Version as of April 15, 2009

POPULATION ASSOCIATION OF AMERICA 2009 ANNUAL MEETING
Detroit, Michigan April 30 to May 2, 2009
Session 59: Improved Methods for Demographic Data Collection and Analysis

Reconciliation of various migration measures: insights from microsimulation of origin-destination specific flows

Beata Nowok[*] *and Frans Willekens*
Netherlands Interdisciplinary Demographic Institute (NIDI), The Hague

Abstract: The conceptual and measurement complexity of migration stems from the fact that people move between places in a time-space continuum. The numerous possibilities to discretize temporal dimension of movements lead to diversity of the employed measures of migration flows. They differ not only across countries, but also over time and between sources within the same country. The aim of the paper is to reconcile different operational measures of origin-destination specific migration flows. We focus on event approach to measuring migration and consider various time-related constraints (e.g. duration threshold of presence and absence in a country). Given the availability and quality problems of migration data, the issue is tackled using a continuous-time microsimulation. We generate individual migration histories of persons moving among a closed system of three areas. Different migration measures are derived and compared for the whole virtual population.

1. Introduction

At first glance, migration seems to be a simple concept and in most studies on the subject the meaning of this term is assumed to be known. Nonetheless the definitional and measurement complexity of migration becomes evident once a precise description is attempted (Boyle et al. 1998, Courgeau 1993, Courgeau 2006, Rees 1977, Rogers et al. 2003, Willekens 1982, Willekens 1985). It stems from the fact that people move between places in a time-space continuum. In order to measure migration all types of movements have to be separated into the dichotomous categories of migration and non-migration. It

[*] nowok@nidi.nl

is necessary to sharpen the migration definition and discretize temporal and spatial dimensions. As universally valid time and space framework does not exist, there are numerous differences in how migration is observed and measured. The linkage between different measures of migration is still an underinvestigated research question. Nonetheless, it is of crucial importance for international comparison of migration patterns, combing data from different national and international sources, and harmonization of data over time. It plays a significant role when missing migration components are estimated from alternative sources, or when statistics have to be transformed into different reference intervals.

Most of the few existing studies that address the issue of relation between different migration measures focus on data derived from a census question on place of usual residence at a specified date in the past. This data type represents a transition approach to measuring migration. The authors have looked into the problem of reconciling data referring to time intervals of various lengths, e.g. one- and five-year period (Kitsul and Philipov 1981, Kraly and Warren 1992, Liaw 1984, Long and Boertlein 1990, Newbold 2001, Rees 1977, Rogers et al. 2003, pp. 2-6, Rogerson 1990, p. 5). In Europe, however, most of the countries derive their migration statistics from population registers and apply an event approach to measuring migration (Kupiszewska and Nowok 2008, Nowok et al. 2006). Register based measures refer to changes of country of usual residence for a specified duration. The main definitional difference between those measures refers to a minimum length of stay in the country that constitutes migration. In view of obtaining an overall and consistent picture of the migration patterns occurring within Europe, some studies have been carried out based on origin-destination-specific flows as reported by sending and receiving countries. In general, they try to use information included in good quality data to correct inadequate ones (and estimate missing values). Poulain and Dal (2008) developed a correction factor method, which enables to produce a unique figure for each migration flow between pairs of countries in the system. The values of the correction factors indicate the level of discrepancies between figures reported by different countries. Another approach, applied e.g. by Raymer (2007), is based on log-linear model that focuses on the underlying structures found in origin-destination flows. Recently an attempt has been made to combine both of the above mentioned approaches (Raymer and Abel 2008). Note that these methods do not provide the answers about the linkage of one measure of migration to another. They rely on statistics that result from different data production processes. These statistics inherit all possible types of deficiency present at each stage of data collection and processing. The definitional problems constitute only a part of it and the final estimates follow definition that can be specified only in an approximate way.

In this paper an attempt is made to create a framework for reconciling event-type migration measures that differ in terms of time criteria. An inherent spatial aspect of migration is, however, not put aside. Given the availability and quality problems of migration data, the issue is tackled using microsimulation approach. The origin-destination migration histories of individuals are projected in continuous-time by randomly drawing their trajectories from a statistical model. Continuous-time microsimulation allows application of a whole range of migration definitions to the simulated data. Different migration measures are derived for the whole virtual population. If there is a close match between the model and the true probabilistic mechanism underlying individual mobility then the link between migration measures for the virtual population should be transferable to the real populations. Measures on migration flows are presented from international perspective, but note that there are no essential conceptual and measurement differences between internal and international migration. Moreover, some of the measurement problems are present also in other fields, where durations of episodes proceeding and following an event are of crucial importance.

In the next section a time-space perspective of international migration is set out. Then an example of individual migration history is used to present various operationalizations of time in migration measures. It is followed by description of microsimulation model and presentation of the results.

2. Time and space

Migration forms a small subset of the total movements of population occurring in time-space continuum. Consequently, time and space constitute crucial defining features of migration. Our primary interest is in international migration, for which the geographical criterion is straightforward. Migration requires the crossing of international border to be considered as international. The temporal dimension, however, poses a significant challenge. The complexities stem from the temporal character of most movements. A change of usual country of residence, even across international frontiers, is no longer the once-in-a-lifetime event. People change residence repetitively for various durations and only few residential moves are genuinely permanent. Moreover, for an increasing number of people it is difficult to determine their current country of usual residence, as they distribute their time across a set of places.

Nonetheless, there must be some permanence to a movement classified as migration (Boyle et al. 1998). The disagreement centres on the minimum duration of residence that constitutes migration. Note that potential migration is usually interrupted by other forms of temporary mobility, which introduces an ambiguity into the minimum length of residence criterion. Spatial supranational perspective on international temporary movements increases the complexity of the time-related issues. Each move is observed by two countries, namely country of origin and country of destination. For the former it may satisfy requirements to be counted as emigration and for the latter as immigration. Emigration from a sending country, however, does not have to be immigration to a receiving country. It is due to the fact that the first move can be followed by an onward one and an individual is absent from an origin country for a lasting duration, but not present long enough in a destination one.

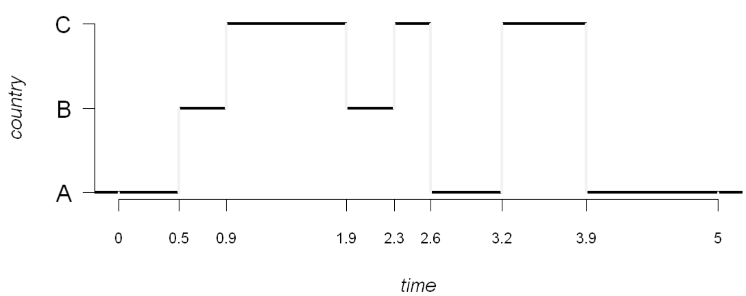
3. Measures of migration: from biographies to statistics

In order to produce statistics migration has to be defined in such a way that it can be practically measured. The fuzzy concept of usual place of residence, which serves as a focal point in migration definition, needs to be operationalized and put in a discrete time framework. Information concerning the duration of time spent in and out of a country is therefore generally taken, directly or indirectly, into account. The variety of different measures that results is much larger that it seems at first sight. Moreover, the scarce metadata that are provided with migration statistics usually do not allow to disentangle the details of definition underlying the data. The key issue is what the potential distortions of data comparability caused by the unrevealed details are and whether they matter. Needless to say, they depend on a threshold of a minimum duration of residence, which is used in migration definition.

