merriam-Webster Online Dictionary; «www.m-w.com», accessed on 25 April 2006).

[3] In the current study, 'deterministic' is thus understood as 'not allowing for randomness or uncertainty', in contrast to the dictionary-based definition, describing 'determinism' as "a theory or doctrine that […] social phenomena are causally determined by preceding events or natural laws" (idem). Notably, determinism in the latter interpretation may refer to stochastic explanations of the phenomena under study, provided that the 'natural laws' involved contain an element of randomness, as, for example, in many areas of contemporary theoretical physics. A discussion on uncertainty and (in-)determinism in migration forecasting is offered in Bijak (2005).

[4] The Merriam-Webster Online Dictionary explains the term 'stochastic' as "(1) involving a random variable; (2) involving chance or probability", both definitions being equally useful for the purpose of the present study («www.m-w.com», accessed on 25 April 2006).

As mentioned before, the theories of migration are too fragmented and too rough to be used directly as a forecasting tool within the nomological approach. Therefore, the only option to obtain plausible predictions of population flows is the use of models or other heuristic methods. There are several such methods and models, both deterministic and probabilistic, which can be potentially interlinked, with the outcome of simpler methods used as input in the more general ones. An attempt to summarise the discussion about various models and methods, and the possible relationships between them, is presented in Figure 2.

Figure 2. Typology of selected migration forecasting methods and models



Source: own elaboration

In general, the distinction between deterministic and stochastic models follows the way they are presented in Sections 4 and 5, with some remarks that need to be made. On one hand, as it has been mentioned before, some of the deterministic macro-level demographic and demo-economic models can acquire stochastic features, when they get probabilistic input. On the other hand, stochastic forecasts can (and often are, as it is discussed further) be interpreted in a deterministic fashion, disregarding the whole context of uncertainty and forecast errors. Moreover, some methods use research techniques of different scientific disciplines in parallel. An example is the ethnosurvey, the ultimate results of which rely both on the probabilistic (survey-based) and deterministic (qualitative) analysis of migration flows. Furthermore, the econometric and time series models can be prepared alternatively within the sampling-theory or the Bayesian statistical paradigm.

4. Deterministic methods and models used in migration predictions

4.1. Judgemental migration scenarios

Judgemental scenarios used in demographic forecasting describe possible future trajectories of particular components of population change (fertility, mortality, and migration). They are constructed on the basis of qualitative and quantitative argumentation about what is considered to be the plausible development of the variables in question. The scenarios have to be coherent with the underlying judgement and assumptions ('the story') behind them, and show the demographic consequences of the latter using a 'what-if' approach. Scenarios usually serve as input for deterministic population projections, showing the outcome of various variants of change in the demographic parameters: usually, the baseline, high, and low ones.

Probably the first judgemental scenario of expected future migration flows set up in terms of quantitative demography can be found in the 'political arithmetick' (*sic!*) study of Sir William Petty (1682) concerning the future growth of the City of London. Although internal migration from rural England to London is hidden in different assumptions regarding the future demographic growth for both areas, population inflows to the Britain's capital are explicitly mentioned as the causes of the city's expansion.

According to the knowledge of the author[5], deterministic scenarios of international migration are contemporarily used in the official population projections in a vast majority of developed countries, at least in Europe. Apart from the national forecasters, judgemental scenarios of population flows are also widely applied in the supra-national demographic projections prepared by international organisations and agencies (for example, United Nations, 2005; Eurostat, 2005[6]), as well as by individual researchers (e.g., Bijak *et al.*, 2006).

With respect to the construction of scenarios, Kupiszewski (1998) pointed out that in some cases they can take into account analogies of the projected migration flows to the ones that have already occurred in the past in other countries, under similar conditions. An example is the enlargement of the European Union in 2004. An analysis of similarities to the previous extensions of the then-EEC (to include Ireland in 1973, Greece in 1981, Portugal and Spain in 1986) would allow for assuming a reasonable scenario, in which the post-accession increase of migration from Central to Western Europe will likely be temporary and rather moderate in size. The analysis may be performed within the formal framework of econometric modelling.

---

[5] The mentioned issues have been discussed for example during the "Meeting of the Working Group on Population Projections" in Eurostat (Luxembourg, 15–16 July 2004), and the "Joint Eurostat – UN ECE Work Session on Demographic Projections" (Vienna, 21–23 September 2005), in both of which the author participated.

[6] In Eurostat (2005), the methodology differed for the 'old' 15 EU member states and for the 12 accession and candidate countries as of 2004. For the latter, the projections were purely based on judgemental scenarios, while for the former they involved averaging of forecasts yielded using three methods: extrapolation of trends, econometric analysis of migration determinants, as well as national forecasts (Lanzieri, 2004).

An example of a scenario-based study dealing with the forecast of the East-West migration in Europe after the EU enlargement is the research of Layard *et al.* (1992), who extrapolated earlier empirical findings concerning migration from Southern to Northern Europe in the 1950s and 1960s, as well as from Mexico to the United States in the 1970s and 1980s. Their estimate of a gross inflow of Central and Eastern Europeans to the West over 15 years totalled less than 3 percent of the population of the sending countries (Layard *et al.*, 1992; after: Alvarez-Plata *et al.*, 2003: 11).

A similarly straightforward calculation has been made by Franzmeyer and Brücker (1997), who built a gravity model of net migration between ten Central and Eastern European countries, i, and the 'old' European Union, EU-15. The model was based exclusively on the difference of the GDP *per capita* (PPP-adjusted), $Y_i/Y_{\text{EU-15}}$, and it was calibrated on the basis of the earlier empirical findings of Barro and Sala-i-Martin (1995). The key assumption was made on the elasticity of migration on income differentials, so that an income gap of 10 percent was presumed to drive between 0.08 and 0.16 percent of the population of the worse-off country out of the place of origin. This supposition is the main reason for extremely high migration forecasts yielded by the model: population flows from Central and Eastern Europe to the EU-15 have been estimated as 590–1,180 thousand persons a year, depending on the pace of income convergence.

The example of the study of Franzmeyer and Brücker (1997) shows that the judgemental scenarios should not only be consistent with the underlying 'stories', but also controlled with respect to the results they produce. Neither reasonable assumptions, nor the use of a particular theory (here: gravity with respect to income differentials) alone do not guarantee that the outcome will be plausible from a demographic point of view[7].

4.2. The Delphi method and surveys among experts

Expert judgement used in setting the scenarios can be either made by the authors of forecasts themselves, or derived from a survey carried out among a larger group of specialist from various countries and fields of expertise. The latter approach may include obtaining migration scenarios using a Delphi method, designed to "elicit and refine group judgement" in an anonymous and interactive communication process among experts in the field (Dalkey, 1969: v). The exchange of knowledge is made in subsequent rounds, between which all feedbacks from the participants are controlled, and the final output is formed from the aggregation of all individual opinions. A migration forecasting example is the study of Drbohlav (1996), who used a two-round Delphi method on a sample of 70 experts in the first round and 39 in the second one, and obtained rough estimates of the magnitude and timing of the expected East-West migration flows in Europe, as well as the envisaged directions of migration policy changes in the future. Another migration-related Delphi research has been conducted in 1991

[7] This issue is also discussed in Section 6.

in Russia by V. Tichonov. The study produced an estimate of emigration from the former USSR in the period 1992–1997, whereby about half of 30 experts envisaged between 2 and 4 million emigrants (Vishnevsky and Zayonchkovskaya, 1994; after: Willekens, 1994: 28)[8].

Also Bauer and Zimmermann (1999) made an attempt to gather information on future migration flows from Central to Western Europe from a survey conducted among experts, and use it as a supplementary source of information for their econometric model. Despite the very low response rate (20 questionnaires received out of 446 sent), the authors found the results consistent with the IOM (1998) survey aimed at estimating the 'migration potential', which is briefly discussed in the next subsection.

4.3. 'Migration potential' assessment surveys

Other survey-based studies that are used to assess the future international migration flows are the analyses of 'migration potential'. Examples of such research with respect to the East-West flows in Europe are presented in the reports of Fassmann and Hintermann (1997), and of C. Wallace (IOM, 1998). Such studies are typically based on questionnaires filled in by a representative random sample of respondents, who are asked questions on whether they consider undertaking migration, from what reasons, under which circumstances, etc.

Results of the surveys are usually presented in a manner that does not address the issue of uncertainty. In fact, although the authors of such studies usually do not refer to their outcomes as forecasts *per se*, the results they present are often interpreted in this way by the users. For example, in summary studies dealing with the expected future migration flows, the survey-based estimates are often treated on similar footing with other forecasts. Moreover, in some of such studies the differentiation between 'migration potential' and the actual forecast is not made explicit. Importantly, this concerns several analyses prepared specifically for the policy-makers, as the research of Alvarez-Plata *et al.* (2003), or the Dutch CPB (2004) report.

Fassmann and Hintermann (1997) surveyed 4,392 persons in the Czech Republic, Hungary, Poland and the Slovak Republic. Depending on a country, they identified between 17 and 30 percent of respondents, expressing an overall willingness to leave their country, as a 'general migration potential', between 6 and 18 percent as the 'probable potential', concerning people who started to gather information about the possible destination, and between 1 and 2 percent as the 'real potential' – those, who have actually undertaken any particular steps to move. The preferred destinations of prospective migration were Germany and Austria, which is not surprising in the light of their geographic proximity to the countries under study, and of the

[8] Apparently, the majority of population flows concerning the ex-USSR in the period 1992–1998 proved to be internal migration, and migration between the former republics. For example, in Russia alone, over 20.5 million people migrated internally in that period, 5.5 million immigrated from abroad, and 2.6 million emigrated (Wegren and Cooper Drury, 2001: 16, 39). It is, however, worth noting that the two latter figures relate to a large extent to the population exchange with the other republics of the former Soviet Union (*idem*).

presence of strong migrant networks. Notably, only the 'real potential' category corresponds to the magnitudes of migrants that can be seen as plausible from the demographic point of view, whereas two other ones reflect rather 'wishful thinking' of the respondents. An additional problem of the study of Fassmann and Hintermann is a too small sample size with respect to the disaggregation of the results jointly by sex, age groups, regions, motives, education levels, etc.

In the Wallace's study (IOM, 1998), between 1,000 and 1,200 respondents in each of the eleven selected Central and Eastern European countries were asked six multiple-choice questions. In particular, the survey aimed at identifying, whether the respondents were going to migrate, for how long, why (or why not), do they have relatives or friends abroad, and have they already taken any preparatory steps. Depending on the type of migration potential distinguished by the author, the country-specific propensity to migrate ranged between 7 and 26 percent for permanent emigration, between 18 and 57 percent for long-term temporary labour migration, and between 13 and 68 percent for short-term labour migration (*idem*: 11). The preferred destinations were: the U.S.A. for permanent emigration, and Germany and Austria for labour migration.

The magnitude of the 'migration potential' assessment shows that the IOM (1998) survey identified dissatisfaction of the respondents with the state of affairs in a given country, rather than real migration perspectives. This conclusion can be supported by a fact that the highest percentages of 'potential migrants' have been found in Croatia and the then-Federal Republic of Yugoslavia, at that time both coping with a tense political situation. Very high estimates have been also obtained for Romania, the country to become, upon its accession to the European Union, the poorest EU Member State in terms of GDP *per capita*.

The IOM (1998) survey also contains some unexpected findings. For example, in Poland as many as 29 percent respondents identified 'ethnic tensions' as an important push factor out of the country (*idem*: 48). In an almost ethnically homogenous country this is a very surprising outcome, which likely indicates some problems with the execution of the survey, or with the translation of a questionnaire in such a multi-country study[9].

All the problems mentioned above support the conclusion that the results of the survey should be interpreted with caution. Another, general problem relates to the uncertainty issue: although the sample-based surveys as such have potential in addressing it in a statistical manner, this possibility is not addressed in the 'migration potential' assessment studies presented above.

[9] Credits go to Iza Koryś for drawing my attention to this curiosity.

4.4. Macro-level mathematical models in demography

With respect to mathematical models of migration, Kupiszewski (2002b: 16–17) noted that they predominantly stem from two different disciplines: demography and human geography. The former approach focuses predominantly on population distributions by sex and age, as well as on the impact of migration on the overall demographic dynamics and *vice versa*, while the latter – on spatial outcomes of the redistribution of migrants. Although both methodologies apply mathematical tools to model and forecast migratory flows within their framework, the classical demographic models of population dynamics use multi-dimensional linear algebra and calculus in a deterministic fashion. In turn, many geographical models are based on the Markov chains, thus using a stochastic approach, and as such are discussed in the next section.

Another important distinction is the one between macro-level (aggregate) and micro-level models[10]. In general, the macro-level demographic models of population change stem from the cohort-component approach, pioneered by Bernardelli (1941) and Leslie (1945) (after: Jóźwiak, 1992: 20). The cohort-component model is usually employed to forecast population size and age structure on the basis of judgementally-assumed scenarios of change in the particular components of demographic dynamics (originally, births and deaths) concerning the region under study. On their basis, the survivorship of particular birth cohorts is calculated in order to yield population size and age structure in subsequent projection steps[11].

