

Distr.: General  
11 October 2012

English only

---

## **Economic Commission for Europe**

### **Conference of European Statisticians**

#### **Group of Experts on Migration Statistics**

#### **Work Session on Migration Statistics**

Geneva, 17-19 October 2012

Item 4 of the provisional agenda

#### **Circular migration**

## **THE DATA BASIS FOR MORE ADVANCED STATISTICS ON INTERNATIONAL MIGRATION – THE CASE OF NORWAY**

**Note by Statistics Norway<sup>1</sup>**

### **I. Introduction**

Political discussions are taking place in different fora on ‘return migration’, ‘circular migration’ and ‘Intra-EU mobility’. One can assume that there is a wish to cover as many aspects of these concepts as possible with statistics. Statistics are useful not only as a basis for designing policies but also for measuring the effect of actions. In any case, a growing demand for high quality statistics on international migration is in evidence.

We will start by interpreting and discussing the demands from external users. Even if different concepts are used, many of the demands on the properties of the data and the statistics will be the same. The rest of the paper aims is to answer some of the questions about the availability and quality of existing data and statistics and the possibility of producing new, special statistics.

In reality, the issue of possibilities for new statistics is a very far reaching question, namely how the statistical system of a country functions. However, the aspects of this broad field that can be covered here are limited to the field of demography and in particular migration, where there is a need for a deeper understanding of some elements.

The aim is to say something general about this topic, and for that reason the first part is devoted to a general, theoretical discussion. This is followed by information on the situation

---

<sup>1</sup> Prepared by Kåre Vassenden.

in Norway, including some comparisons with other countries, in particular the Nordic countries and the Netherlands.

## II. A model of the migration between countries

The starting point is a country and its relation to the outside world in terms of migration of people. If we take one specific migrant, the first move is leaving the *home country*, and if he subsequently returns to the same country the return can be called an *in-return* or a *return from*. The opposite is the case when the country in question starts as a receiving country in its involvement with the migrant. If the migrant then returns it will be an *out-return* (*return to*). For the sake of simplification, we choose here to denote the initial sending country 'A' and the initially receiving country 'B'. This model can be called a one-country model, as it includes only one identified country.

In principle, each country has two return flows: Natives returning home after a stay abroad (in-returns), and foreign nationals returning home after a stay in the visited country (out-returns). At least for the time being only the latter returns are policy relevant in Norway.

However, the broad grouping of the outside world into one big unit is insufficient in many cases. If this grouping is broken down to around 200 single countries, the model will consist of country A that sends out someone to country B. The person will stay there permanently, or will return to country A or go to a third country (C). If C is chosen, the migrant will once again have the possibility to stay permanently, return home to A, return to B or even find a fourth country. It is possible to repeat some of the moves and establish long complex chains. In the discussion here we focus mostly on Norway as a B country and the out-returns.

## III. Interpretation of the different return migration concepts

What distinguishes 'return migration' from just ordinary 'immigration' or 'emigration' is the *history* or *chain* of moves. 'Return migration' entails at least two migration events for a migrant. A relevant question is whether moves during childhood should be omitted from the account. The other extreme would be to also include the parents' moves in the count, which means that even a simple out-migration performed by a second generation immigrant would be a 'return'.

There is clearly more than one basic perception of the term 'return migration' in use. In addition to the aforementioned distinction between in and out-returns there is a difference between return to or from the rest of the world versus return to or from a specific country. In the first case, the options are simply staying or returning, where return covers moves to all other countries in the world. With a model consisting of only specified countries, the options are staying, going/returning to A, or going to C. In this model, the term 'return' does not cover cases B-C, only B-A. Users of the term 'return migration' should be aware that the different meanings exist side by side, and that there is a potential for misunderstandings because the definitions cover different flows.

Even if A-B-C is not a return, A-B-C-A would be an indirect return. This trajectory can be observed from country A, but not from B.

The requirements for statistics on circular migration are stricter. The concept presupposes a model with at least two countries. The word 'circular' can easily be understood to mean that the migrant covers at least three countries (A-B-C-A), but in reality most cases under this label are simply back-and-forth movements (A-B-A).