To tackle these questions we start with a presentation of possible meanings of a duration component in migration definition. In this connection, it is vitally necessary to provide clarification of the concepts used throughout the paper. We treat a relocation event as a starting and reference point. We assume that relocation is an unambiguously defined event that occurs at a specific point in time. It is defined as a change of residence (address) involving the crossing of national boundary. Thus, relocation is characterized by three variables: date of relocation, country of origin and country of destination. Other characteristics, such as the reason for relocation or nationality of the migrant may be added. There are no duration restrictions connected with a relocation event, as opposed to an event that is referred to as migration. The duration criteria come as conditions for what happens after and possibly before a relocation event. A resulting migration event is a composite event. Thus, to undertake a migration an individual has to experience a particular history over a certain period of time, rather than an event at a specific point in time. Note that migration with minimum duration threshold equal to zero reduces to relocation. Duration refers to presence – time spent continuously in a reference country, or absence – time spent continuously out of a reference country. Episode of presence or absence denotes a continuous uninterrupted stay in one country, in other words time between subsequent relocations. Absence from a reference country may be connected with stays in different countries and refer to a sequence of episodes rather than a single one. Next to duration, a country of stay is therefore, an additional condition that may be imposed on absence. An episode of presence in a reference country exceeding minimum duration of stay constitutes a usual residence in this country, which is then not canceled by episodes of absence for period shorter than the cut-off duration. The duration of residence in a country may refer to an uninterrupted stay in that country or may refer to a sequence of stays interrupted by short periods of absence. Since migration, similarly as relocation, is characterized by origin and destination country a location of observer may be situated in one of them. A country where an observer is located is a country of reference. From the perspective of an origin country migration is seen as emigration and from the perspective of a destination country as immigration.

For illustrative purpose we use an example relocation career of an individual distributing time between different countries. It is to show how this relocation path may be viewed from the perspective of those countries and how it contributes to different origin-destination specific migration measures that are presented. Consider a relocation history of an individual within an observation period of five years, $\omega[0,5]$. Assume that a person moves between three countries, $S=\{A,B,C\}$. An example relocation path is presented in Figure 1.

Figure 1 Example relocation path



We can describe this sample path in a compact way including timings of relocation and places of residence as proposed by Tuma and Hannan (1984)

$$\omega[0,5]=\{0, A, 0.5, B, 0.9, C, 1.9, B, 2.3, C, 2.6, A, 3.2, C, 3.9, A\}. \qquad (1)$$

Alternatively, places of residence and timings of relocations can be presented as two separate vectors. At the onset of the observation, time *0,* a person is living in *A,* at time *0.5* moves to *B* and stays there till *0.9*, when he or she relocates to *C* and so on. Finally at *3.9* a person comes back to *A* and stays there until the end of an observation period, time *5*. Thus, (1) includes a complete history of state occupancies and times of changes. Equivalently it can be also presented in terms of waiting times between subsequent relocations

$$\omega[0,5] = \{0, A, 0.5, B, 0.4, C, 1, B, 0.4, C, 0.3, B, 0.6, C, 0.7, A\}, \qquad (2)$$

which is more convenient when duration of episodes is of primary interest. Note that in practice this relocation path (1) is observed from three different country specific perspectives, which leads to three different incomplete trajectories:

$$\omega_A[0,5] = \{0, A, 0.5, B.........................C, 2.6, A, 3.2, C, 3.9, A\},$$
$$\omega_B[0,5] = \{....A, 0.5, B, 0.9, C, 1.9, B, 2.3, C..............................\},$$
$$\omega_C[0,5] = \{............B, 0.9, C, 1.9, B, 2.3, C, 2.6, A, 3.2, C, 3.9, A...\}.$$

Whether the country perspective leads to discrepancies between migration figures recorded for the same flow by different countries depends on the operational definition of migration. Consider first the *relocation* measure without duration conditions imposed. There are no discrepancies between the origin-destination specific data that are produced by different countries based on relocations. For instance, a number of out-moves from *A* to *B* recorded by an origin country *A* is consistent with a number of in-moves to *B* from *A* reported by a destination country *B.* An analogous situation exists when duration conditions are formulated in terms of duration of presence and absence due to stay in a country of origin or destination. A duration condition refers therefore to episode following relocation only or both following and preceding relocation. The corresponding definitions are phrased as follows: (I a) relocation that is followed by a stay of specified duration in a destination country; (I b) relocation that is followed and preceded by a stay of specified duration in a destination and origin country respectively.

Operationalizations that consist in attaching duration conditions referring to presence and absence in the country without any additional restriction on country of stay leads inevitably to discrepancies between figures produced by different countries. It stems from the fact that absence from the country that is long enough to be treated as emigration may be connected with a few short term stays (episodes) in other countries, neither of which qualifies as migration. The country-specific definitions of migration are as follows: (II a) relocation that is followed by a presence/absence of specified duration; (II b) relocation that is followed by a presence/absence of specified duration and proceeded directly by an absence/presence for a specified duration. The distinction between presence and absence is drawn to indicate the difference between immigration and emigration. Note that in case of immigration, figures following definition (II a) are equal to those following definition (I a).

So far we have presented definitions in which continuous uninterrupted presence or absence directly following and preceding relocation were considered. In order to disregard relocations for short-lasting periods migration can be defined as (III) relocation that is followed by a presence/absence of a specified duration, provided a person is a non-resident/resident of a destination/origin country. A residence is established by a continuous presence of a minimum duration and cancelled by a continuous absence of a minimum duration. Absences from a country of residence for shorter duration do not entail cancelation of residence in this country. Thus a person coming back after a short duration spent abroad is not a migrant.

The duration variables associated with the relocations from our example are shown in Table 1 together with the resulting contributions to different origin-destination migration measures produced by different reference countries (see part (a) and part (b) of the table respectively; compare also with Figure 1). The same minimum duration threshold of six months (*0.5* year) is applied by all countries. The individual in question experiences seven relocations at dates indicated in column 2. Consider the first relocation (line 1, unless otherwise indicated) directed from country A to country B (A:B) and taking place at time *0.5* ((a) column 2). Before relocation a person spends half a year in A ((a) column 5) and then moves to B, where he or she spends less than five months ((a) column 4). Thus, duration of stay (episode) following relocation is shorter than the minimum threshold of six months and this relocation is not a migration according to measure (I a) and (I b) ((b) column 2 and 3). Nonetheless, he or she relocates later on between B and C and comes back to A only at time *2.6* ((a) column 2). A person is therefore absent from A after his or her first observed relocation for *2.1* year ((a) column 6: *2.1,* based on column 2, line 5 and 1 respectively*: 2.6-0.5=2.1*), which makes relocation qualify as migration according to measure (II a) and (II b), but only from the perspective of country A ((b) column 4 and 7). The duration of presence in country B is too short to consider this relocation as migration ((b) column 5 and 8).