Despite the continuous development of a cohort-component model, population forecasts until the 1970s either ignored the migration component or treated it simplistically. According to Rogers (1975: 1–2), more sophisticated methods in the mathematical studies of populations, including net migration, have been for the first time presented comprehensively in the report of the United Nations (1970)[12]. Since then, migration has been incorporated in an increasingly larger number of official population projections or forecasts. A survey on the methodology of such studies carried out in 30 developed countries in the early 1990s (Keilman and Cruijsen, 1992) showed that in a vast majority of cases, deterministic scenarios of all components of demographic change have been used as input. For international migration, Cruijsen and Keilman (1992: 20–22) found that at that time six countries did not include the migratory component in the projection models. Among the remaining ones, only Belgium used a regression model to extrapolate migration, while the other relied on assumptions on political

[10] A detailed overview of selected micro- and macro-level models of migration, with special focus on the latter (the Poisson regression, gravity models, and spatial interactions), is offered in Stillwell and Congdon (1991).

[11] The single-region mathematical demography, including the cohort-component approach, is presented in details for example in Keyfitz (1968), Rogers (1975: 7–55), and Jóźwiak (1992: 21–50).

[12] Paradysz (2004: 130) noted that an earlier study addressing the same problem in a comprehensive manner was the one of Korčak-Čepurkivs'kij (1970), which, however, was published only posthumously, as the author has been persecuted under the Stalinist regime, and spend 18 years in a Gulag forced labour camp (cf. the interview with A. Višnevskij in the *Demoscope weekly* 197–198, 4–17 April 2005; «www.demoscope.ru/weekly/2005/0197/analit01.php», accessed on 25 August 2006).

plans or targets, ‘visual’ interpolation of curves, or simply assumed constancy of migration rates throughout the forecast horizon.

Kupiszewski (2002b: 39–49) noted that the basic cohort-component model has ultimately evolved to include migration, following various methodological approaches. In the *migrant pool* method, applied for example in the 1970s and 1980s in the works of the U.S. Bureau of the Census, the numbers of migrants from all regions under study are projected, merged, and subsequently distributed among the receiving regions using an allocation algorithm. The *population accounting models* (Rees and Wilson, 1973; Rees and Willekens, 1986) are based on the rates of transitions or movements of people between different regions. In this approach, the balance of changes in population size on one hand, and births, deaths, in- and out-migrations on the other, needs hold, both on an aggregate level, as well as for each age group specifically. In turn, the *multi-regional model* (Rogers, 1975) is a generalisation of the cohort-component approach, based on the concept of a multi-dimensional life table, and treating jointly population of a system of regions, between which migration can occur. The projected numbers of migrants in particular age groups are thus an outcome of an analysis of demographic interdependencies involving all components of demographic change, including initial assumptions on migration propensities. The multi-regional approach is in turn generalised in the multi-state models (Rogers, 1980), where regions can be substituted by any other ‘states’ that an individual can ‘occupy’, e.g., marital status, economic activity, educational level, etc[13]. Examples of further extensions include non-linear multidimensional models, like the exponential LIPRO (LIfestyle PROjection) model developed by van Imhoff (1990), and van Imhoff and Keilman (1991).

Attempts to synthesise the multi-regional and population-accounting approaches are presented in details by Kupiszewska and Kupiszewski (2005), who themselves constructed the multi-level model MULTIPOLES (MULTIstate POpulation model for multi-LEvel Systems). The model treats migration on three different geographic levels (between regions, between the countries under study, and the population exchange with the rest of the world). Models like the MULTIPOLES offer an internally-coherent possibility of an analysis and forecasting of migratory phenomena on a very detailed level of disaggregation (by regions, sex and age), taking into account other demographic features of the population under study.

The deterministic nature of the presented mathematical models of population dynamics stems from their algebraic formulation, as well as from the dominant forecasting practice in demography. As it has been noted at the beginning of this section, the cohort-component or multi-regional models are until now usually fed with judgemental scenarios of particular components of demographic change, including migration. However, there exist probabilistic exceptions to this tendency, which are presented in Section 5[14].

[13] For a mathematical discussion of multi-regional and multi-state models, see Jóźwiak (1992: 51–94, 95–111).
[14] A discussion on the stochastic versions of population dynamics models is also offered by Jóźwiak (*idem*: 113–121).

4.5. Demo-economic modelling attempts

In addition to purely demographic forecasts, several authors made efforts to create models that would combine population and economic aspects of social development. From the point of view of migration forecasting, an interesting example is the recent study of Fachin and Venanzoni (2002). Their model for Italy, the IDEM (Integrated Demographic and Economic Model), combines a multi-regional cohort-component model of population dynamics with an economic input-output analysis (the Leontieff's table).

Migration between regions is one of key intermediary modules in IDEM, linking demography with economic aspects: labour supply and productivity. As the analysis is limited to a single country, the focus is on internal rather than international migration. In fact, the latter type of flows is treated very simplistically, as constant yearly inflows (in terms of numbers) and outflows (in terms of rates). In this way, international migration is exogenous, as are two other components of demographic change – mortality and fertility. Nonetheless, one can easily imagine an extension of the model to a multi-national economic system, where migration between particular countries would also play an important role.

In the IDEM, migration is modelled and forecasted in terms of origin-and-destination-specific rates, disaggregated by age according to the patterns observed in particular regions. The rates, as well as other exogenous and intermediary variables, are estimated and predicted using econometric models (in the case of migration, the logit regression). Nevertheless, despite of the stochastic character of the forecasting tools for particular model components, the very nature of the whole construction of the IDEM is algebraic and deterministic, similarly as in the case of macro-level models of mathematical demography.

As regards the possibilities of future research on possible ways to introduce a stochastic framework in the demo-economic, as well as in the multi-regional or multi-state demographic models, it is definitely worth considering to refer to over 50 years of experience of the theory of econometrics. Systems of linear or non-linear equations, as involved in the deterministic models presented above, can be made probabilistic in a relatively simple manner, following several methodological options widely used by econometricians. There have been a variety of research paradigms in econometrics over the last half of the century, changing from the simultaneous equations models, through the 'atheoretical' vector autoregression (VAR) models, to the structural modelling within the latter, involving, among other issues, the 'from general to specific approach' and the analysis of cointegration[15]. These developments offer very interesting paths of potential methodological advancements of the demographic and demo-economic forecasting in the future.

[15] See for example the discussion in Charemza and Deadman (1992/1997: 13–22, 151–166).

5. Probabilistic migration forecasts: assessing uncertainty

5.1. Markovian and related models of aggregate population flows

As noted in the previous subsection, one important group of stochastic models of migration stems from the tradition of human geography, with focus on spatial redistribution of population through migration. In particular, there are numerous examples of models applying the *Markov chains*, originally due to A. Markov (1906)[16]. The basic assumption of a Markov chain – a stochastic process X_t, which in a given discrete moment of time $t \in \{0, 1, \ldots\}$ can be in any of m states, $x_t \in \{1, \ldots, m\}$ – is the *Markov property* in its simplest form:

(7) $$p(X_t = x_t \mid X_{t-1} = x_{t-1}, \ldots, X_0 = x_0) = p(X_t = x_t \mid X_{t-1} = x_{t-1}).$$

The Markov property assumes that the process X_t 'lacks memory' from its past, so that the probability of being in a state x_t in the moment t depends only on the state that was occupied in the preceding period, t–1. This allows for formulating a transition matrix $\mathbf{P_t} = [p_{i,j,t}]_{mxm}$, where $p_{i,j,t} = p(X_t = j|X_{t-1} = i)$. If all $p_{i,j,t}$ are time-invariant (i.e. for every t, $\mathbf{P_t} = \mathbf{P}$), the Markov chain is called *homogenous*. In such cases, given the probability distribution of the process over the space of states (in migration studies: of regions) at the time t, $\mathbf{x_t}$, the model for use in forecasting the future distribution of population n periods ahead, $\mathbf{x_{t+n}}$, can be written as:

(8) $$\mathbf{x_{t+n}}' = \mathbf{x_t}' \mathbf{P}^n.$$

In general, the pioneering works in the applications of Markov chains to modelling social mobility[17] are due to Prais (1955) for movements of people between income classes, as well as Blumen *et al.* (1955) for changes of jobs. Also in migration studies, population flows between regions (states of a Markov chain) have been modelled in various ways, originating from chains with homogenous transition matrices. Specifically for inter-regional migration, the early examples are the studies of Rogers (1966), Brown (1970), and Joseph (1975). The pioneering work of Blumen *et al.* (1955) is also important, for it includes the notion of heterogeneity of the population under study. The population is divided into 'movers' and 'stayers', only the former ones changing the states they belong to. This approach has been later generalised in a model with different transition matrices for various subpopulations (Goodman, 1961). The population heterogeneity has been also formally examined by Spilerman (1972), who applied a regression analysis based on several exogenous variables within the Markov chain framework.

[16] See, for example: 'Markov chain' and 'Markov process' in the Springer online *Encyclopaedia of Mathematics* («eom.springer.de/M/m062350.htm», «eom.springer.de/M/mm062490.htm», accessed on 25 August 2006), as well as the review of probabilistic literature of the Kolmogorov Library («www.kolmogorov.pms.ru/uspensky-predvarenie.html», accessed on 25 August 2006), albeit the latter quoting the 1907 edition of the seminal work of Markov (1906).

[17] The overview is presented after Kupiszewski (2002b: 28–34).

The homogeneity of a Markov chain and the related stationarity of the process, meaning in particular that the conditional distribution of $\mathbf{x_t}$ given $\mathbf{x_0}$ stabilises as t increases to infinity, is desired from a mathematical point of view, although rarely observed in real-life migration modelling problems (cf. Huff and Clark, 1978). On the other hand, Brown (1970: 401) noted that certain properties of Markov chains, like mean times of passage between states, time of recurrence to a given state, or stationary distributions of a chain, are very useful for interpreting the properties of the system under study. For these reasons, there have been many attempts to retain certain assumptions of the model and lift some other ones, leading for instance to migration models using heterogeneous Markov chains with non-stationary transition matrices. An application of this approach can be found for example in models characterised by the 'cumulative inertia' property (McGinnis *et al.*, 1963). This notion links the transition probabilities with the duration of stay in particular states, assuming that the longer the stochastic process is in a given state (a person does not migrate), the less likely it will move to another state (region) in the future.

Characteristics of the 'cumulative inertia' and 'mover-stayer' approaches are combined in the semi-Markov processes proposed by Ginsberg (1971). These models are a generalization of Markov chains, and relate the probability of leaving a state to the time previously spent in it, as well as to the state entered next (destination). Although for such processes the Markov property does not hold, according to Ginsberg (*idem*), including duration-of-stay effects is a more suitable premise for modelling social processes than the Markovian 'lack of memory'.

Further to semi-Markov extensions, an attempt to construct a stochastic demo-economic model for forecasting interregional migration, has been made by Plane and Rogerson (1985). They combined a Markov chain with an economic gravity model (cf. Section 2), in which changes in the spatial distribution of economic opportunities modify the spatial distribution of the population. Their basic model has the following form:

(9)
$$p_{i,j}^{t} = \frac{p_{i,j}^{t-1}(A_j^t / A_j^{t-1})^{\gamma}}{\sum_k p_{i,k}^{t-1}(A_k^t / A_k^{t-1})^{\gamma}},$$

where $p^t_{i,j}$ is the probability of migration from i to j in the period (t, t+1), while A^t_j is the 'attractiveness' of the j-th region (Plane and Rogerson, 1985; after Rogerson, 1984: 115). A^t_j was defined here as $A^t_j=(\Delta E^{t-1}_j+s^{t-1}_j E^t_j)/(U^{t-1}_j+s^{t-1}_j E^t_j)$, with E^t_j and U^t_j respectively denoting employment and unemployment in the j-th region at the time t, and s^t_j being a rate, at which employees separate from their jobs in the period (t, t+1). Alternatively to (9), Rogerson (1984) formulated a logit model, additionally including the job application and competition process, through the region-specific rates of the numbers of job searchers to job vacancies.

Recently, Constant and Zimmermann (2003) combined a Markov chain model for the moves of migrants between Germany and their countries of origin, with an estimation of transition probabilities using logit models on the basis of micro-level data. For that particular example the authors found that “a Markov chain is an appropriate representation of the structure of the behavioral process of repeat migrants” (*idem*: 22).

With respect to the practical applicability of Markovian and related models in migration studies, Kupiszewski (2002b: 34–35) noted that they form a very general and elegant tool for modelling any transitions involving human populations, including the migratory ones. Moreover, Lindsay and Barr (1972) have presented a study, in which the use of Markov chains for modelling migration proved to be more realistic than the random Monte Carlo simulations based on dynamic transition probabilities, which would change over time according to pre-defined deterministic patterns. In these Monte Carlo simulations, the probabilities were derived using the gravity model, additionally acknowledging the presence of various barriers between the regions.