In addition, the circular migration concept stresses the temporary nature of the migration, and possibly even a repetition of change in place of usual residence. Information on reason for immigration or actual activity in country B should be available, as circular migration is first and foremost work-related. A key point is whether the migrant brings relevant work experience, training and/or economic wealth (including social benefits and pensions) with him back to the home country. Citizenship and other formal and accumulated rights are also relevant. In general, it seems that the concept demands a great deal of information on the person's socioeconomic situation at the time of the move.

One particular demand is the wish to know about migration between a limited group of countries. These are normally neighbouring countries, and as such this need distinguishes itself from the typical circular migration. In fact, every time the relevant geographical area for the movements is smaller than the rest of the world, specified geographical information needs to be included in the data on movements. In addition to the aforementioned elements, information on accompanying family members and cross-border work (commuting) has been mentioned in requests concerning intra-continental migration.

Other kinds of similar demands emanate from some individual national statistical institutes (NSIs) that need figures from other countries to supplement their own imperfect data. Also, the NSIs are faced with the EU regulation on migration statistics, which stipulates a maximum share for unknown country variables. Last but not least, authorities and others ask where there are nationals abroad and how many there are. The return aspect is not necessarily included in any of the last four uses, but the quality of geographical information is essential in all of them.

From this review, it is clear that a geographical breakdown of 'the rest of the world' and sufficient quality on the duration of stay is compulsory for the advanced migration statistics. In addition, there is often a need for information on reasons, rights, permissions and more general personal information. Henceforth in this paper we will give priority to some core topics related to the migration events, and will not go into detail about the personal information.

## **IV. A general discussion about the demands viewed from a statistics' provider's perspective**

### **A. The situation for NSIs**

It is important to maintain a distinction between the administrative system of a country and the NSI, and between the choices made at these two levels, respectively. As a general rule, NSIs do not have many choices in relation to data sources; they have to take the data they can get. NSIs sometimes nevertheless have some influence on the administrative system. In any case, in the production of statistics there are decisions to be made even by the NSI.

The demands on NSIs in this field range from the basic, coarse numbers on international migration to a complete overview of the movements, in addition to comprehensive information on the migrants. The absolutely highest expectations that can be observed, however, require more in reality than a normal NSI can realistically comply with. At best, national administrative systems and NSIs only have satisfactory information on the events taking place in their own country. Knowing about events that take place in other countries is a real challenge. Without a centralised, integrated or coordinated international administrative and statistical system it is not possible to follow the migrants' total trajectories from one country to another.

## B. General demands on systems and data

For a country to be able to address the challenging requests, its general, basic system must be of sufficient quality. The first question is whether an NSI has data on immigration and emigration at all. A negative answer means no statistics on return or circular migration.

In reality, the main issue is the quality of the *national administrative register system* in general and the *national population registration system* in particular. Included here is the possible cooperation between countries at the administrative level. A well-functioning register system is based on a main ID number series. This number must be relatively stable over both time and sector, and in any case links between replaced and valid numbers and between different number series must exist. The ID number(s) should be widely used in the administrative registers of a country. Only with this functionality is it possible to find the same person at different times and in different connections.

From the perspective of NSIs, it is important that a large range of administrative data sets exist, and that technically these can be linked to each other. NSIs must have access to all the relevant data and be allowed to link them. Only in this way is it possible to produce the comprehensive set of information on persons that is demanded by the circular migration concept.

Upon receipt of the data there will be some adaption of the register data to statistical needs. NSIs may have different policies, procedures and conditions for some important choices in the production of statistics. This could include the length of the waiting period between the reference period/date and the extraction of data, choice of variable for date of event (actual or official), handling of annulments and corrections and handling of events with an event date before the reference year ('the lag'). The final quality of the official statistics is a function of the quality of both the administrative and statistical systems.

Official statistics are sometimes produced in a way that does not take other data needs into account, but NSIs normally store raw or processed data than can be exploited for the production of new, special statistics. However, third parties seeking to compare the quality of data between countries have to rely on the published statistics extensively.

## C. The size of the correction factor in the population balance

If population numbers and the growth components are independent of each other, it is possible to calculate the annual population growth both as the difference between the two population figures and as the sum of the components. The difference between these two measures constitutes the correction factor. If this factor is large (in absolute terms), this means that the component figures do not fully explain the population growth. If the component-calculated growth is lower than the population numbers indicate, the growth is under-explained. Over-explaining is equally invidious for the quality of the statistics on population balance.