Table 1 Individual relocation history

(a) duration variables (in years)

No.	Relocation		Duration of stay in origin and destination country		Duration of presence in/absence from a reference country					
	Date	Direction from:to			A		B		C	
			after	before	after	before	after	before	after	before
[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]
1	0.5	A:B	0.4	0.5	2.1	0.5	0.4	Inf		
2	0.9	B:C	1.0	0.4			1.0	0.4	1.0	Inf
3	1.9	C:B	0.4	1.0			0.4	1.0	0.4	1.0
4	2.3	B:C	0.3	0.4			2.7	0.4	0.3	0.4
5	2.6	C:A	0.6	0.3	0.6	2.1			0.6	0.3
6	3.2	A:C	0.7	0.6	0.7	0.6			0.7	0.6
7	3.9	C:A	1.1	0.7	1.1	0.7			1.1	0.7

(b) contribution to different origin-destination migration measures; minimum duration of stay threshold amount to six months (0.5)

No.	(I a)	(I b)	(II a)			(II b)			(III)		
			A	B	C	A	B	C	A	B	C
[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]
1	-	-	A:B	-		A:B	-		A:B	-	
2	B:C	-		B:C	B:C		-	B:C		-	B:C
3	-	-		-	-		-	-		-	-
4	-	-		B:C	-		-	-		-	-
5	C:A	-	C:A		C:A	C:A		-	C:A		C:A
6	A:C	A:C	A:C		A:C	A:C		A:C	A:C		A:C
7	C:A	C:A	C:A		C:A	C:A		C:A	C:A		C:A

NB: Empty cells indicate relocations that are not observed, i.e. relocations between the other two countries;
- = relocation that do not qualify as migration; Inf = individual enters the country for the first time

In the presented example the discrepancies between different measures stems from different strictness of the imposed conditions and location of an observer. Nonetheless, these two factors do not always have to lead to disagreement in statistics. There are no differences between measures (II a), (II b), and (III) produced by country A (Table 1 (b) columns 4, 7 and 10), and also no differences between measures (IIa) and (III) produced by country A and C and referring to flows between them (Table 1 (b) columns 4, 6, 10 and 12). In practice, however, the discrepancies resulting from the location of an observer are greater than in the presented instance due to variability of cut-off duration.

Note that migration measure (III) is the only one that produces consistent origin-destination migration history. Nonetheless, as the knowledge about the relocations abroad is limited, an individual may leave a country to one direction and comes back from a different one (Table 1 (b) column 10, line 1 and 5). In addition, although an individual cannot be a resident of two countries at the same time, unless countries use a different minimum duration threshold, he or she may have no country of residence. For the considered 0.5 cut-off duration an individual in question does not have a country of residence in period *0.5-0.9*. In case of other measures the origin-destination characteristic of subsequent flows are far more problematic. It may happen, for instance, that a person migrates in the same direction two times in a row (see measure (IIa) according to country B, Table 1 (b), column 5, line 2 and 4).

In order to ensure consistent and complete trajectories of unique places of residence and changes thereof, a definition (III) based on lengths of presence and absence has to include additional condition on duration of stay in a destination country and this definition has to be used by all countries with the same value of duration threshold. Thus, migration is (IV) relocation that is followed by a continuous stay of specified duration in a destination country provided a person is a non-resident thereof. For consistency reason, the origin of the relocation must be a previous country of residence and not a previous country of a short stay that does not constitute a residence. In our example the first migration would be at time *0.9* from *A* (a previous country of residence), instead of B (a previous country of a short stay), to *C*.

In event approach presented above the operational definitions of migration describe rules for selecting migration from all relocations. In transition approach, characteristic for census and household surveys, the selection rules applies to population stocks. Basically only the immigration data are therefore available. They include persons with different place of usual residence at a specified date in the past. The duration criterion applies to population membership, e.g. only those are interviewed about their previous place of residence who have been already living in the country for at least one year or who are going to live in total for at least one year. Additional minimum duration criteria may be imposed on a residence at a reference date in the past. This type of measure is, however, beyond the scope of this paper.

It has to be emphasized that a migration definition recommended by the United Nations (1998) does not correspond precisely to any of the described measures. According to United Nations recommendations a migrant is characterized as “a person who moves to a country other than that of his or her usual residence for a period of at least a year (12 months)". The definition resembles measure (IV), but a stay in a destination country does not have to be continuous, which pose a serious problem of ascertaining a country of usual residence in some cases. In result additional time-limits of presence and absence have to be applied anyway. The measure (III) is equivalent to a definition used in a previous set of the United Nations recommendations adopted in 1976. A migrant was defined as "a person who has entered a country with the intention of remaining for more than one year and who either must never have been in that country continuously for more than one year or, having been in the country at least once continuously for more than one year, must have been away continuously for more than one year since the last stay of more than one year". It was perceived as logically impeccable (United Nations 1998, para. 22) and as such is preferred by us to the current one that is ambiguous.

4. Microsimulation of origin-destination flows

The purpose of the microsimulation model is to gain insights into the dependence of aggregate migration counts produced by different countries on various operationalizations of the concept in terms of time, which works at individual level. A microsimulation model is designed to produce origin-destination migration histories of individuals within specified time horizon, as presented in Section 3. We rely on the general notion that different migration measures represent the same continuous data generating process. They differ depending on how the data happened to be collected and how the statistics happened to be produced. In principle, the discrepancies between various measures can be expressed in terms of the parameters of the underlying process, but they become intractable when complexity of measures and processes increases. For a considered closed system of three countries we have six possible origin-destination specific flows. This system can be viewed as a multiregional model or a multistate model (Rogers 2008) with three potential regions of residence or states respectively. Since timing and duration is of crucial importance we implement a continuous-time approach. The timing of migrations, which can occur at any arbitrary moment, is described by a duration model. The assumed origin-destination instantaneous transition rates (hereinafter referred to also as hazard rates or intensities) determine when the next transition will occur and to which state. They are denoted by $\lambda_{ij}(t)$, where *i* is country of origin, *j* country of destination, and *t* process time (time since event-origin). In the simplest case the transition rates are the same for all individuals and constant over time.

The transition rates uniquely define distribution of random interarrival time. We can therefore simulate the random variates from the interarrival time distributions. To generate random duration *X* we can use the inverse transform method (Ross 2006). We take a random number *U* distributed uniformly on (0,1) and set $X=F^{-1}(U)$, where F^{-1} denote an inverse function. If migration intensities are assumed to be constant, then the times between successive relocations comes from the exponential distribution. We assume also that these times are independent and identically distributed. In our setting where migration may have several alternative destination states (competing risk setting) we have to determine not only time to movement, but also its direction. It can be done simultaneously by simulating the sojourn times to each possible destination, based on origin-destination specific intensities. The destination state is the one with the shortest simulated waiting time. It can be done alternatively by first simulating time to migration regardless of destination, using intensity of leaving current stay. Then additional random number is used to determine the direction based on the probability distribution of the possible destination states. We apply the former approach.

We generate the migration trajectories of a large number (6 000) of virtual individuals who start their migration experience in three different countries (2000 in each). Then we derive different migration measures presented in Section 3 for each of the countries for each year of the simulated period of time. We run simulation with purely theoretical input. Different assumptions are made on population heterogeneity, levels of origin-destination specific intensities of relocation, and dependence of these intensities on process time (e.g. a power dependence of the hazard on time for Weibull distribution).

Simulation and all calculations are carried out using the R statistical software package (http://www.r-project.org/), the GNU implementation of the S language.