Despite their advantages, Markovian models are contemporarily of a rather marginal use in geographical studies, regardless of their relative popularity in the 1960s and 1970s. There are several reasons for that. Firstly, the assumption that a stochastic process lacks memory longer than from the preceding period is strong, and sometimes very far from the reality of migratory phenomena, for example with respect to return migration. Secondly, the ‘ideal’ properties of Markov chains, like the homogeneity of populations under study and stationarity of the stochastic process, are also rather artificial in the real-world modelling. Moreover, the above-listed attempts to overcome these problems, although very well-designed from a mathematical point of view, require very detailed empirical data, that are usually not available (Kupiszewski, 2002b: 34–35).

5.2. Micro-level methods: event-history analysis and ethnosurvey

In the micro-level event-history analysis, migration can be one of many possible demographic events that may happen to an individual. In the event-history modelling, people can move not only between geographic regions, but also between other socio-economic and demographic ‘states’, like the marital status, type of economic activity, level of education, etc., where the moves are modelled on the basis of estimated transition probabilities. The methodology is thus a natural analytical framework for use in multi-state demographic models (Rogers, 1980).

In human geography, the origins of the event-history approach date back to the studies of migration and residential histories. For example, Ginsberg (1978) used a probabilistic framework (a semi-Markov process mentioned in the previous subsection) to analyse residence histories of individuals, with focus on times between particular moves. In demography, the event-history framework in its wider sense, encompassing many possible aspects of human life, had numerous practical applications up to date. Among the first ones, there was a study of interrelations between migration, family formation and stage of career in the individual's life course (Courgeau, 1985)[18].

The event-history models can be formulated for either continuous or discrete time. In the former case, following Courgeau (1995: 23), let the duration of stay of an individual *a* be denoted by T_k^a, where *k*–1 is the number of migrations undertaken previously by that person. Further, let the respective region of residence be described by a random variable I_k^a. Additionally, let various characteristics of the individual *a*, $x_i^a(t)$, as well as sets of information this person has about the destination region *j*, $y_j^a(t)$, be given. The equation linking the instantaneous migration rate $m_{i,j,k}(t)$ of the individual *a* with the remaining quantities has the form:

$$(10) \qquad m_{i,j,k}^{a}(t \mid x_i^a(t), y_j^a(t)) = \lim_{dt \to 0} \frac{P(T_k^a < t + dt, I_{k+1}^a = j \mid T_k^a \geq t, I_k^a = i, x_i^a(t), y_j^a(t))}{dt}.$$

In an analogous model with discrete time, the probabilities of migration (transition) $p_{i,j,k,t}$ would be modelled instead of instantaneous rates, which are in fact also probabilities, although specifically defined.

Courgeau (*idem*) observed that (10) is "a multivariate model for failure-time data with competing risks", which may take into account not only variables related to migration as such, as duration of stay, sequence of migrations, or preferences of regions, but also different personal characteristics of migrants (sex, age, etc.). Additionally, other important events from the individual biographies, related for example to family formation, childbearing, changes in education level, as well as to employment status, can be considered in the model, provided that appropriate data are available.

The instantaneous probabilities of migration between particular regions, given in (10), can be estimated for example from representative retrospective surveys. Especially in the case of international migration, these rates should be ideally calculated on the basis of large comparable cross-country studies, which are nevertheless very costly and thus performed very seldom[19]. An alternative approach to prepare forecasts of migration on the basis of an event-history model are Monte Carlo micro-simulations, which are comprehensively presented for example by van Imhoff and Post (1998). As noted by Courgeau (1995: 24), the micro-

[18] Methodology, and an overview of possible demographic applications of the event-history analysis is provided in Courgeau and Lelièvre (1992), while a detailed analysis of migration dynamics is discussed in the life-course context by Mulder (1993).

[19] One well-known example is the Fertility and Family Survey (FFS), carried out in the 1990s in 20 European countries under the auspices of the Population Activities Unit (PAU) of the UN Economic Commission for Europe, following a common research design. Migration-related questions have been included in one of the optional modules of the original FFS. According to the author's knowledge, up till the time of writing of this study, the FFS has been repeated only once in a handful of originally-participating countries. Although the PAU also conducts another European-wide survey on *Generations and Gender*, the latter does not include any information about migration. More detailed information is available on the PAU website: «www.unece.org/pau» (accessed on 9 August 2006).

simulation methods in migration forecasting, allowing “to obtain an estimation of future regional population, under certain hypotheses”, date back to the pioneering work of Hägerstrand (1957). The hypotheses predominantly concern probabilities of making a migratory move within a given time after the previous one, and can be extended to include the impact of other socio-demographic characteristics. On the basis of such probabilities, the migratory behaviour of each individual from a population under study can be simulated using the Monte Carlo method. The individual-level trajectories can be subsequently aggregated in order to yield the required migration forecast. An important element assumed here, not always realistically, is the population homogeneity (Courgeau, 1995: 24).

Another micro-level research framework designed specifically for studying international migration is an ethnosurvey (Massey, 1987), combining the features of a quantitative survey with ethnographic methods. As traditional analytical techniques do not properly deal with undocumented migration, circular movements, and dynamic character of migratory processes, there was a need for a research design that would allow to capture these phenomena. The qualitative part of the ethnosurvey methodology involves semi-structured interviews, participant observations and case studies, while the quantitative one – gathering data on individuals, their life-event histories, as well as on households and communities, by the means of parallel sampling in the origin and destination areas. The main drawbacks of the ethnosurvey approach are: (1) it is cost- and labour-consuming, and (2) it is very difficult to obtain representative results from the samples used (*idem*: 1515).

Among other applications, the results of an ethnosurvey can be directly used to predict international migration flows. For example, Massey and Zenteno (1999) prepared a forecast of a Mexican migration to the United States, using a system of the following dynamic equations for the probability of immigration of an individual i in the year t, $p(Mig_{it})$, as well as for the probability of return, $p(Ret_{it})$, depending on the personal characteristics (*idem*: 5330):

(11a) $$p(Mig_{it}) = f(Age_{i,t}, Sex_i, Itrips_{i,t-1}, Iexp_{i,t-1}, Ctrips_{t-1}, Cexp_{t-1}),$$

(11b) $$p(Ret_{it}) = f(Age_{i,t}, Sex_i, Itrips_{i,t}, Iexp_{i,t-1}, Ctrips_{t-1}, Cexp_{t-1}).$$

In the above equations, the variables *Age* and *Sex* are self-explanatory, $Itrips_{i,t}$ denotes the number of migratory moves made by the i-th person in the year t, while $Ctrips_t$ – the number of such moves made by the other members of the community of origin. Further, $Iexp_{i,t}$ and $Cexp_t$ depict the migratory experience, measured by the number of moves, cumulated respectively by the i-th person, and by other community members in the year t. The probabilities (11a) and (11b) can be estimated within the event-history framework using a logistic regression, in order to obtain the rates of out- and return migration. Notably, although the quantitative part of an ethnosurvey is defined in probabilistic terms, there are also qualitative components of the research, not discussed in the current study, which make the whole methodology *de facto* a stochastic-descriptive (hence, deterministic) hybrid.

5.3. Selected attempts to bridge the micro and macro perspectives

With respect to the general discussion on combining the advantages of micro- and macro-level approaches in population modelling, Courgeau (2003) observed that both methodologies are complementary. In a comprehensive multi-level research framework, both individual and aggregate characteristics should be introduced into the model, as it has been applied for example in a study of inter-regional migration in Norway presented by Courgeau and Baccaïni (1998). In such a way, the macro-level variables provide an exogenous context for the analysis of individual biographies within the event-history approach. Interestingly, Courgeau (2003: 81–82) sees large methodological potential in applying the Bayesian paradigm and subjective probability in multi-level studies, involving hierarchical Bayesian modelling.

Another framework for migration modelling and forecasting, using the tools of theoretical physics has been proposed by Weidlich and Haag (1988) within an approach labelled as 'sociodynamics'. The authors developed a model that aimed at linking micro-level migratory decisions of the individuals with their macro-level outcomes for inter-regional population flows. On the micro level, migratory processes are described in terms of dynamic utility and mobility functions, applying the *master equation* method used in statistical physics.

Master equations are first-order differential equations defined for probability distributions of the process over a vector of N states, $\mathbf{n} = [n_i]_{1xN}$. In migration studies, they refer to the population distribution over the space of N regions, with n_i denoting the number of people in the i-th one. The probabilities are related to the rates of transition from one state (region) to another, w_{ij}, according to the formula (Weidlich and Haag, *idem*: 9–10, 23):

(12) $$\frac{dp(\mathbf{n},t)}{dt} = \sum_{i \neq j} w_{ij}(\mathbf{n}^{(ij)},t) \cdot p(\mathbf{n}^{(ij)},t) - \sum_{i \neq j} w_{ij}(\mathbf{n},t) \cdot p(\mathbf{n},t),$$

where $\mathbf{n}^{(ij)}$ depicts the state of the process after a single move of one person from the i-th state (region) to the j-th one, $\mathbf{n}^{(ij)} = [n_1, \ldots, (n_i-1), \ldots, (n_j+1), \ldots, n_N]'$.

On a macro level, the transition rates w_{ji} are derived using a regression analysis with several socio-economic and geographic (distance) explanatory variables, which constitute the background of migration processes. In Weidlich and Haag (*idem*), the model was tested on the examples of selected European countries: West Germany, France, Italy, and Sweden, as well as of Israel and Canada. Nevertheless, despite the potential usefulness of the model, and the level of detail and mathematical precision of a dynamic analysis it offers, its complexity rendered it hardly exploited in practical applications up to date.

At the time of writing of this paper, several leading European demographic research institutes are making a joint effort in order to develop a unified framework for demographic predictions within the EU-sponsored project 'Bridging the micro-macro gap in population forecasting'

(MicMac). The analysis aims at producing a forecasting model that would "offer a bridge between aggregate projections of cohorts […] and projection of the life courses of individual cohort members […]" (van der Gaag *et al.*, 2005: 3). The outcome would combine the features of multi-regional models with the event-history analysis, paying special attention to the uncertainty issue. Unfortunately for migration predictions, this component of population change is treated in MicMac much less comprehensively than mortality and fertility. Nevertheless, such a micro-macro approach has no doubt considerable potential also for modelling and forecasting of population flows[20].

5.4. Econometric forecasts of international migration

Econometric models are a natural tool not only to predict migration, but also to verify particular economic theories (cf. Section 2) on the basis of empirical data. The recent boom of using econometric models to forecast international migration in Europe dates back to the 1990s, and focuses almost universally on population flows from Central and Eastern European countries to the West after the expected enlargement of the European Union. Many such studies have been published in Austria and Germany, both located very close to the potential sending countries, and already accommodating large groups of immigrants from Central and Eastern Europe. Due to a multitude of such research, the current subsection presents only a handful of models in more detail, focusing on different methodological approaches to econometric modelling of migration, and, in most cases, explicitly referring to the European circumstances[21].

To start with, a very simple model has been proposed by Fertig and Schmidt (2000), who estimated immigration rates to Germany, m, from the four EU candidate countries at the time of preparing the study (the Czech Republic, Estonia, Hungary, and Poland). Their model covers the country-specific, time-specific, and cross-sectional effects, in addition to the overall mean migration rate, and can be written in a simple form:

(13) $$m_{i,t} = \mu + \varepsilon_i + \varepsilon_t + \varepsilon_{i,t}.$$

Denoting the country of origin by i, the model (13) assumes that $\varepsilon_i \sim N(0, \sigma_i^2)$, $\varepsilon_{i,t} \sim N(0, \sigma_{i,t}^2)$, and ε_t is a Gaussian autoregressive process AR(1). The forecast produced on its basis yielded that in the period 1998–2017 the average annual population inflow to Germany would range between 15 and 57 thousand immigrants a year (*idem*: 37).

The idea of Fertig and Schmidt (2000) has been further corroborated, among others, by Dustmann *et al.* (2003), who also forecasted European migration after the EU enlargement. In

[20] A brief description of the project, its rationale, basic assumptions, and methodology are offered for example by van der Gaag *et al.* (2005). More details can be found on the project's website: «www.micmac-projections.org» (accessed on 30 April 2006).

[21] More complete and detailed literature surveys of existing econometric forecasts are presented for example in Alvarez-Plata *et al.* (2003), in a report of the CPB (2004), as well as in Brücker and Siliverstovs (2005).

the latter study, the error term was decomposed into origin-specific, destination-specific, time-specific, and various cross-sectional components, additionally involving relative income *per capita*, and dummies for particular countries. Their approach also produced rather moderate forecasts of post-EU-enlargement migration flows.