From a migration statistics perspective, a relatively small correction factor in the population balance means that the two migration components have done their bit to explain why a population has grown. In addition, a small factor indicates that few events are missing (or superfluous where the figure is positive). The main cause of a large absolute value of the correction factor is usually that the emigration component is too small.

The correction factors may give an indication of the quality of the *statistics* (and published statistics always matter), but do they give any indication of the quality of the underlying migration *data*? At least a small factor shows that the data are complete in terms of events. A large factor, however, is not reliable as an indicator of quality as some individual level

data sets contain more events than are included in the official statistics. However, there is a risk that these additional records have more quality problems, in one way or another.

#### **D. Event data consistent over time**

We have seen that there is a distinction between annual statistics and individual level data. However, there are also differences between types of data sets. Some crucial aspects of the data used to produce official statistics must be closed when compiling the statistics. High quality research data sets, on the contrary, must be open for modifications based on the current flow of annulments, corrections or delayed records.

Event data at individual level must be dated and consistent over time (and preferably also between sectors). When the events follow each other in a logical order, they fulfil the requirement for 'life course' or 'event history' data. This general quality requirement is particularly important when the task is to produce circular migration statistics.

#### **E. Special demands on migration data. Are the relevant events included in the data?**

A basic question is whether the registration system is interested in and can identify only the migration events required by the statisticians and demographers. The biggest problem is, of course, possible *missing* registrations, because these cannot just be created later. If, on the contrary, a registered event is superfluous according to the relevant statistical criteria, it is sometimes possible to delete the event at a later date.

The issue relates firstly to the laws and regulations governing the registration system and secondly to the capability of the system to capture all the movements it is supposed to in a timely way.

The body of rules defines the *legal residence*. Very often there is conflict between this concept and the *actual residence* of a person, but it should be noted that even actual residence is sometimes difficult to define. In any case, it is important for the statisticians that the concept 'legal residence' does not become so 'legal' that the residence aspect of it is lost. However, there should be no illusions about the potential to achieve perfect quality.

Another issue regarding the body of rules is the shortest length of stay for qualifying for registration. Here the demands on the population registration system head off into different directions. On the one hand they face a need to assign ID numbers and keep records of as many relevant people as possible for purely practical administrative reasons. On the other, the criteria for assigning rights often call for a restrictive line. The solution is to build up an inclusive population register and let the different rights be specified explicitly by designated variables. Such an approach probably increases the possibility for statistical demands to be met.

The second quality issue is whether the system manages to capture the events specified by the body of rules. It is a well-known fact that many people do not notify the population registry when they are supposed to. The reasons for the extent of the problem are to be found in the structure of incentives, the information provided, habits from home country and so on. It is no doubt possible to do something about some of these factors, and thereby improve the quality of the population registration. In any case there must be legal authority, systems and methods to maintain the quality of the register even when inhabitants do not provide the necessary updates.

## F. Geographical specification of migration data

For a typical NSI, it is not always obvious that statistics distributed by previous/next country are very important. An NSI's main task is to serve its own country, and its loyalty lies with national considerations. Normally, it is most relevant to look at the moves to and from the rest of the world. When more information is needed about the migrants, the information on citizenship goes a long way to covering these needs. By using the citizenship variable, the marginal gain of including the country variables is rather small – except for the national citizens.

In this *national perspective*, the issue is ‘what is going on with a specific country’. To answer this, different information on, for instance, migration is useful, but how the information is adapted is the country's own business. Even if the rest of the world is divided into single countries in the statistics, from a national perspective it does not matter if the mirror figures of other countries are not consistent.

If the question is what is going on *between* countries, the perspective is *international*. In that perspective ‘country of previous/next residence’ is a totally fundamental variable. When a solitary country takes on the task of compiling such statistics it is possible to talk about a simplified version of the international perspective. A lack of comparability with other countries is not really an issue in this version. Circular migration projects presuppose an international perspective, but this may be of the simplified version.

With a full international perspective, country variables must not only exist in the statistics – they should be of such a quality that they do not deviate too much from the statistics of other countries. With this perspective, figures reflecting relations with other countries turn into a common responsibility. With a full international perspective, cooperation, harmonisation and comparability are more in focus. The final goal is a coordinated set of data.