5. Reconciling different migration measures

There are three main areas of definitional discrepancies between migration statistics we are confronted with. First, duration conditions may be imposed differently on presence and absence following and proceeding relocation, which determine a type of migration measure. Second, for a given measure a minimum duration criterion may be of different length. Third, a location of an observer may be placed in a different country of reference (country of origin vs. country of destination). All these aspects are described based on simulation results run under the following two sets of assumptions:

i. homogenous population, origin-destination specific intensities[†], intensities constant over time (exponential duration model);
ii. homogenous population, origin-destination specific intensities, intensities decreasing with time (Weibull duration model[‡]).

All results are shown as proportion of relocation that satisfy a particular duration criterion for a given measure. Relocation and migration numbers are calculated as a mean of annual values over ten years.

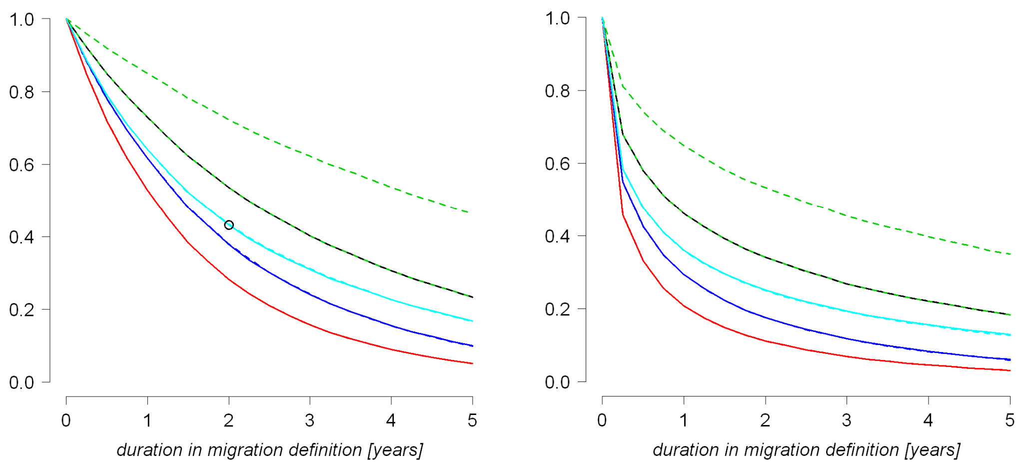
Consider first the total flows in the system and look at the impact of the duration threshold used in migration definition on the flow volume. Figure 2 shows ratios of total number of migrations in these three countries according to the measures introduced in Section 3 to a total relocation number. The total migration number is calculated from the perspective of origin countries only (emigration) or destination countries only (immigration), which will be dealt with later on. The ratios are presented for different length of duration criterion. For instance, under the initial assumptions (i), if the measure (III) is applied and the minimum duration criterion is set to two years than 43 % of all relocations that occur in the system are considered as migrations (see the marked point in Figure 2, part (a), left panel). For the assumed set of origin-destination intensities there is a substantial impact of duration threshold used in migration definition on the migration

[†] Results are presented for two sets of intensities. Initial set of intensities is as follows: λ_{ij}: λ_{12}=0.16, λ_{13}=0.04, λ_{21}=λ_{23}=0.15, λ_{31}=λ_{32}=0.25; additional one includes intensities that are ten times lower than the initial ones.
[‡] The hazard rates are monotone decreasing (shape parameter is 0.5) and reduce to the constant hazard rates used in exponential duration model when shape parameter is equal to unit.

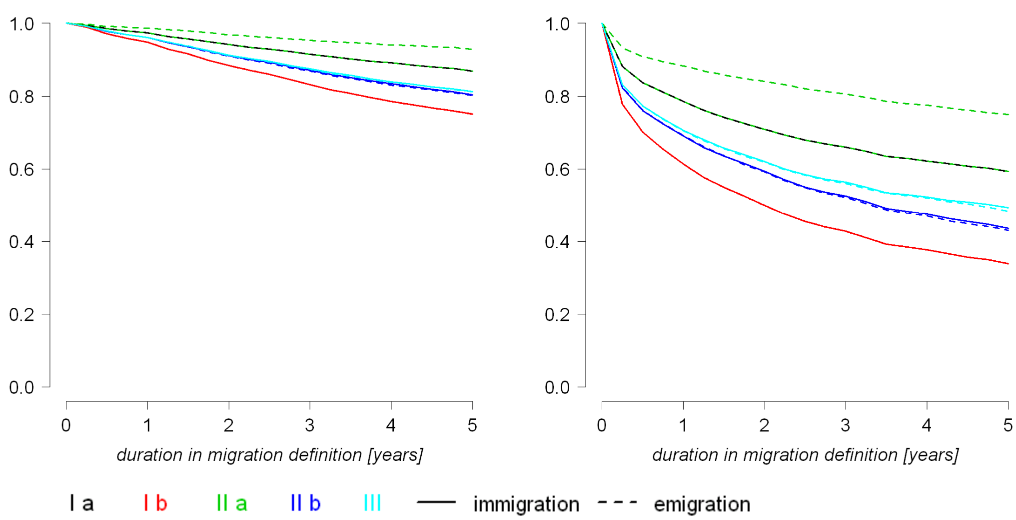
volume. It refers to all measures regardless of the shape of the hazard function (see Figure 2 (a)). A decrease in the values of intensities leads to smaller discrepancies, especially when the hazard rates are constant (see Figure 2 (b) for all intensities ten times lower than the initial ones). This result suggests that the length of duration criterion (and measurement differences as well) do not play an important role when the mobility level of the population is low. In the example with constant rates, imposing one year duration criterion on relocation results in decline in migration figures by 1-5 %.

Figure 2 Proportion of total relocation number in the system that satisfy a particular duration criterion; left panel: exponential duration model, right panel: Weibull duration model