In a more comprehensive modelling framework, Sinn *et al.* (2001) prepared the forecast of foreign population stocks in Germany (B), considering five largest EU candidate countries of that time: Poland, Romania, the Czech Republic, Hungary and the Slovak Republic. In the analysis, a partial adjustments model has been applied:

$$(14) \qquad B_t = \lambda\,[\alpha_0 + \alpha_1\, YV_t + \alpha_2\, \mathrm{G}_t + \alpha_3\, EU_t + \alpha_4\, FR_t + (1/\lambda - 1 + \alpha_5)\, B_{t-1}] + \varepsilon_t,$$

where λ satisfies the long-term relationship: $B_t = B_{t-1} + \lambda\,(B_t^* - B_{t-1})$, and B_t^* is an equilibrium trajectory of the foreign population stocks under study, $B_t^* = \alpha_0 + \alpha_1\, YV_t + \alpha_2\, \mathrm{G}_t + \alpha_3\, EU_t + \alpha_4\, FR_t + \alpha_5\, B_{t-1}$. A reference in B_t^* to the number of migrants in the preceding period, B_{t-1}, reflects the 'migrant network' hypothesis (*idem*: 8–9, cf. Section 2). The other predictors used in the model are: YV – fraction of PPP-adjusted GDP *per capita* in the sending country to the German one, G – the output gap in Germany, EU – a dummy variable regarding the EU membership, and FR – a dummy related to the freedom of movement of the labour force. The model has been in part estimated on the basis of empirical data, and in part (parameters α_3 and α_4) calibrated using the historical information on German population stocks originating from Greece, Italy, Portugal, Spain, and Turkey. As a result, the forecast yields that by 2015 the B_t would increase from the initial 459 thousand to 3.2–4.1 million people (*idem*: 18–21).

Another forecast of post-enlargement migration to the EU-15 from ten countries of Central and Eastern Europe has been prepared by Alvarez-Plata *et al.* (2003). Their study explains the share of migrants from country i residing in country j, expressed as a percentage of the total population of the latter ($ms_{i,j}$). The dependent variable has been modelled as:

$$(15) \qquad ms_{i,j,t} = \alpha + (1 - \delta)\, ms_{i,j,t-1} + \beta_1 \ln(w_{j,t}/w_{i,t}) + \beta_2 \ln(w_{i,t}) + \beta_3 \ln(e_{i,t}) + \beta_4 \ln(e_{j,t}) + \beta_5 \ln(P_{i,t}) + \gamma Z_{i,j} + u_{i,j,t},$$

where $u_{i,j,t} = \mu_{i,j} + v_{i,j,t}$ and $v_{i,j,t}$ denotes the Gaussian white noise. The remaining explanatory variables are: w – real income levels, e – employment rates, P – population sizes, and Z – dummy variables denoting geographic and cultural proximity of particular countries. Additionally, the forecast assumed a long-term convergence of economic explanatory variables (in particular, w), which concern ten Central and Eastern European countries, to the average EU-15 levels. As a result, net migration from the countries under study was forecasted to decline exponentially from 367 thousand persons a year, shortly after the introducing the freedom of labour force movement in the enlarged EU, to below zero by 2030 (*idem*: 60). It is worth noting that in model (15) the population size is exogenous, taken from the predictions of the World Bank, and hence is not internally consistent with the forecasted migration numbers.

The features of several types of the models presented above, (13), (14), and (15), have been combined in the study of Brücker and Siliverstovs (2005), who made an attempt to explain the share of migrant stocks from the country *i* in the country *j*, expressed as a percentage of the total population of the latter (thus, defined as in Alvarez-Plata *et al.*, 2003). In addition to including exogenous variables (real income levels and fractions, employment rates, etc.), the model has been prepared using the partial adjustments framework, where destination-specific shares of immigrants change stepwise towards their long-term equilibrium levels, and the error term is decomposed into the country-specific effect and the white noise. Their study also examines the impact of heterogeneity across countries on the estimation of migration models, as well as on the forecasts obtained.

Similar models, although with slightly varying sets of explanatory variables, have been proposed also by other authors. In addition to income (or GDP) *per capita* differences between the sending and receiving countries, present in virtually all models, Hatton and Williamson (1998) added a share of employment in agriculture, a birth rate of a sending country (lagged 20 years), migrant stock at the destination, and country-specific dummies. In their model, the dependent variables were emigration rates related to the population of the source country. Orłowski (2000) added population size and unemployment rates of the destination region, as well as geographic distance, in order to explain emigration rates. Boeri and Brücker (2001) found a significant effect of such variables as: employment rates, institutional restrictions to migrate, presence of migrant networks, cultural (language) proximity, as well as of standard of living, on the stocks of Central and Eastern European migrants residing in the 'old' EU-15 countries. Alecke *et al.* (2001) used unemployment rates in both the sending and receiving countries, alongside with country-specific dummies, in order to model the origin-destination emigration rates. Several models involving various macroeconomic determinants of country-specific net international migration rates in Europe have been also tested by Jennissen (2004).

Another option of econometric migration forecasting is based on generalised linear models (GLM). Willekens and Baydar (1986) suggested that in such models, there is a need to separate effects that are specific to the regions of origin, destination and their interactions, as well as the effects of the 'environment' (exogenous socio-economic, political, cultural, and other variables). They observed that overall migration can be decomposed into the *level effect* concerning the number of migrants, the *generation effect* related to the shares of particular sending areas in all outmigrations, and the *distribution effect* concerning the share of these outmigrations among all possible destinations. Letting N_t for the total number of migrants, $w_{i,t}$ for the probability that a migration originates from region i in the year t, and $p_{i,j,t}$ for the probability that a migration originating from i ends in j during t, the model for migration from i to j in time t ($m_{i,j,t}$) can be thus written as (*idem*: 207):

(16a) $$m_{i,j,t} = N_t \cdot w_{i,t} \cdot p_{i,j,t} \,,$$

or, for a system of regions, in a matrix notation (with $\mathbf{M_t}=[m_{i,j,t}]_{n \text{ x } n}$, diagonal $\mathbf{W_t}=[w_{i,t}]_{n \text{ x } n}$, and $\mathbf{P_t}=[p_{i,j,t}]_{n \text{ x } n}$):

(16b) $$\mathbf{M_t} = N_t \cdot \mathbf{W_t} \cdot \mathbf{P_t} \,.$$

For the elements of $\mathbf{W_t}$ and $\mathbf{P_t}$, additional explanatory analysis can be performed. Willekens and Baydar (*idem*: 225) suggested decomposing the logits of $w_{i,tn}$ and $p_{i,j,t}$ into parts related to regions, time, and their interrelations, possibly applying certain restrictions on the variables under study.

Another general category is formed by the simultaneous-equations econometric models of migration, the examples of which, however, are less numerous than of single-equation models. Already in the 1960s, Okun (1968) modelled interrelations between inter-state migration in the U.S.A. and the inequality of income levels *per capita* among the states. Later, Greenwood (1973) presented a complex model with seven equations and seven endogenous (jointly dependent) variables: the civilian labour force (*CLF*) out-migration (*OM*) and in-migration (*IM*), income change (*ΔInc*), employment change (*ΔEmp*), unemployment change (*ΔUnemp*), as well as *CLF* change (*ΔCLF*) and natural increase (*NatInc*). Equations for the two latter variables are identities ($\Delta CLF \equiv \Delta Emp + \Delta Unemp$; $NatInc \equiv \Delta CLF + IM - OM$), while the remaining ones are structural, involving several exogenous variables (proxies of age and education structures, government expenditures, and several dummies), as well as the stochastic component (Greenwood, 1973: 92–95). The simultaneous-equations models can be potentially very useful for forecasting migration within larger supra-national systems, yet the weak international comparability of migration data forms a clear practical limitation in that respect.

5.5. Stochastic forecasts of migration time series

In addition to econometric models involving additional explanatory variables, and thus modelling conditional distributions, another important group of migration forecasting approaches is based on the analysis and extrapolation of time series. The most common methodology, originally due to Box and Jenkins (1976), applies autoregressive integrated moving average (ARIMA) models, predominantly within the framework of the sampling-theory statistics. In general terms, the ARIMA(p,d,q) model can be formulated as (e.g., Greene, 2000: 776–777):

(17) $$\left(1-\sum_{i=1}^{p}\phi_i L^i\right)\left[(1-L)^d x_t\right] = c + \left(1-\sum_{j=1}^{q}\theta_j L^j\right)\varepsilon_t \,,$$

where x_t is the variable under study, ε_t is the error term (usually assumed to be a Gaussian white noise), L denotes the lag operator ($L^k x_t = x_{t-k}$), and $(1 - L)^d x_t$ is the d-th difference of x_t.

There are several examples of migration studies based on ARIMA models with various parameters p, d, and q. For instance, the forecast of Dutch migration prepared by de Beer (1997) uses an AR(1) autoregressive process $x_t = c + \phi x_{t-1} + \varepsilon_t$, for the total volume of emigration and immigration (x_t). In the same study, a moving average process MA(1), $x_t = c + \varepsilon_t - \theta \varepsilon_{t-1}$, has been found best fitting the data in the case of net migration. In a population forecast prepared for Finland, Alho (1998) applied ARIMA(0,1,1) models $x_t = c + x_{t-1} + \varepsilon_t - \theta \varepsilon_{t-1}$ for logarithms of immigration, as well as for emigration volumes. Keilman *et al.* (2001) prepared a Norwegian forecast using an ARMA(1,1) model for the logarithm of immigration: $x_t = c + \phi x_{t-1} + \varepsilon_t - \theta \varepsilon_{t-1}$, and an ARIMA(0,1,0) model, that is a random walk with drift: $x_t = c + x_{t-1} + \varepsilon_t$, for the logarithm of emigration. Further to the presented studies, the most important classes of stochastic models for forecasting demographic rates (random walks, linear stationary processes, and ARIMA) are also discussed in Alho and Spencer (2005: 198–225), together with the possibilities of their extension by allowing for example for heteroskedasticity of the error term, ε_t.

The time series models presented above have their multivariate generalisations, allowing for including other variables in addition to the key one under study (migration). The time series models presented above have their multivariate generalisations, allowing for including other variables in addition to the key one under study (migration). An example of this approach is the vector autoregressive (VAR) modelling, originally due to Litterman (1979) and Sims (1980) (after: Greene, 2000: 741), which has been used in the forecast of migration between Australia and New Zealand prepared by Gorbey *et al.* (1999). Their VAR(4) model, based on quarterly data, has the form:

$$\mathbf{X_t} = \mathbf{C_0} + (\mathbf{C_1} L + \mathbf{C_2} L^2 + \mathbf{C_3} L^3 + \mathbf{C_4} L^4)\, \mathbf{X_t} + \boldsymbol{\varepsilon}_\mathbf{t}. \quad (18)$$

where $\mathbf{C_i}$ are the matrices of cross-variable coefficients, L denotes the lag operator, and $\boldsymbol{\varepsilon}_\mathbf{t}$ is a multi-dimensional Gaussian white noise. The choice of variables in the VAR models has been based on the formal tests of the unit root, and the Granger-Sims tests of causality (*idem*: 78–84), which yielded the following sets of interdependent variables, $\mathbf{X_t}$, applied in the forecasts:

$$\mathbf{X_t} = [NMR_t, YRCGQ_t]', \quad (18a)$$
$$\mathbf{X_t} = [NMR_t, YRGQ_t, AUGQ_t]', \quad (18b)$$
$$\mathbf{X_t} = [D4NMR_t, YRCGQ_t, ZUGQ_t]', \quad (18c)$$
$$\mathbf{X_t} = [NMR_t, YRGQ_t, UDIFQ_t, ERGQ_t]'. \quad (18d)$$

In (18a)–(18d), the following abbreviations of particular variables are used: *NMR* – net migration rate, *D4NMR* – yearly difference of *NMR*, *YRGQ* – growth of the real GDP ratio for the two countries, *YRCGQ* – growth of the real GDP ratio *per capita*, *UDIFQ* – differences in unemployment rates, *AUGQ* and *ZUGQ* – country-specific unemployment growth indices for

Australia and New Zealand, and *ERGQ* – growth of the earnings ratio for the two countries. Although Gorbey *et al.* (*idem*) observed that migration between Australia and New Zealand is largely visa-free and resembles internal population flows, the same modelling framework can be also tested for typical international migration. The possible modifications include adjustments taking into consideration exogenous, not modelled, variables related to migration policies and to the freedom of movement of persons.

Apart from migration models based on the analysis and extrapolation of time series, there is an example of a forecasting framework that makes an attempt to partially depart from the sampling-theory statistical paradigm. Lutz *et al.* (1996, 1998, 2000, 2004) developed a concept of 'expert-based probabilistic population projections', in which subjective expert judgement is applied to prepare stochastic forecasts. In general, let v_t denote the phenomenon under study, for example international migration rates. Within the expert-based framework, the forecasting model has the form $v_t = \bar{v}_t + \varepsilon_t$, where $\bar{v}_t$ is the average trajectory of the process, assumed *a priori* by specialists in the field, and ε_t follows a stochastic process, e.g., AR(p) or MA(q).

Lutz *et al.* (2004) applied $\varepsilon_t \sim$ MA(30) for yearly data, assuming additionally that the standard deviation of ε_t, $\sigma(\varepsilon_t)$, is equal to a pre-defined value $\sigma^*(\varepsilon_t)$, also pre-set on the basis of subjective expert opinion. For international migration forecasts, the authors assumed that the average of the process is time-invariant ($\bar{v}_t = \bar{v}$), while $\sigma^*(\varepsilon_t)$ was calculated in such a way that 80 percent of the density of the probability distribution of v_t was concentrated between zero and the judgementally-chosen value v_{max}. As subjectivity is explicitly expressed in the 'expert-based' forecasting approach, it can be seen as a hybrid between the sampling-theory and Bayesian methods. Notably, a partial Bayesian interpretation of 'expert-based probabilistic population projections', has been given by Tuljapurkar (1997: 760), in terms of attaching "*a priori* probabilities to each static or dynamic scenario" in the forecasts presented by Lutz *et al.* (1997).