We have seen that information on place of previous/next residence is essential for advanced migration data, but how necessary is it to define these places as single countries? No doubt in a few cases it could be sufficient to use groups of countries, provided that the groups are not too big. From some development perspectives it should be enough to know that someone returned to the African horn instead of Eritrea, for example. However, the established single countries are much more standardised and in most cases more useful as a geographical specification, and are definitely to be preferred.

The countries that have been split constitute a special kind of forced grouping. In an ideal world it should have been possible to recode out-dated country values into countries that exist today. In practice, however, it is not possible to avoid, for example, Soviet Union as a value in data and statistics of a certain age. A compromise here is to regard an outdated value as valid if the migration took place when the country existed, but as a grouping if the value is used by an NSI after the split.

Finally, some states in the world are composed of territories that have traditionally been seen as separate countries in national and international demographic statistics. It should be seen as a kind of grouping if these countries are included in one entity. Simply using the name of the biggest entity as the name of the grouping does not change the realities in respect of comparability. One example of this is the name ‘Denmark’, which in reality refers to the grouping ‘Denmark, Faroe Islands and Greenland’ in some NSIs.

## G. The problem of a high proportion of unknown country

Even if the necessary country variables are included in the data set it is problematic if too many migration events have an unknown country of previous/next residence. For natural reasons, this problem applies in particular to emigration data. The value 'unknown' can be limited to 'totally unknown', or it can also include 'partial' unknown values (i.e. different kinds of groupings).

In general, comparing mirror statistics is useful to check a country's own migration figures against the figures for the same flows produced by other countries. However, the higher the share of unknown values, the less that can be gleaned from the mirror statistics; high coverage is imperative to reaping the full gains of mirror statistics.

A question asked by producers of statistics is whether it is possible to *impute* country of next residence in some way or another. It is tempting to just let the emigrants go to the country they came from instead of an unknown place. However, when the most interesting issue is whether the emigrant left for the home country or a third country it is simply not possible to use such a method. If someone has immigrated from a specific country, however, then emigrated to an unknown country and finally once again immigrated from the first country, it should be safe to assume that the country in the middle of this chain was the same as the two others. The drawback, however, is that this possibility only applies to frequent migrants and to their 'old' events.

The last possibility would be to assign a country of next residence randomly according to the distribution found in the cases with known destination. However, this would only give an accurate result if we know that the known cases are representative of the unknown cases. In any case, this method reduces the advantage of having individual level data. Estimation possibilities are outside the scope of this paper.

## V. Norway's situation

### A. The basic system

What characterises Norway and some other countries (primarily the Nordic countries) with regard to administrative register sources is that they cover many aspects of society, the quality of the registers is relatively high and there is contact between registers that enables a degree of coordination. The Population Register plays a key role with regard to person data in Norway. It has assigned ID numbers and registered central demographic and legal person information since 1964.

Most of the individual data that Statistics Norway receives uses the Personal Identification Number (PIN) as identification. The *D-number* is used for persons who do not qualify for a PIN but have economic relations with the Norwegian authorities – a similar role to the Swedish 'coordination number' in Sweden. The fact that very little information is connected to the D-numbers and that it is difficult to know if people with only a D-number actually stay in Norway or not is, however, problematic.

The PIN has been a key factor in the production and development of social statistics in Norway. The existence of this ID number series in practically all relevant input registers enables the simple and secure forming of links. Statistics Norway has access to a broad range of data that makes it possible to meet demands on socioeconomic information on such areas as migrants. In addition to the data retrieved externally, Statistics Norway has considerable amounts of data from earlier data collections. These data also represent an important source for the production of statistics.

## B. Event data

Norwegian event data mainly date back to around 1967 or some years later, but they are probably less exploited than the Swedish event data. The Swedish population register data were relatively stable from the beginning, and the NSI has been very systematic and now possesses a large range of good quality data ready for use. By comparison, the oldest Norwegian data appear more 'raw'. Before these data are fit for advanced life course analysis, further scrutiny, rectifying of errors and omissions and generally more processing are all needed. In the past, Denmark also had a project with similar aims ('Danish Demographic Database').

## C. The size of the correction factor in the population balance and the implications for migration data

We have seen that the correction factor can be used as an indication for problems with the data quality (aside from this indicator having an independent role). In table 1, Norway comes out with a rather small annual correction factor compared with other countries, and there are no extremes among the Norwegian figures. At least for the recent decade, the reason for the result is a deliberate policy to include the lag in the official event statistics. In any case, this bodes well for further examination.