(a) initial λ_{ij}



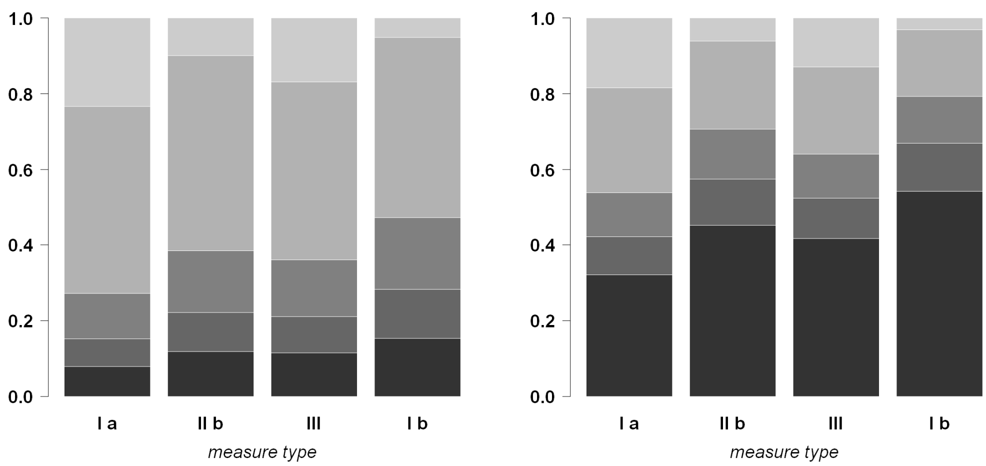
(b) $\lambda_{ij}/10$



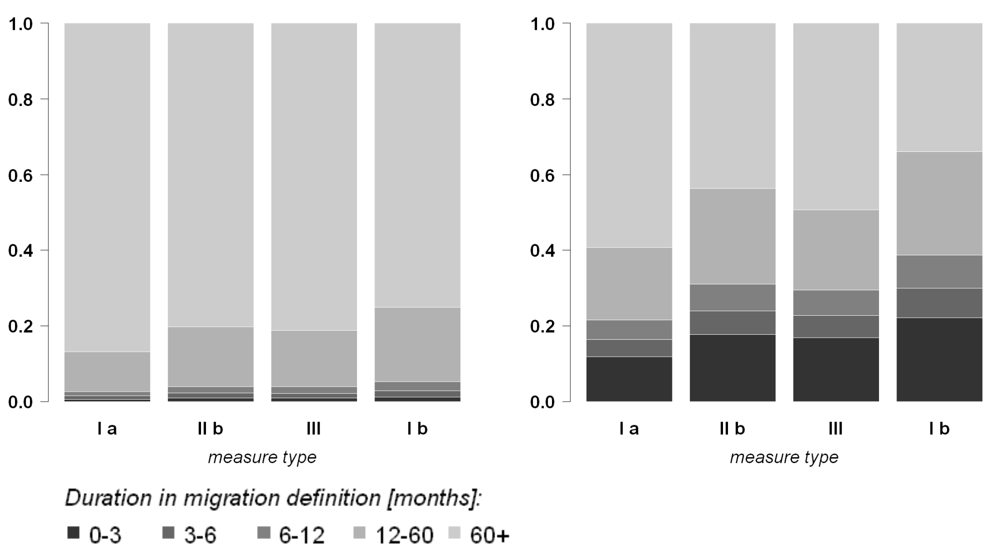
The low intensities mean basically that share of long-term relocations in the total number of relocations is larger. In other words, more persons relocate after an extended duration of stay, long enough to satisfy the duration of stay criterion. Hence, the smaller the migration intensity, the higher the proportion of relocations that fulfill a given duration of stay criterion. Shares of relocation numbers that fulfill duration criteria of different length for various measures are shown in Figure 3. The dominance of the long-term migrations is particularly clear for constant rates (Figure 3 (b), left panel).

Figure 3 Shares of relocation numbers that fulfill duration criteria of different length for various measures; left panel: exponential duration model, right panel: Weibull duration model

(a) initial λ_{ii}



(b) $\lambda_{ii}/10$

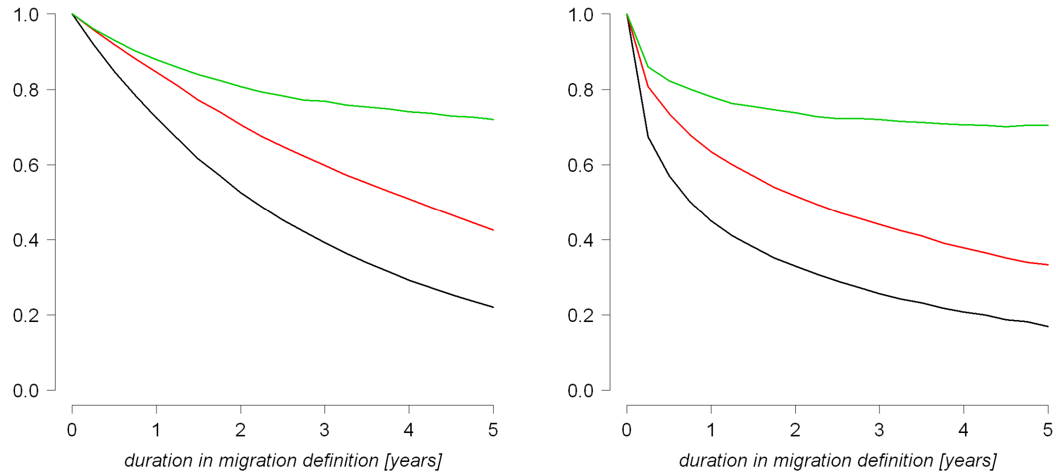


For intensity decreasing over time it is noteworthy that the differences between measures using short duration conditions are greater than for those using long ones, e.g. discrepancy between the relocation number and the number of migrations applying one year duration condition is wider than between the latter and the number of migrations applying five year duration condition. In addition, the dependence of intensity on time has a substantial impact on the shares of relocations that happen before and after minimum duration of stay. As relocations occurring before a minimum duration are not visible for the statistical systems, it has a crucial importance in practice.

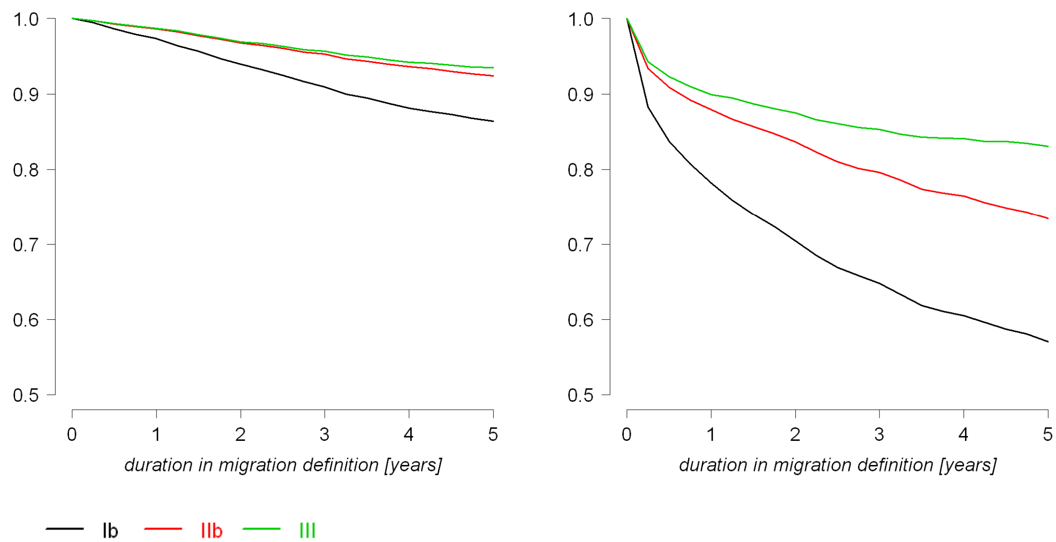
All migration measures introduced in section 3 can be seen as conditional relocation with different restrictiveness of the imposed condition. Needless to say, the longer the duration, the larger the differences between migration figures produced according to different definitions. Figure 4 shows ratio of different measures (I b, II b, III) to relocations that are followed by a stay of specified duration (I a) for duration condition up to five years. Whether one measure can be used as a good approximation of another one when modeling, depends on characteristics and parameters of an underlying relocation process. For example, measure (III) which follows a far more complex definition than measure (I a), could be approximated by the latter for duration longer than one year (see Figure 4). For the whole domain of duration threshold, measure (III) may be roughly estimated as an average of the most restrictive (II a) and the least restrictive (I a) measures (not shown in figure).