The features of various approaches to probabilistic population forecasting have been also synthesised within the framework of the EU-financed research project 'Uncertain Population of Europe' (UPE) (Alho *et al.*, 2005). The UPE predictions combine the cohort-component model of population dynamics with probabilistic forecasts of fertility, mortality, and migration, based on the analysis of time series, historical forecast errors, and on expert opinion. The results have been calculated on the basis of 3,000 simulations.

One specific innovation of the UPE Project is a comprehensive empirical analysis of correlations between forecast errors for various components of population change (including international migration), as well as between the countries under study (Alho *et al.*, 2005: 2). With respect to migration, the balance of flows for particular countries was modelled using linear trends, random walk models, or autoregressive models, with dummy variables for some country-specific years. Among the mentioned models, the AR(1) has been ultimately chosen for forecasting net migration among the 18 countries of the European Economic Area (Keilman and Pham, 2004)[22].

5.6. Bayesian models and forecasts of population flows

The existing examples of using Bayesian methods to model and forecast international migration are scarce. Therefore, the current overview applies a somewhat broader perspective, including additionally selected models of non-migratory population flows.

To begin with, Gorbey *et al.* (1999) estimated their VAR(4) models of migration between Australia and New Zealand defined in (18) also within the Bayesian framework. The coefficients of the models have been assumed to follow *Minnesota priors*. For the parameters on the first lags of the same, *i*-th variable, normal distributions N(1, $\sigma_{i,i,1}^2$) have been assumed *a priori*. In all other cases, for interrelations between the *i*-th variable and the *k*-th lag of the *j*-th variable, that is for $k > 1$ or $k = 1$ and $i \neq j$, normal prior distributions N(0, $\sigma_{i,j,k}^2$) have been used. In this way, the time series of each variable is *a priori* believed to be most likely generated independently by a random walk process. For the specific variances of the Gaussian priors, it has been assumed that $\sigma_{i,j,k} = \gamma\, g(k)\, f(i,j)\, s_i/s_j$, with $\gamma = 0.4$, $g(k) = k^{-1}$, $f(i,i) = 1$, $f(i,j) < 1$ for $i \neq j$, and s_i denoting a standard error in the autoregressive model for the *i*-th variable.

The *ex-post* comparison of various sampling-theory and Bayesian models yielded that the best-performing ones were the ones given by (18b), with $\mathbf{X_t}$ comprised of net migration rates, growth of the real GDP ratio for the two countries, and quarterly unemployment growth in Australia. Among them, the Bayesian model produced slightly greater *ex-post* forecast errors than the corresponding traditional VAR, one possible explanation being the likely disagreement between the prior distributions and the data sample.

Nevertheless, from a strictly Bayesian point of view, there is one problem with the coherence of such approach. As the s_i values are estimated from the same sample as the model itself, the priors are data-based. Therefore, the presented approach is not fully Bayesian, as in that case the prior distributions should be specified independently from the observations[23].

The research of Brücker and Siliverstovs (2005), mentioned earlier in this section, also compares the results of estimation of migration models in the sampling-theory and Bayesian approaches. In their comparison of various estimation methods, the hierarchical Bayes estimator (likely the mean from the appropriate posterior distribution), as well as the sampling-theory fixed-effects estimator, performed best in terms of *ex-post* prediction errors. The problem is, however, that the authors consider the Bayesian framework merely as an

[22] More information is available from the project's website: «www.stat.fi/tup/euupe» (accessed on 5 May 2006).
[23] I am very grateful to Prof. Jacek Osiewalski for drawing my attention to this problem.

alternative methodology of estimation. The study lacks both the discussion of the prior distributions used, as well as the *a posteriori* uncertainty assessment, which elements are both inherent in any Bayesian analysis.

In addition to the studies of 'pure' migratory movements, examples of Bayesian gravity models based on the Poisson regression, applied to forecasting other types of population flows, have been offered by Congdon (2000, 2001)[24]. It should be noted, that although his studies focus on the flows of patients to hospitals, the models are sufficiently general to cover other types of spatial movements in a similar manner. The common assumption is that the number of patients coming to the hospital *j* from the region *i* follows a Poisson distribution with a mean $\mu_{i,j}$, for which various alternative models can be built, as for example (Congdon, 2001: 297–298):

(19a) $\ln(\mu_{i,j}) = k + \alpha_0 \ln(P_i) + \delta \ln(R_{i,j}) + \xi_1 E_{1,j} + \xi_2 E_{2,j} + \phi S_{i,j}$, or:

(19b) $\ln(\mu_{i,j}) = k + \alpha_0 \ln(P_i) + \alpha_1 YAN_i + \alpha_2 Aged_i + \delta_i \ln(R_{i,j}) + \xi_1 E_{1,j} + \xi_2 E_{2,j} + \phi S_{i,j}$.

The basic set of explanatory variables is comprised of P_i – population size of the *i*-th region, and $R_{i,j}$ – a proxy of the medical services supply, calculated as the 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 further exogenous predictors are: YAN_i – an index of demand for the health-care services, $Aged_i$ – a fraction of population aged 65 years or more, $E_{1,j}$, $E_{2,j}$ – 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 assumed to be Gaussian, diffuse for the constant, $k \sim N(1, 100^2)$, and carrying more information in the case of the remaining parameters: $\alpha_i, \beta, \delta \sim N(1, \sqrt{10}^2)$, $\gamma \sim N(2, \sqrt{10}^2)$, and $\xi_1, \xi_2, \phi \sim N(0, \sqrt{10}^2)$. The model (19b), in its variant based on car-time distance, has been found fitting the data best.

It is worth noting that the scarcity of Bayesian forecasts of international population movements resembles the situation in the other fields of demographic analysis. The rare existing examples include forecasts of fertility (Tuljapurkar and Boe, 1999), mortality (Girosi and King, 2005), or the whole population size in the case of limited information, as in the study of the Iraqi Kurdish population by Daponte *et al.* (1999). Population change has also been analysed in the Bayesian framework by Bernardo and Muñoz (1993), and Clark (2003), the latter from an ecological perspective, not limiting the analysis to the human species.

Interestingly, according to Alho (1999: 1), one of the first attempts to formulate the population prediction problem in probabilistic terms instead of producing the baseline-high-low intervals has been made by Törnqvist (1949) within the Bayesian framework. More

[24] A summary of various probability models that can be used for migration forecasts is offered in Willekens (2005), who argues for using Poisson regression for counts of migrants, logit or logistic models for proportions of migrants in a given population, and Poisson models with offset for occurrence-exposure rates. The discussion further distinguishes models for state occupancies, transition probabilities, and transition rates.

contemporarily, a complex Bayesian analysis of demographic change using a cohort approach has been proposed by Nakamura (1986), while the application of Bayesian inference in the studies on the Lexis diagram has been shown by Berzuini *et al.* (1993). The latter concept has been recently extended in a form of a specialised software 'BAMP' designed for the age-period-cohort modelling and forecasting (Schmid and Knorr-Held, 2001). Some other examples of Bayesian models that can be used in demography and related fields are given by Courgeau (2004), and include multi-level or hierarchical models (Draper, 1995; Goldstein, 2003), as well as survival models (Ibrahim *et al.*, 2001; Gustafson *et al.*, 2003).

6. Evaluation of existing migration forecasting methods and models

6.1. Problems with the use of survey results as migration forecasts

Subsequent parts of this section are devoted to the advantages and drawbacks of various methods and models presented before, also with respect to the way they deal with the uncertainty issue. The relatively largest number of problems with the interpretation of the results considers survey studies used to predict population flows. As noted by Kupiszewski (2002a: 633–637), the categories used in such analyses are defined very vaguely. For example, the commonly-used 'migration potential' can be defined differently in various studies, depending on what is asked in the questionnaire. In this way, the formulation of questions may heavily influence the results of a survey. There are no doubt different numbers of 'potential migrants' giving a positive answer to the question, whether they intend to migrate any time in the future, or whether they have already actively searched for any real possibilities of settling abroad.

Moreover, the migration *intentions* declared by the respondents can not be directly transformed into the actual behaviour, i.e. into the magnitude of population flows in the future. In that respect, either the declared intentions can reflect 'wishful thinking' rather than reality, or additional factors may get involved in the decision-making process between the moment of the survey and the expected time of migration. The problem is thus not with survey-based research as such, but with interpreting its results as migration forecasts, irrespective of clear disclaimers and warnings provided by the authors of such studies (e.g., IOM, 1998: 11). On the top of that, due to high costs of conducting a survey, the sample sizes are usually to small to allow for obtaining significant results involving breakdowns by sex, age groups, regions, motives, etc.

6.2. Deterministic character of most of the existing predictions

As noted in Section 4, most of the practical applications of the cohort-component and multi-regional demographic models are based on judgemental scenarios of particular components of demographic change, including migratory flows. Although the input for the demographic models could easily consist of stochastic forecasts of these variables, a majority of demographic applications stick to the multi-variant projections, where uncertainty is not properly quantifiable. The exceptions are based on the event-history analysis, in which micro-level simulations are used to estimate probabilities of transition between particular states, thus operating within a probabilistic framework (Section 5). Also many geographic models, despite being based on stochastic processes (notably, Markov chains), and thus having large potential of simulation-based uncertainty assessment, do not seem widely explored in this context up to date.

Moreover, the judgemental scenarios that often form the input to the models of population dynamics often assume the constancy of migratory flows starting from a given period of time. Although such an approach is understandable, as given the lack of specific knowledge of the researcher about the more distant future, the constancy assumption seems to be the neutral option, it is also likely to generate very high *ex-post* forecast errors. Nevertheless, as international migration will most probably continue to be a dynamic process, this approach cannot be expected to produce reasonable results in the long term.

To address the problems listed above, an appropriate option would be to involve forecasts based on econometric or time series models, applying either the sampling-theory or Bayesian framework, but nevertheless bearing in mind several reservations. For example, although the analysis of uncertainty is embodied in the very nature of econometric models, many studies devoted to migration forecasting do not pay proper attention to this issue. In the discussion of the results, the presentation of forecast errors in terms of prediction intervals is often missing (e.g., in Orłowski, 2000; Boeri and Brücker, 2001; Sinn *et al.*, 2001; Alvarez-Plata *et al.*, 2003; Brücker and Siliverstovs, 2005). In most cases, this is replaced by a scenario analysis, based on different assumptions on the GDP growth and convergence. An exception is the study of Fertig and Schmidt (2000), who based their high-migration variant on the mean forecast plus one standard deviation. In contrast to the practice in econometric forecasting of migration, uncertainty is explicitly addressed in the predictions based on the pure time series models (e.g., de Beer, 1997; Alho, 1998; Keilman *et al.*, 2001; Lutz *et al.*, 2004; Alho *et al.*, 2005).

All remarks made with respect to the sampling-theory econometric and stochastic forecasting models remain in force for their Bayesian equivalents. The latter address the issue of uncertainty in a form of whole predictive distributions, from which the credible intervals can be derived, in order to assess the uncertainty span. In addition, the Bayesian inference allows for a formal incorporation of expert judgement into the model, with respect to the characteristics of the processes and interactions between variables. This is possible through assuming appropriate informative prior distributions of the model parameters. The expert opinion can be obtained for example from surveys among experts or Delphi studies, similar to the ones presented in Section 4, or subjectively assumed by the forecaster.

6.3. Specification problems in econometric models of migration

Despite the clear advantage of econometric studies in the context of uncertainty, they are often being criticised for the shortcomings of model specification, especially with respect to demographic variables and country-specific effects, which are missing in most of them (Fertig and Schmidt, 2000; Alecke *et al.*, 2001). Limiting the possible explanations of population flows to the economic aspects is also challenged by Kupiszewski (2001), as often leading to very high forecast errors. Ideally, the forecasting model should control at least the basic demographic characteristics, as the populations size and age structure. In that respect, Kupiszewski (2002a: 637–641) argues that if crude numbers of migrants are forecasted instead of the occurrence-exposure rates and there are no demographic constraints on migration, this may lead to extreme results, like the ones obtained in the judgemental scenarios of Franzmeyer and Brücker (1997).

To avoid extremities, a reasonable level of control of the limiting factors should be exerted, depending on various dimensions (demographic, social, economic, environmental, and cultural) of the forecasting context, as well as on the scale of the predictions (Cohen, 1998). Examples of applications of models successfully taking into account various constraints of the socio-demographic setting include for example Massey and Zenteno (1999). A possible analysis of the impact of the age composition of the population on migration may follow for example the hypothesis of Plane (1993), who asserted that age-specific internal mobility rates of young people generations (especially aged 20–24 years) are visibly lower in the 'baby-boom' generations than in the other ('baby-bust') ones. However, such hypotheses still need to be carefully verified on the basis of empirical data, before being applied in the migration forecasting practice.