**Table 1. The correction factor in the population balance as a percentage of the alternative for population growth that provides the highest absolute figure. Negative percentage means that the growth is over-explained by the component figures.**

Year	Norway	Sweden	Denmark	Finland	Iceland	Greenland	Faroe		Belgium	Slovenia	Austria
							Islands	Netherlands			
1996	-0.6	-2.4	1.5	7.9	-4.5	-20.6	0.0	6.1	12.5	-50.9	0.0
1997	-3.9	-9.4	1.4	7.3	5.2	-49.3	0.0	2.9	17.9	18.7	0.0
1998	-1.2	1.0	-0.2	8.8	2.8	81.8	0.0	0.9	0.3	50.1	0.0
1999	0.1	-6.3	3.1	5.6	0.5	-77.7	0.0	3.3	-2.9	89.6	0.0
2000	0.1	-0.8	3.2	-1.7	2.6	-53.1	0.0	2.6	1.2	5.6	0.0
2001	-0.2	-0.1	0.2	2.5	-3.6	15.2	0.0	4.4	0.9	50.1	21.2
2002	0.2	-0.7	2.6	0.3	-2.9	-38.2	0.0	3.7	5.7	34.0	2.9
2003	-0.2	-0.2	5.8	0.4	-3.9	22.5	0.0	11.3	6.9	8.4	6.4
2004	-0.2	-0.3	0.9	0.3	2.2	-36.8	-5.4	13.2	4.9	-13.7	5.6
2005	-0.3	-1.1	0.9	0.9	0.2	60.3	7.8	16.1	7.5	0.0	10.6
2006	-0.2	-0.1	0.8	1.2	0.2	4.3	0.0	22.8	5.2	0.0	3.4
2007	-0.2	-0.1	-9.1	1.2	0.5	11.8	-55.4	10.2	4.3	-99.3	-2.1
2008	-0.1	-0.4	-10.2	-0.2	-2.1	13.1	-0.3	6.4	12.4	0.0	-1.2
2009	-0.1	-0.5	-23.0	0.1	-4.0	-1.5	25.4	4.5	10.4	0.0	2.4
2010	-0.3	-0.3	-17.7	0.1	8.4	-20.1	47.2	-0.7			
2011	-0.4	0.2	-31.6	-0.8	1.4		0.0	0.6			

Sources: National statistical databases and the Nordic statistical database.

Each of the figures in table 1 reflects differences in in-data, policies and/or criteria for the production of official statistics.

## D. Registration of residence

Probably only in a few countries is the concept of residence as strong as in Norway. Not only does it entail inhabitants having to give notification every time they move, it also means that the Population Register constitutes a public authority, which through legislature decides who is resident in Norway.

The body of rules on where persons shall be registered as resident is based on where they take their regular night-rest, but there are several exceptions to this main rule. For the purpose of preventing tax evasion, the rules make it difficult to be registered as emigrated. The result is that the registration of residence does not reflect the actual situation as closely as statisticians would prefer.

The biggest challenges, however, are linked to the moves that are not notified to the Population Registry; this particularly applies to emigration events. Cases are constantly being discovered where someone has left Norway for such a long period that they should have been registered as emigrated. This problem does not apply to emigration to the other Nordic countries, however. If someone moves from one Nordic country to another, the move will not be registered as an emigration event in the sending country before the receiving country has accepted the move as immigration and then notified the sending country about it.

In cases where someone has left the country without notifying the registry, the registry makes a decision that the person in question has emigrated. This 'cleaning up' is a good thing for the quality of the statistics, but unfortunately the data quality of these cases will not be the same as for ordinary notifications. Checks vary considerably between years, local offices and seasons.

A new possibility for analysing the data quality was introduced in January 2011. The new variable 'Source for registration of emigration' in reality has these values: 'Notification from the migrant', 'Notification from one of the other Nordic population registration authorities' and 'Administrative decision'. In the time the variable has existed each of the three categories gets approximately the same share of the emigration events. This means that in this period some information is missing or is more formal than actual for up to a third of the emigration events.