Figure 4 Ratio of different measures (I b, II b, III) to relocations that are followed by a stay of specified duration (I a); left panel: exponential duration model, right panel: Weibull duration model

(a) initial λ_{ii}



(b) $\lambda_{ii}/10$



Consider now impact of location of observer, namely a country of origin or a country of destination, on the number of migrations according to different measures. For most types of measures, the location perspective does not influence the total number of migrations in a system. It stems from the fact that conditions on presence in and absence from a country for a flow in one direction are equivalent to those for a flow in the opposite direction. A migration measure, with condition referring to duration of presence (in destination country)/absence (from origin country) following relocation is the only exception (II a) (see Figure 2). In this case emigration figure is higher due to the fact that absence from an origin country following relocation may be longer than presence in a destination country and the opposite situation is not possible.

The impact of location of observer on migration numbers for origin-destination specific flows is of higher complexity because of the interplay of origin-destination specific intensities. Consider flows to country B for exponential and Weibull duration model presented in Figure 5 and Figure 6 respectively (all origin-destination specific flows according to origin and destination country for the two basic models (i) and (ii) can be consulted in Annex 1). In most cases the number of migrations viewed from origin and destination perspective does not agree. For flow between country A and B, the outflows from the perspective of the sending country are higher than the inflows from the perspective of the receiving country for all presented measures (II a, II b, III). For emigration from country C to B, there are all possible relations between immigration and emigration numbers depending on a measure type: emigration higher than immigration (II a), emigration lower than immigration (II b), and emigration approximately equal to immigration (III). The analogous relations for total flow to country B will represent a combined effect of the two single flows.

Figure 5 Origin-destination specific flows according to origin and destination country; proportion of relocation number that satisfy a particular duration criterion (a-c); ratio of immigration to emigration number (d); exponential duration model (i)

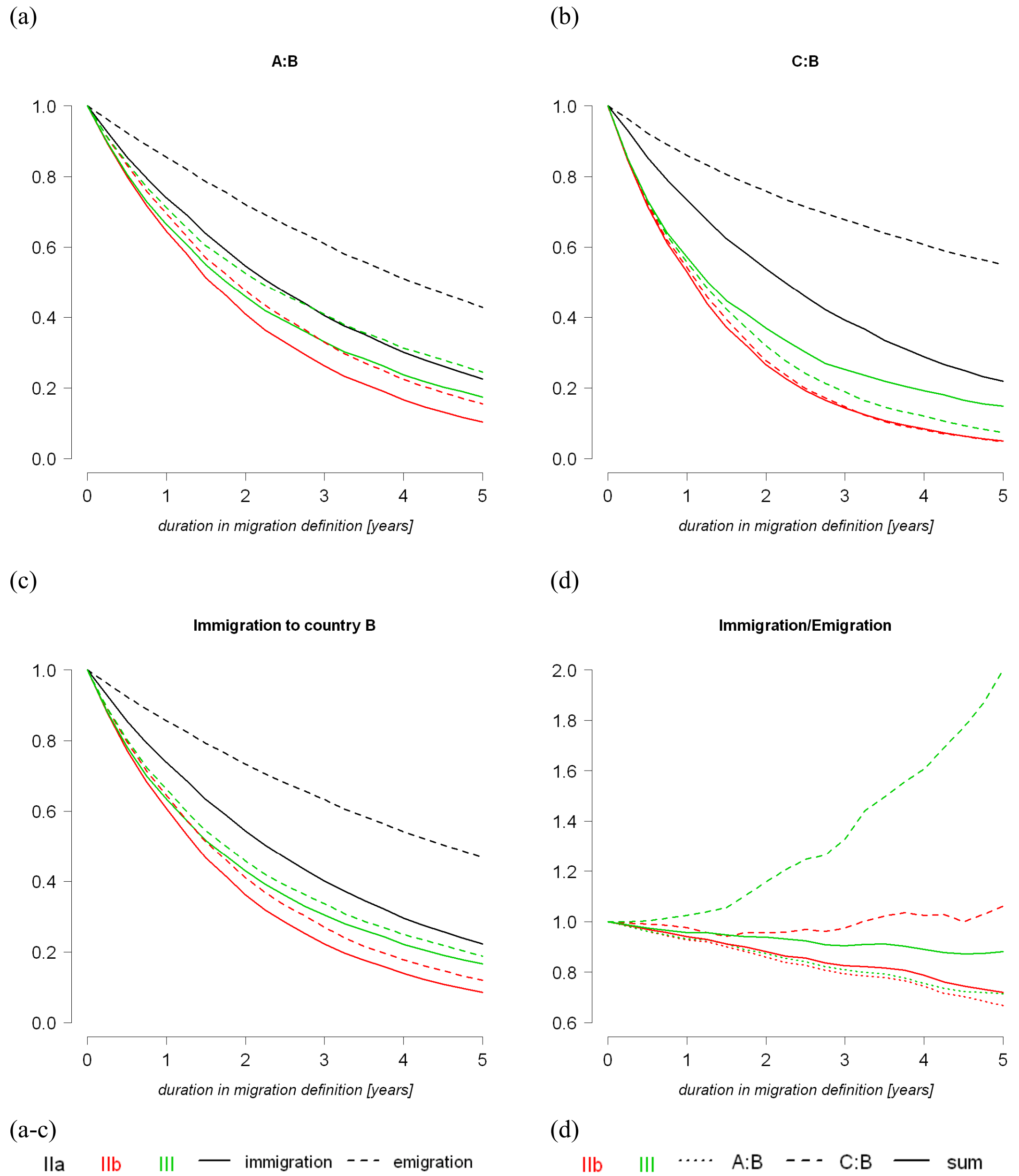
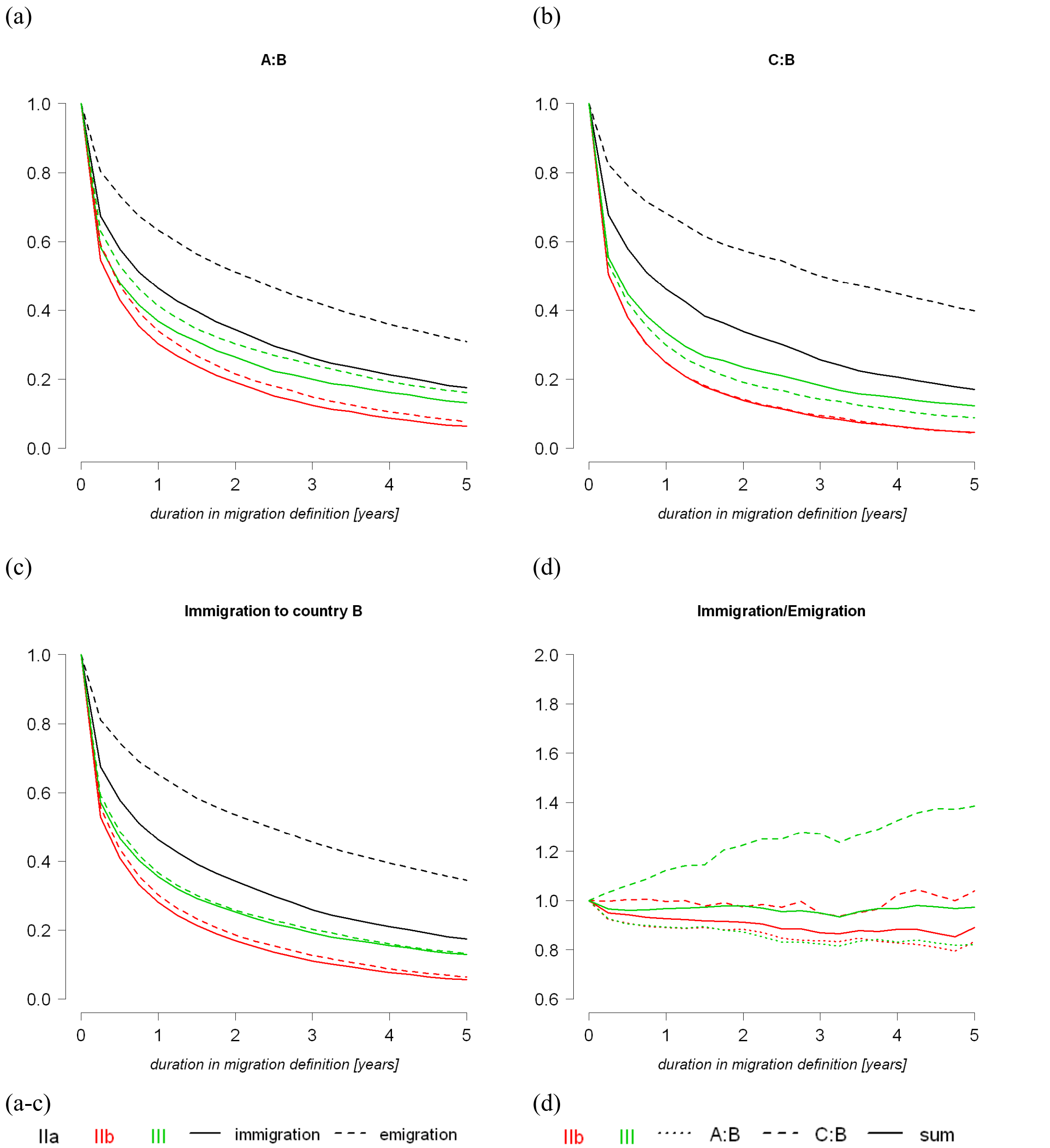