When migration rates or shares of migrants in the total population are forecasted within a model which uses population size as one of the explanatory variables, as for example in Alvarez-Plata *et al.* (2003), the movements of people occur *de facto* outside the model. In a coherent framework, migration flows would increase the population of the destination country and diminish in the source country by the same number, changing also the denominators of migration rates or shares of migrants between subsequent forecast periods. Otherwise, treating population size as exogenous is another source of bias, as in the study of Alvarez-Plata *et al.* (*idem*), who used the population stocks projected by the World Bank as predictors. The consequences of this omission can be very serious especially for regions with rapid population growth or decline.

A way to overcome the incoherence caused by population moves in such models is to examine the forecasted rates or shares in a broader model of demographic dynamics. In such a way, the rates or shares would be transformed into numbers of migrants according to the changes of the population size predicted within the same model. For this purpose, a multi-regional framework of Rogers (1975) can be used, or its extensions, like the exponential LIPRO model (van Imhoff, 1990; van Imhoff and Keilman, 1991), the MUDEA (MUlti-level DEmographic Analysis) model (Willekens, 1995), which combines the features of the population accounting and the multi-regional methodologies, or the generalised multi-level MULTIPOLES approach (Kupiszewska and Kupiszewski, 2005). Another possibility is to use a statistically-coherent framework proposed within the framework of the 'Uncertain Population of Europe' project (Alho *et al.*, 2005), taking into account interrelations of migratory processes between the countries and regions under study.

6.4. Implications for migration studies

In an attempt to create a complex conceptual model for analysing and forecasting international migration, Salt and Singleton (1995) suggested that population flows should ideally be predicted separately for different categories of migrants. Van der Gaag and van Wissen (1999) successfully tested this idea on several models for immigration into five EU countries, distinguishing nationals, EU-foreigners and non-EU-foreigners, as well as examining several migration motives (economic migration, asylum-seeking, elderly migration).

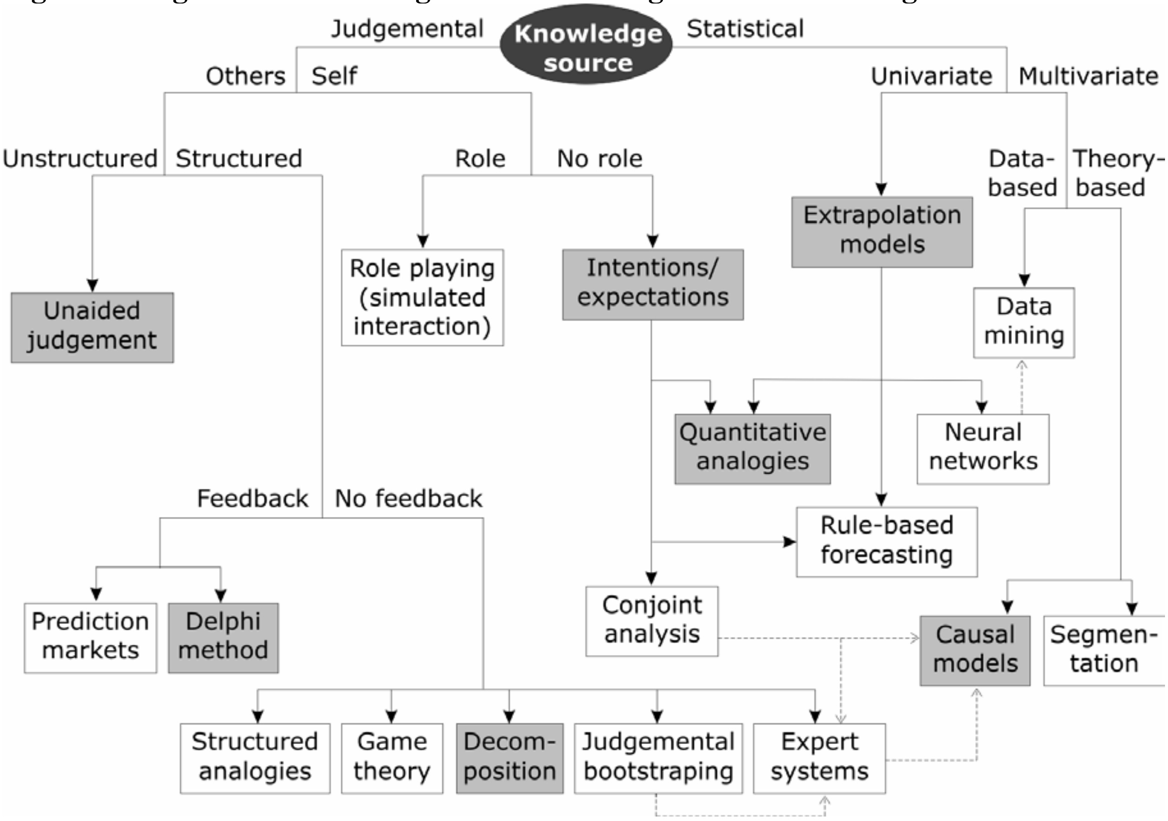
With respect to various methodological options of stochastic modelling and forecasting of demographic phenomena, Lee (1998) suggested that the choice of a 'proper' method should depend on the availability of good-quality data sets. Longer series of observations enable using the time series approach, while for shorter samples or less reliable data, the expert-based probabilistic scenarios are recommended. In order to synthesise these methodologies, Tuljapurkar *et al.* (2004) suggested that the new directions in demographic forecasting should include hybrid models, combining time series and the expert-based scenarios.

As a consequence of the argumentation presented above, the methodology of migration forecasting should predominantly focus on applying econometric and time series models. The deterministic methods are not recommended due to their disregard for the uncertainty issue, as for example in the case of judgemental scenarios. In turn, the survey-based approach does not qualify as a proper forecasting method in the light of the discussion offered before. Nevertheless, a survey among experts or a Delphi study concerning either the parameters of a forecasting model, or migratory flows and their uncertainty assessment (advocated for example by Lutz *et al.*, 1996, 2004) could potentially be an interesting extension of the analysis.

A detailed discussion of the ways of incorporating judgement in migration forecasts, including insights into the Bayesian approach, has been provided by Willekens (1994), whose remarks may serve as a point of reference for an expert opinion-based study. Although such possibilities are generally very broad, they are clearly worth addressing in separate research. On the other hand, the remaining probabilistic methods and models (Markov chains, the event-history analysis, or the 'sociodynamic' approach) are complex and require much more detailed statistical information than there usually is available for international population flows in Europe. An investigation into some of these options, although very promising, like in the case of the event-history approach bridged with macro-level studies of demographic dynamics, would require a separate, very extensive (and expensive) research project.

In general, there is still a large potential of applying alternative forecasting methods, which have not been used in migration context up to date. A full account of methodological possibilities, provided by Armstrong and Green (2005), is illustrated in Figure 3, together with the indication of the position of methods already used in migration forecasting. And thus, the ‘intentions/expectations’ analysis covers ‘migration potential’ assessment surveys. Judgemental migration scenarios can be classified either as ‘unaided judgement’, or ‘quantitative analogies’, depending on whether or not the forecasts are supported by formal methods, and use available information on analogous situations from the past, respectively from other countries. Among the judgement-based methods in migration studies, there are also examples of the use of the Delphi approach, as indicated in Section 4. The same applies to the judgemental decomposition of migration into various groups of migrants, which are further analysed using formal methods based on available statistics, as for example in the study of van der Gaag and van Wissen (1999).

Figure 3. Migration forecasting methods in the general methodological framework



Notes: Grey shading denotes types of methods used in migration forecasting, straight lines – the relationships between various methods, while dotted lines – possible relationships.
Source: Armstrong and Green (2005), own elaboration (migration forecasting methods)

At the same time, the vast majority of the existing studies of migration either fall into the class of ‘causal models’, involving econometric analyses, or ‘extrapolation models’, including deterministic demographic and demo-economic forecasts, and predictions based on stochastic processes: Markov chains and time series analysis. ‘Extrapolation’ also covers the micro-macro bridging perspectives, as well as the outcomes of event-history studies or the quantitative parts of ethnosurveys. As a consequence, this category refers to a very broad group of methods, covering deterministic and probabilistic approaches alike, according to their potential of addressing the uncertainty issue in a coherent quantitative manner.

Specifically in the probabilistic approach, further paths of methodological developments could possibly include the investigation of data mining techniques based on various sources ('letting the data speak for themselves'), stochastic versions of neural networks, and rule-based forecasting, or generalised versions of expert systems, devised in a way addressing the uncertainty issue. As noted by Armstrong and Green (2005), in the rule-based forecasting, "expert domain knowledge and statistical techniques are combined using an expert system to extrapolate time series. Most series features are identified by automated analysis, but experts identify some factors. In particular they identify the causal forces acting on trends". The latter approach seems thus especially promising in the context of such a multi-dimensional phenomenon as international migration, where the judgemental element is very important. Again, all the mentioned options require conducting detailed research in the future, dedicated particularly to the examination of the applicability of specific methods in international migration forecasting.

References

Alecke, B., Huber, P., and Untiedt, G. (2001). What difference a constant makes? How predictable are international migration flows? In: *Migration policies and EU enlargement. The case of Central and Eastern Europe*. OECD, Paris: 63–78.

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

Alho, J. (1999). *On Probabilistic Forecasts of Population and Their Uses*. Paper for the 52nd Session of the International Statistical Institute, Helsinki, 10–18 August.

Alho, J., Alders, M., Cruijsen, H., Keilman, N., and Nikander, T. (2005). *New forecast: population decline postponed in Europe*. Paper for the Joint UNECE-Eurostat Work Session on Demographic Projections, Vienna, 21–23 September.

Alho, J., and Spencer, B. (2005). *Statistical Demography and Forecasting*. Springer, Berlin-Heidelberg.

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

Armstrong, J. S., and Green, K. C. (2005). Methodology Tree. In: J. S. Armstrong, and K. C. Green, *Forecasting Principles* (online resource), Marketing Department of The Wharton School, University of Pennsylvania, Philadelphia, PA; «www.forecastingprinciples.com/ methodologytree.html» (accessed on 24 August 2006).

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

Bauer, T., and Zimmermann, K. F. (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. Institut zur Zukunft der Arbeit, Bonn.

Bernardelli, H. (1941). Population waves. *Journal of the Burma Research Society*, 31: 1–18.

Bernardo, J. M., and Muñoz, J. (1993). Bayesian analysis of population evolution. *The Statistician*, 42(4): 541–550.

Berzuini, C., Clayton, D., and Bernardinelli, L. (1993). Bayesian Inference on the Lexis Diagram. *Bulletin of the International Statistical Institute*, 55(1): 149–165.

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

Bijak, J., Kupiszewska, D., Kupiszewski, M., Saczuk, K., and Kicinger, A. (2006). Population and labour force projections for 27 European countries, 2002–2052: impact of international migration on population ageing. *European Journal of Population*, forthcoming.

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

Boeri, T., and Brücker, H. (2001). *Eastern Enlargement and EU Labour Markets: Perceptions, Challenges and Opportunities*. IZA Discussion Paper 256. Institut zur Zukunft der Arbeit, Bonn.

Borjas, G. (1990). *Friends or Strangers: The Impact of Immigrants on the US Economy*. Basic Books, New York.

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

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

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

Castro, L. J., and Rogers, A. (1983). Patterns of Family Migration: Two Methodological Approaches. *Environment and Planning A*, 15(2): 237–254.

Charemza, W. W., and Deadman, D. F. (1992). *New Directions in Econometric Practice*. Edward Elgar, Aldershot [(1997). *Nowa ekonometria*. PWE, Warszawa].

Chojnicki, Z. (1977). Podstawy metodologiczne prognozowania w geografii ekonomicznej [Methodological foundations of forecasting in economic geography]. *Przegląd Geograficzny*, 49(2): 247–261.

Clark, J. S. (2003). Uncertainty and Variability in Demography and Population Growth: A Hierarchical Approach. *Ecology*, 84(6): 1370–1381.

Cohen, J. E. (1998). Should Population Projections Consider "Limiting Factors" – and If So, How? *Population and Development Review*, 24(Suppl.): 118–138.

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.

Constant, A., and Zimmermann, K. F. (2003). *The Dynamics of Repeat Migration: A Markov Chain Analysis*. IZA Discussion Paper 885. Institut zur Zukunft der Arbeit, Bonn.

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

Courgeau, D., and Lelièvre, E. (1992). *Event history analysis in demography*. Oxford University Press, Oxford.

Courgeau, D. (1995). Migration theories and behavioural models. *International Journal of Population Geography*, 1(1): 19–27.

Courgeau, D., and Baccaïni, B. (1998). Multilevel analysis in the social sciences. *Population–E*, 10(1): 39–71.

Courgeau, D. (2003). From the macro-micro opposition to multilevel analysis. In: D. Courgeau (ed.), *Methodology and Epistemology of Multilevel Analysis*. Kluwer, Dordrecht: 43–92.

Courgeau, D. (2004). Probabilités, demographie et sciences sociales. *Mathématiques et Sciences humaines*, 42(3) : 27–50.

CPB [Centraal Planbureau] (2004). *Arbeidsmigratie uit de Midden- en Oost-Europese toetredings-landen* [*Labour migration from the Central and Eastern European accession countries*]. Notice to the Dutch Ministry of Social Affairs and Employment. CPB, The Hague.