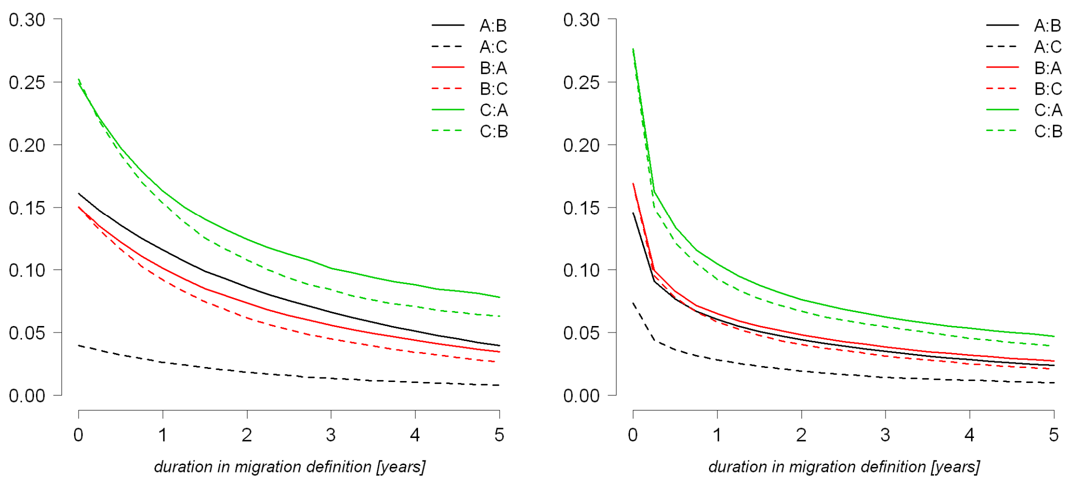
Figure 6 Origin-destination specific flows according to origin and destination country; ratio of different measures to relocations (a-c); ratio of immigration to emigration number (d); Weibull duration model (ii)



Thus, the effect of applying duration criterion of different length on discrepancy between immigration and emigration data is origin-destination specific. When harmonizing data, applying the same adjustment factor to all single flows is not appropriate, unless there are no origin-destination preferences. In addition, contrary to usual expectations, perfect quality data on origin-destination specific flows produced by sending and receiving countries using the same measure and duration threshold do not have to be equal.

The comparability problems presented above were presented for theoretical values of origin-destination instantaneous rates of relocation. In practice they are not known. We are confronted with conditional rates of relocation and the problem is to express one rate in terms of the other. Figure 7 shows the conditional rates of relocation estimated based on time to events considered as migration under the measure (III), which are substantially lower than relocation rates for longer durations.

Figure 7 Emigration rates estimated from simulated events included in measure (III); left panel: exponential duration model, right panel: Weibull duration model, estimation under the assumption of constant hazard rates



If conditional rates are used instead of the unconditional ones for harmonization of the available data the adjustment factors are underestimated. In addition, further disturbances are caused by the fact that conditional relocation rates differ even if the unconditional rates are the same, as is visible in Figure 7.

6. Conclusions

The complexity of migration process poses a considerable challenge of providing an unambiguous definition of migration, which could be successfully applied in practice. In result there are various definitions with possible multiple interpretations. The measure details make for some non-obvious implications, which are of importance in particular for comparison of flows recorded by different countries and modeling the differences between them. Note that these comparability issues remain valid also when the minimum duration threshold used in migration definition by different countries is precisely the same. Differences in the threshold length will cause further disturbances with impact depending on the shape of hazard rates and increasing with the intensity level. In general, all types of discrepancies between different measures are sensitive to characteristics of relocation intensities.

For most presented measures origin-destination migration paths of individuals are not consistent. For instance, statistics may indicate that a person migrates in exactly the same direction a few times in a row. Thus, individuals experience transitions that are theoretically not possible. Being in a particular state and at the same at risk of migrating to other states is ambiguous. In order to ensure origin-destination consistency of migration trajectories more complex and unequivocal conditions have to be specified on establishing and giving up residence in a country. Consistent does not always, however, means complete, because relocations outside the country may not be recorded.

When migration is defined in terms of duration of stay directly proceeding and following relocation, a country of observation does not influence a recorded number of migrations. A migration definition based on length of presence in and absence from a country lead inevitably to discrepancies between migration figures reported by sending and receiving countries. Thus, immigration and emigration figures reported for precisely the same flow do not have to be equal, even if data are of perfect quality. Origin-destination specific intensities lead to origin-destination specific discrepancies between immigration and emigration data.

All migration measures can be seen as conditional relocation with different strictness of the imposed condition. The strength of the condition is reflected in relation between different measures for immigration and emigration figures separately. Nonetheless owing to origin-destination specific relocation rates the relative discrepancies between measures will vary by origin and destination. It refers also to relations between immigration and emigration measures. As not all conditions are amenable to easy calculations, some approximation by other measures may be used to calculate impact of different minimum duration threshold.

References

Boyle, P., Halfacree, K. & Robinson, V. (1998) *Exploring Contemporary Migration,* London, Longman.

Courgeau, D. (1993) Measuring flows and stocks of internal migrants: Selected statistical issues. In D. J. Bogue, E. E. Arriaga & D. L. Anderton (Eds.) *Readings in Population Research Methodology.* New York, United Nations Fund for Population Activities.