Cruijsen, H., and Keilman, N. (1992). A comparative analysis of the forecasting process. In: N. Keilman, and H. Cruijsen (eds.), *National population forecasting in industrialised countries*. NIDI & CBGS publication 24. Swets & Zeitlinger, Amsterdam: 3–25.

Dalkey, N. C. (1969). *The Delphi Method: An Experimental Study of Group Opinion*. Research Memorandum RM-5888-PR. RAND Corporation, Santa Monica, CA.

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

Dawid, A. P. (1984). Statistical Theory: The Prequential Approach. *Journal of the Royal Statistical Society A*, 147(2): 278–292.

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.

DeJong, G. F., and Fawcett, J. T. (1981). Motivations for Migration: An Assessment and a Value-Expectancy Research Model. In: G. F. DeJong, and R. W. Gardener (eds.), *Migration Decision Making: Multidisciplinary Approaches to Microlevel Studies in Developed and Developing Countries*. Pergamon Press, New York: 13–57.

Draper, D. (1995). Inference and hierarchical modelling in the social sciences (with discussion). *Journal of Educational and Behavioral Statistics*, 20(2): 115–147, 233–239.

Drbohlav, D. (1996). The Probable Future of European East-West International Migration – Selected Aspects. In: F. W. Carter, P. Jordan, and V. Rey (eds.), *Central Europe after the Fall of the Iron Curtain; Geopolitical Perspectives, Spatial Patterns and Trends*. Wiener Osteuropa Studien 4. Peter Lang & Co., Frankfurt: 269–296.

Dustmann, C. (1997). Return migration, uncertainty and precautionary savings. *Journal of Development Economics*, 52(2): 295–316.

Dustmann, C., Casanova, M., Fertig, M., Preston, I., and Schmidt, C. M. (2003). *The impact of EU enlargement on migration flows*. Home Office Report 25/03. Home Office, London.

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

Fachin, S., and Venanzoni, G. (2002). *IDEM: an Integrated Demographic and Economic Model of Italy*. Paper for the 14th International Conference on Input-Output Techniques, Montreal, Canada, 10–15 October.

Faist, T. (2000). *The Volume and Dynamics of International Migration and Transnational Social Spaces*. Oxford University Press, Oxford.

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

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

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

Gawryszewski, A. (1989). *Przestrzenna ruchliwość ludności Polski 1952–1985* [*Spatial mobility of the population of Poland, 1952–1985*]. Ossolineum, Wrocław.

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

Ginsberg, R. B. (1978). Probability Models of Residence Histories: Analysis of Times between Moves. In: W. A. V. Clark, and E. G. Moore (eds.), *Population Mobility and Residential Change*. Studies in Geography 25. Northwestern University, Evanston, IL.

Girosi, F., and King, G. (2005). *Demographic Forecasting*. Manuscript. Center for Basic Research in the Social Sciences, Harvard University, Boston, MA; «gking.harvard.edu/files/smooth.pdf» (accessed on 10 June 2006).

Goldstein, H. (2003). *Multilevel statistical models*. Arnold, London.

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

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

Greene, W. H. (2000). *Econometric Analysis*. Prentice Hall, Upper Saddle River, NJ.

Greenwood, M. J. (1973). Urban economic growth and migration: their interaction. *Environment and Planning A*, 5(1): 91–112.

Gustafson, P., Aeschliman, D., and Levy, A. R. (2003). A simple approach to fitting Bayesian survival models. *Lifetime data analysis*, 9(1): 5–19.

Hägerstrand, T. (1957). Migration and area. Survey of a sample of Swedish migration fields and hypothetical considerations on their genesis. In: D. Hannerberg, T. Hägerstrand, and B. Odeving (eds.), *Migration in Sweden. A Symposium*. "Lund Studies in Geography", Series B, no. 13, Gleerup, Lund: 27–158.

Harris, J. R., and Todaro, M. P. (1970). Migration, unemployment and development: A two-sector analysis. *American Economic Review*, 60(1): 126–142.

Hart, R. A. (1975). Interregional economic migration: Some theoretical considerations (Part I). *Journal of Regional Science*, 15(2): 127–138.

Hatton, T. J., and Williamson, J. G. (1998). *The Age of Mass Migration: Causes and Economic Impact*. Oxford University Press, Oxford.

Huff, J. O., and Clark, W. A. V. (1978). The Role of Stationarity in Markov and Opportunity Models of Intraurban Migration. In: W. A. V. Clark, and E. G. Moore (eds.), *Population Mobility and Residential Change*. Studies in Geography 25, Northwestern University, Evanston, IL.

Ibrahim, J. G., Chen, M.-H., and Sinha, D. (2001). *Bayesian survival analysis*. Springer, Berlin-Heidelberg.

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

Isard, W. (1960). *Methods of regional analysis: an introduction to regional science*. John Wiley, New York [(1965). *Metody analizy regionalnej. Wprowadzenie do nauki o regionach*. PWN, Warszawa].

Janicki, W. (2005). *Migracje wewnętrzne w Unii Europejskiej i ich uwarunkowania* [*Internal migration in the European Union and its determinants*]. PhD Dissertation. Maria Curie-Skłodowska University, Lublin.

Jennissen, R. (2004). *Macro-economic determinants of international migration in Europe*. Dutch University Press, Amsterdam.

Joseph, G. (1975). A Markov Analysis of Age/Sex Differences in Inter-regional Migration in Britain. *Regional Studies*, 9(2): 69–78.

Jóźwiak, J. (1992). *Mathematical models of population*. NIDI Report 26. Netherlands Interdisciplinary Demographic Institute, The Hague.

Keilman, N., and Cruijsen, H. (eds.) (1992). N*ational population forecasting in industrialised countries*. NIDI & CBGS publication 24. Swets & Zeitlinger, Amsterdam.

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

Keilman, N., and Pham, D. Q. (2004). *Empirical errors and predicted errors in fertility, mortality and migration forecasts in the European Economic Area*. Discussion Paper No. 386. Social and Demographic Research, Statistics Norway, Oslo.

Keyfitz, N. (1968). *Introduction to the Mathematics of Population*. University of California Press, Berkeley, CA.

Korčak-Čepurkivs'kij, Ju. A. (1970). Rozrachunok tempu zrastanja kilkosti i vikovoho skladu tych, ščo živut' u ustabilizovanomu naselenni z postijnym rivnem mihracij. *Demohrafična Doslidžennija*, 1: 24–41.

Korcelli, P. (1994). On interrelations between internal and international migration. *Innovation*, 2: 151–163.

Koryś, P., and Okólski, M. (2004). *Czas globalnych migracji. Mobilność międzynarodowa w perspektywie globalizacji* [*Time of global migration. International mobility in the perspective of globalisation*]. ISS Working Papers "Prace Migracyjne" 55. Center for Migration Research, Warsaw.

Kritz, M., Lim, L. L., and Zlotnik, H. (eds.) (1992). *International migration systems: A global approach*. Clarendon Press, Oxford.

Kupiszewska, D., and Kupiszewski, M. (2005). *A revision of a traditional multiregional model to better capture international migration: The MULTIPOLES model and its applications*. CEFMR Working Paper 10/2005. Central European Forum for Migration Research, Warsaw.

Kupiszewska, D., and Nowok, B. (2005). *Comparability of statistics on international migration flows in the European Union*. CEFMR Working Paper 7/2005. Central European Forum for Migration Research, Warsaw.

Kupiszewska, D. (2006). International migration from new to old Member States in the year of EU enlargement. Submitted to the *Statistics in Focus* series of the Eurostat.

Kupiszewski, M. (1998). Czy na podstawie doświadczeń innych krajów można przewidywać zmiany migracji międzynarodowych po przyjęciu Polski do Unii Europejskiej? [Can the experience of other countries help forecast international migration changes after Poland's accession to the EU?]. In: P. Korcelli (ed.), *Przemiany w zakresie migracji ludności jako konsekwencja przystąpienia Polski do Unii Europejskiej* [*Changes in migration as a consequence of Poland's accession to the European Union*]. Bulletin of the Committee of Spatial Organization of the Polish Academy of Sciences, no. 184, Warsaw: 67–78.

Kupiszewski, M. (2001). Demograficzne aspekty wybranych prognoz migracji zagranicznych [Demographic aspects of selected forecasts of international migration]. In: A. Stępniak (ed.), *Swobodny przepływ pracowników w kontekście wejścia Polski do Unii Europejskiej* [*Free flow of workers in the context of Poland's accession to the European Union*]. Office of the Committee of European Integration, Warsaw: 73–98.

Kupiszewski, M. (2002a). How trustworthy are forecasts of international migration between Poland and the European Union? *Journal of Ethnic and Migration Studies*, 28(4): 627–645.

Kupiszewski, M. (2002b). *Modelowanie dynamiki przemian ludności w warunkach wzrostu znaczenia migracji międzynarodowych* [*The role of international migration in the modelling of population dynamics*]. Institute of Geography and Spatial Organisation, Polish Academy of Sciences, Warsaw.

Kupiszewski, M. (2005). *Migration in Poland in the Period of Transition – The Adjustment to the Labour Market Change*. Discussion Paper No. 266, Project on Intergenerational Equity. Institute of Economic Research, Hitotsubashi University, Tokyo.

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, Luxembourg, 15–16 July.

Layard, R., Blanchard, O., Dornbusch, R., and Krugman, P. (1992). *East-West Migration: The Alternatives*. MIT Press, Cambridge, MA.

Lee, E. S. (1966). A Theory of Migration. *Demography*, 3(1): 47–57.

Lee, R. D. (1998). Probabilistic Approaches to Population Forecasting. *Population and Development Review*, 24(Suppl.): 156–190.

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

Lewis, A. W. (1954). *Economic development with unlimited supplies of labour*. School of Economic and Social Studies, Manchester.

Lindsay, J., and Barr, B. M. (1972). Two Stochastic Approaches to Migration: Comparison of Monte Carlo Simulation and Markov Chain Models. *Geografiska Annaler B*, 54(1): 56–67.

Litterman, R. B. (1979). *Techniques of Forecasting Using Vector Autoregressions*. Working Paper No. 115. Federal Reserve Bank of Minneapolis, Minneapolis, MN.

Lowry, I. S. (1966). *Migration and Metropolitan Growth: Two Analytical Models*. Chandler, San Francisco.

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

Lutz, W., Sanderson, W. C., and Scherbov, S. (1997). Doubling of World Population Unlikely. *Nature*, 387(6635): 803–805.

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

Lutz, W., Saariluoma, P., Sanderson, W., and Scherbov, S. (2000). *New Developments in the Methodology of Expert- and Argument-Based Probabilistic Population Forecasting*. IIASA Interim Report IR-00-20. International Institute for Applied Systems Analysis, Laxenburg.

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

Mabogunje, A. L. (1970). Systems Approach to a Theory of Rural-Urban Migration. *Geographical Analysis*, 2(1): 1–18.

Markov, A. A. (1906). Rasprostranenie zakona bol'ših čisel na veličiny, zavisjaščie drug ot druga [Extension of the law of large numbers to dependent variables]. *Izvestija Fiziko-matematičeskogo obščestva pri Kazanskom universitete*, II Series, 15: 135–156.

Massey, D. S. (1987). The Ethnosurvey in Theory and Practice. *International Migration Review*, 21(4): 1498–1522.

Massey, D. S. (1990). Social structure, household strategies, and the cumulative causation of migration. *Population Index*, 56: 3–26.

Massey, D. S., Arango, J., Hugo, G., Kouaouci, A., Pellegrino, A., and Taylor, J. E. (1993). Theories of International Migration: Review and Appraisal. *Population and Development Review*, 19(3): 431–466.

Massey, D. S., and Zenteno, R. M. (1999). The dynamics of mass migration. *Proceedings of the National Academy of Sciences of the U.S.A.*, 96(9): 5328–5335.

Massey, D. S. (2002). A Synthetic Theory of International Migration. In: V. Iontsev (ed.), *World in the Mirror of International Migration*. MAX Press, Moscow: 142–152.

Mazurkiewicz, L. (1986). *Teoretyczne podstawy modeli przestrzennego oddziaływania* [*Theoretical foundations of the models of spatial interactions*]. Institute of Geography and Spatial Organisation, Polish Academy of Sciences. Ossolineum, Wrocław.

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

Milne, W. J. (1993). Macroeconomic Influences on Migration. *Regional Studies*, 27(4): 365–373.

Mincer, J. (1978). Family Migration Decisions. *Journal of Political Economy*, 86(5): 749–773.

Morrison, P. A. (1973). Theoretical issues in the design of population mobility models. *Environment and Planning A*, 5(1): 125–134.

Mulder, C. H. (1993). *Migration Dynamics: A Life Course Approach*. Thesis Publishers, Amsterdam.

Myrdal, G. (1957). *Rich lands and poor*. Harper and Row, New York.

Nakamura, T. (1986). Bayesian Cohort Models for General Cohort Table Analyses. *Annals of the Institute of Statistical Mathematics*, 38: 353–370.

Nowok, B., Kupiszewska, D., and Poulain, M. (2006). Statistics on international migration flows. In: M. Poulain, N. Perrin, and A. Singleton, (eds.), *Towards Harmonised European Statistics on International Migration*. Presses Universitaires de Louvain, Louvain-la-Neuve: 203–231.