Courgeau, D. (2006) Mobility and spatial heterogeneity. In G. Caselli, J. Vallin & G. Wunsch (Eds.) *Demography: Analysis and Synthesis.* San Diego, Academic Press-Elsevier.

Kitsul, P. & Philipov, D. (1981) The one-year/five-year migration problem. In A. Rogers (Ed.) *Advances in Multiregional Demography* Research Report RR-81-6, Laxenburg, Austria, International Institute for Applied Systems Analysis.

Kraly, E. P. & Warren, R. (1992) Estimates of Long-Term Immigration to the United States: Moving US Statistics toward United Nations Concepts. *Demography,* 29, 613-626.

Kupiszewska, D. & Nowok, B. (2008) Comparability of statistics on international migration flows in the European Union. In J. Raymer & F. Willekens (Eds.) *International Migration in Europe: Data, Models and Estimates.* Chichester, John Wiley & Sons, Ltd.

Liaw, K.-L. (1984) Interpolation of transition matrices by the variable power method. *Environment and Planning A* 16, 917-925.

Long, J. F. & Boertlein, C. G. (1990) Comparing migration measures having different intervals. *Current Population Reports* Washington, U.S. Census Bureau.

Newbold, K. B. (2001) Counting Migrants and Migrations: Comparing Lifetime and Fixed-Interval Return and Onward Migration. *Economic Geography,* 77, 23-40.

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

Poulain, M. & Dal, L. (2008) *Estimation of flows within the intra-EU migration marrix.* GéDap-UCL, Discussion paper.

Raymer, J. (2007) The estimation of international migration flows: a general technique that focuses on the origin-destination association structure. *Environment and Planning A,* 39, 985-995.

Raymer, J. & Abel, G. (2008) *Review and improvement of estimation model developed in Task 2 (OD matrix for totals), 2002-2006.* Deliverable 5.1 of the MIMOSA project, 26 May.

Rees, P. H. (1977) The measurement of migration, from census data and other sources. *Environment and Planning A,* 9, 247-272.

Rogers, A. (2008) Demographic Modeling of the Geography of Migration and Population: A Multiregional Perspective. *Geographical Analysis,* 40, 276-296.

Rogers, A., Raymer, J. & Newbold, K. B. (2003) Reconciling and translating migration data collected over time intervals of differing widths. *Annals of Regional Science,* 37, 581-601.

Rogerson, P. A. (1990) Migration analysis using data with time intervals of differing widths. *Papers in Regional Science,* 68, 97-106.

Ross, S. M. (2006) *Simulation*, Academic Press.

Tuma, N. B. & Hannan, M. T. (1984) *Social Dynamics: Models and Methods,* London, Academic Press.

United Nations (1998) *Recommendations on Statistics of International Migration: Revision 1,* New York, Statistical Papers, No. 58, Rev.1 Sales No. E.98.XVII.14.

Willekens, F. (1982) Identification and measurement of spatial population movements. *A National Migration Survey, Manual X: Guidelines for Analysis.* New York, United Nations ESCAP.

Willekens, F. (1985) Comparability of migration data: Utopia or reality? . In M. Poulain (Ed.) *Migrations Internes, Collecte des Données et Méthodes d'Analyse.* Louvain-la-Neuve, Cabay.

Annex 1

Figure 8 Origin-destination specific flows (from:to) according to emigration and immigration country; proportion of relocation number that satisfy a particular duration criterion; exponential duration model (**i**)

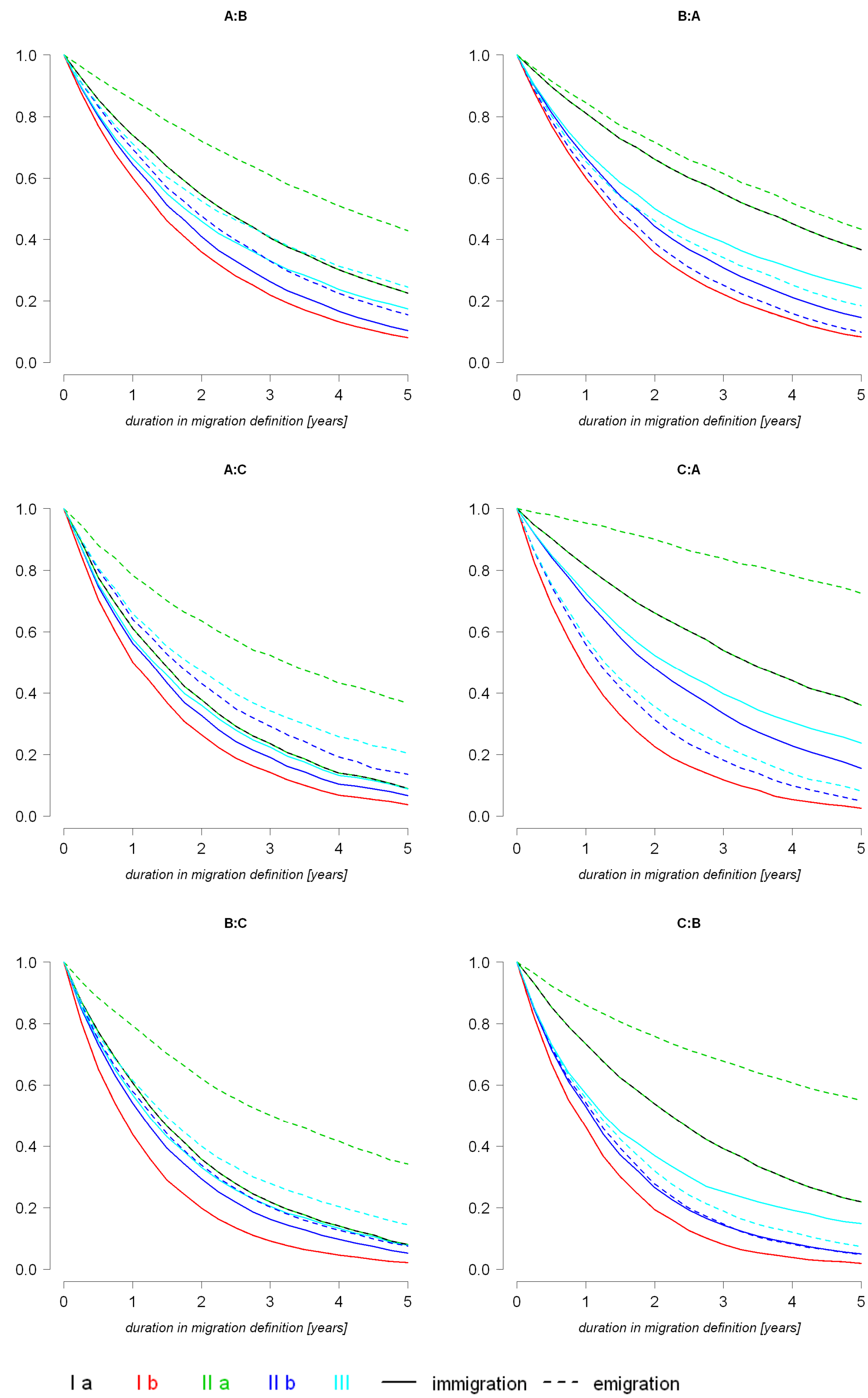


Figure 9 Origin-destination specific flows (from:to) according to emigration and immigration country; proportion of relocation number that satisfy a particular duration criterion; Weibull duration model (**ii**)