NRC [National Research Council] (2000). *Beyond Six Billion. Forecasting the World's Population*. National Academies Press, Washington, D.C.

Öberg, S., and Wils, A. B. (1992). East-West migration in Europe. Can migration theories help estimate the numbers. *Popnet*, 22: 1–7.

Öberg, S. (1996). Spatial and Economic Factors in Future South-North Migration. In: W. Lutz (ed.), *The Future Population of the World: What Can We Assume Today?* Earthscan, London: 336–357.

Okun, B. (1968). Interstate Population Migration and State Income Inequality: A Simultaneous Equation Approach. *Economic Development and Cultural Change*, 16(2): 297–313.

Orłowski, W. (2000). Migration from central and eastern European countries after the accession: Effects for regions, labour markets, and social security systems. *Research Bulletin*, 2(9), Central Statistical Office, Warsaw: 7–21.

Paradysz, J. (2004). Marek Kupiszewski – Modelowanie dynamiki przemian ludności w warunkach wzrostu znaczenia migracji międzynarodowych, Wydawnictwo PAN, IGiPZ, Prace Geograficzne nr 181, Warszawa, 2002, ss. 174 (A Review). *Studia Demograficzne*, 145(1): 129–133.

Petty, W. (1682). *An Essay concerning the Multiplication of Mankind, Together with another Essay in Political Arithmetick concerning the Growth of the City of London with the Causes and Consequences Thereof*. London. Reprinted within the framework of the Gutenberg Project, «www.gutenberg.org/etext/5619» (accessed on 15 March 2006).

Piore, M. J. (1979). *Birds of passage: Migrant labour in industrial societies*. Cambridge University Press, Cambridge.

Plane, D. A. (1993). Demographic Influences on Migration. *Regional Studies*, 27(4): 375–383.

Plane, D. A., Rogerson, P. A. (1985). Economic-Demographic Models for Forecasting Interregional Migration. *Environment and Planning A*, 17(2): 185–198.

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

Pries, L. (1999). *Migration and Transnational Social Spaces*. Ashgate, Aldershot.

Quinn, M. A. (2006). Relative Deprivation, Wage Differentials and Mexican Migration. *Review of Development Economics*, 10(1): 135–153.

Ravenstein, E. G. (1885). The laws of migration. *Journal of the Royal Statistical Society*, 48(2): 167–227.

Ravenstein, E. G. (1889). The laws of migration. *Journal of the Royal Statistical Society*, 52(2): 214–301.

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

Rees, P. H., and Willekens, F. (1986). Data and accounts. In: A. Rogers, and F. Willekens (eds.), *Migration and Settlement. A Multiregional Comparative Study*. Reidel, Dordrecht: 15–59.

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*. John Wiley, New York.

Rogers, A. (1980). Introduction to Multistate Mathematical Demography. *Environment and Planning A*, 12(5): 489–498.

Rogerson, P. A. (1984). New Directions in Modelling of Interregional Migration. *Economic Geography*, 60(2): 111–121.

Salt, J., and Singleton, A. (1995). *Analysis and forecasting of international migration by major groups*. Report to the Working Party on Demographic Projections. Eurostat, Luxembourg.

Schmid, V., and Knorr-Held, L. (2001). *BAMP: Bayesian Age-Period-Cohort Modelling and Prediction*. Software documentation. Department of Statistics, Ludwig-Maximilian-University, Munich.

Sims, C. A. (1980). Macroeconomics and Reality. *Econometrica*, 48(1): 1–48.

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

Sjaastad, L. A. (1962). The costs and returns of human migration. *Journal of Political Economy*, 70(5): 80–93.

Spilerman, S. (1972). The Analysis of Mobility Process by the Introduction of Independent Variables into a Markov Chain. *American Sociological Review*, 37(3): 277–294.

Stark, O., and Bloom, D. E. (1985). The new economics of labor migration. *American Economic Review*, 75(2): 173–178.

Stark, O., and Taylor, E. (1989). Relative deprivation and international migration, *Demography*, 26(1): 1–14.

Stark, O. (1991). *The migration of labor.* Basil Blackwell, Cambridge, MA.

Stark, O., and Wang, Y. (2001). *Inducing Human Capital Formation: Migration as a Substitute for Subsidies*. Economics Series No. 100. Institute for Advanced Studies, Vienna.

Stark, O. (2003). *Tales of Migration without Wage Differentials: Individual, Family and Community Contexts*. ZEF Discussion Paper on Development Policy No. 73. Zentrum für Entwicklungsforschung, Bonn.

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

Stillwell, J. C. H., and Congdon, P. (eds.) (1991). *Migration models: macro and micro approaches*. Belhaven Press, London.

Stillwell, J. C. H., Eyre, H., and Rees, P. H. (1999). *Regional international migration and interregional migration within the European Union: a feasibility study*. Research report for the ERDF Study 98/00/27/174. University of Leeds, Leeds.

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

Stouffer, S. A. (1960). Intervening opportunities and competing migrants. *Journal of Regional Studies*, 2(1): 187–208.

Taylor, J. E. (1986). Differential migration, networks, information and risk. In: O. Stark (ed.), *Research in Human Capital and Development*, Vol. 4: *Migration, Human Capital, and Development*. JAI Press, Greenwich, CT.

Törnqvist, L. (1949). Näkökohdat, jotka ovat määränneet primääristen prognoosiolettamusten valinnan [Points of view that have determined the primary forecast assumptions]. In: J. Hyppölä, A. Tunkelo, and L. Törnqvist (eds.), *Suomen väestöä, sen uusiutumista ja tulevaa kehitystä koskevia laskelmia* [*Calculations Concerning the Population of Finland, Its Renewal and Future Development*]. Tilastollisia tiedonantoja 38. Statistics Finland, Helsinki.

Tuljapurkar, S. (1997). Demography. Taking the measure of uncertainty. *Nature*, 387(6635): 760–761.

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

Tuljapurkar, S., Lee, R. D., and Li, Q. (2004). Random Scenario Forecasts Versus Stochastic Forecasts. *International Statistical Review*, 72(2): 185–199.

United Nations (1970). *Methods of Measuring Internal Migration*. United Nations, New York.

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

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

van der Gaag, N., de Beer, J., and Willekens, F. (2005). *MicMac. Combining micro and macro approaches in demographic forecasting*. Paper for the Joint UNECE-Eurostat Work Session on Demographic Projections, Vienna, 21–23 September.

van Imhoff, E. (1990). The exponential multidimensional demographic projection model. *Mathematical Population Studies*, 2(3): 171–182.

van Imhoff, E., and Keilman, N. (1991). *LIPRO 2.0: An application of a dynamic demographic projection model to household structure in The Netherlands*. NIDI & CBGS Publication 23. Swets & Zeitlinger, Amsterdam.

van Imhoff, E., and Post, W. (1998). Microsimulation methods for population projection. *Population–E*, 10(1): 97–138.

Veblen, T. (1898). Why is Economics not an Evolutionary Science? *Quarterly Journal of Economics*, 12(4): 373–397.

Vishnevsky, A., and Zayonchkovskaya, Zh. [Višnevskij, A. G., and Zajončkovskaja, Ž. A.] (1994). Emigration from the former Soviet Union: the fourth wave. In: H. Fassmann, R.

Münz (eds.), *European Migration in the Late Twentieth Century: Historical Patterns, Actual Trends and Social Implications*. Edward Elgar, Aldershot.

Wallerstein, I. (1974). *The Modern World-System*, Vol. I: *Capitalist Agriculture and the Origins of European World-Economy in the Sixteenth Century*. Academic Press, New York.

Wegren, S., and Cooper Drury, A. (2001). Patterns of Internal Migration During the Russian Transition. *Journal of Communist Studies and Transition Politics*, 17(4): 15–42.

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

Willekens, F., and Baydar, N. (1986). Forecasting place-to-place migration with generalized linear models. In: R. Woods, and P. H. Rees (eds.), *Population structures and models*. Allen and Unwin, London: 203–244.

Willekens, F. (1994). Monitoring international migration flows in Europe. Towards a statistical data base combining data from different sources. *European Journal of Population*, 10(1): 1–42.

Willekens, F. (1995). *Multiregional demographic forecasting*. Population Research Centre Working Paper 95-1, University of Groningen, Groningen.

Willekens, F. (2005). *Models of migration: Observations and judgments*. Paper for the Workshop on the Estimation of International Migration in Europe: Issues, Models and Assessment, Southampton, UK, 28–30 September.

Wilson, A. G. (1967). A statistical theory of spatial distribution models. *Transportation Research*, 1: 253–269.

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

Wilson, A. G. (1981). *Catastrophe Theory and Bifurcation. Applications to Urban and Regional Systems*. Croom Helm, London.

Zelinsky, W. (1971). The hypothesis of the mobility transition. *Geographical Review*, 61(2): 219–249.

Zipf, G. K. (1946). The P_1P_2/D Hypothesis: On the Intercity Movement of Persons. *American Sociological Review*, 11(6): 677–686.

Zlotnik, H. (1998). *The theories of international migration*. Paper for the Conference on International Migration: Challenges for European Populations, Bari, Italy, 25–27 June.

Zolberg, A. R (1989). The Next Waves: Migration Theory for a Changing World. *International Migration Review*, 23(3): 403–430.

■

CEFMR Working Papers Series

1/2003: M. Kupiszewski, *Consequences of EU enlargement for freedom of movement between Council of Europe Member States*, ISBN 83-920313-0-X.

2/2003: M. Kupiszewski, D. Kupiszewska, *Internal migration component in subnational population projections in Member States of the European Union*, ISBN 83-920313-1-8.

3/2003: M. Kupiszewski, J. Bijak, K. Saczuk, R. Serek, *Komentarz do założeń prognozy ludności na lata 2003–2030 przygotowywanej przez GUS*, ISBN 83-920313-2-6.

4/2003: K. Saczuk, *A development and critique of the concept of replacement migration*, ISBN 83-920313-3-4.

5/2003: I. Koryś, *Migration Trends in Selected EU Applicant Countries: Poland*, ISBN 83-920313-4-2.

1/2004: A. Kicinger, K. Saczuk, *Migration policy in the European perspective – development and future trends*, ISBN 83-920313-5-0.

2/2004: A. Kicinger, *International migration as a non-traditional security threat and the EU responses to this phenomenon*, ISBN 83-920313-6-9.

3/2004: J. Bijak, *Fertility and mortality scenarios for 27 European countries, 2002–2052*, ISBN 83-920313-7-7.

4/2004: J. Bijak, M. Kupiszewski, A. Kicinger, *International migration scenarios for 27 European countries, 2002–2052*, ISBN 83-920313-8-5.

5/2004: K. Saczuk, *Labour force participation scenarios for 27 European countries, 2002–2052*, with Annex, ISBN 83-920313-9-3.

1/2005: J. Bijak, D. Kupiszewska, M. Kupiszewski, K. Saczuk, *Impact of international migration on population dynamics and labour force resources in Europe*, with Annex, ISBN 83-921915-0-1.

2/2005: A. Kicinger, *Unia Europejska wobec zagadnienia integracji imigrantów*, ISBN 83-921915-1-X.

3/2005: I. Koryś, *Dimensions of Integration: Migrant Youth in Poland*, ISBN 83-921915-2-8.

4/2005: A. Kicinger, *Polityka emigracyjna II Rzeczpospolitej*, ISBN 83-921915-3-6.

5/2005: B. Nowok, D. Kupiszewska, *Official European statistics on international migration flows: availability, sources and coverage*, ISBN 83-921915-4-4.

6/2005: J. Bijak, *Bayesian methods in international migration forecasting*, ISBN 83-921915-5-2.

7/2005: D. Kupiszewska, B. Nowok, *Comparability of statistics on international migration flows in the European Union*, ISBN 83-921915-6-0.

8/2005: B. Nowok, *Evolution of international migration statistics in selected Central European countries*, ISBN 83-921915-7-9.

9/2005: A. Kicinger, *Between Polish interests and the EU influence – Polish migration policy development 1989–2004*, ISBN 83-921915-8-7.

10/2005: D. Kupiszewska, M. Kupiszewski, *A revision of a traditional multiregional model to better capture international migration: The MULTIPOLES model and its applications*, ISBN 83-921915-9-5.

1/2006: M. Kupiszewski, J. Bijak, *Ocena prognozy ludności GUS 2003 z punktu widzenia aglomeracji warszawskiej*, ISBN 83-60462-00-3.

2/2006: M. Mazur-Rafał, *Zmiana paradygmatu w niemieckiej polityce imigracyjnej w latach 1998–2004? Wnioski dla Polski*, ISBN 83-60462-01-1.

3/2006: J. Bijak, I. Koryś, *Statistics or reality? International migration in Poland*, ISBN 83-60462-02-X.

4/2006: J. Bijak, *Forecasting international migration: Selected theories, models, and methods*, ISBN 83-60462-03-8

All CEFMR Working Papers are available at: «www.cefmr.pan.pl».

ISSN 1732-0631
ISBN 83-60462-03-8

Printed in Poland