## E. Qualities of the migration dates

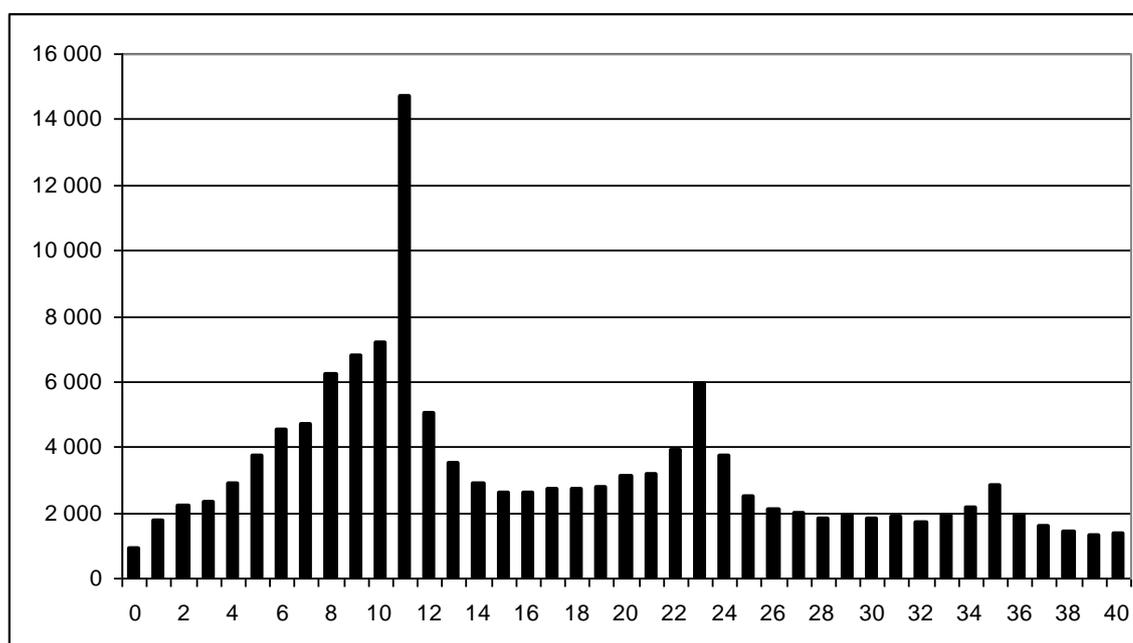
In general, the official date of an event is important for the population registration system. The date is not something a migrant can choose himself as there may be major legal implications for both the migrant and the authorities. For a first time *immigrant*, the official date of immigration is when he turns up at the population registry with a residence permit (where this is required) and other documents. When someone notifies their *emigration* a week or two in advance, the official date will be the date given by the emigrant. If the notification is received by the registry after the given date, the official date will be set to the day of receipt of the notification.

The official event date is normally acceptable when an emigration is notified by the migrant or a Nordic country. When the source of an emigration event is an administrative decision, however, it is up to the registry to set a date. The main principle is that migrants should not gain from not reporting. Until 2009, most cases were based on a passed expiry date of the residence permit, and in those cases the date of emigration was set to the expiry date. Now in 2012, EU citizens do not need a residence permit, in which case other solutions are used. In any case, the date used will often be legal/technical and not as accurate as many parties would like.

A study of all stays in Norway between 1996 and 2010 (where both the in-migration and out-migration were within this period) reveals that by far the most ‘popular’ length of stay is 365 days. The frequency of these stays is more than 2.5 times that of the next length of stay on the list, which is 364 days. Another common length of stay is 730 days, which is exactly two years. The reason for this picture is simply that a lot of immigrants have had a residence permit lasting one year (or two years). Typically, the migrant has left without notifying their move, and later when the population registry has checked expired dates of residence, they have ‘emigrated’ the person and entered the expiry date as the official date of event.

The conclusion that can be drawn is that emigration dates are not very accurate. The consequences are greatest for studies with short stays. However, if we lower the expectations and take a broader view, the data are nevertheless useful.

**Figure 1. Stays in Norway by length of stay in months. Stays with both the in and out-migrations in the period 1996-2010, up to 40 months.**



## VI. The quality of the Norwegian data on source and destination country

### A. The percentage of complete or partial unknown country of previous/next residence

Norwegian population statistics do not differ from other comparable statistics in other countries, and have not given particularly much attention to the country variables. However, the tendency towards an increase in the percentage of unknown variables has been a cause for concern.

The *immigration* statistics of this study’s eight comparison countries have rather low percentages for unknown previous country; mostly less than 1 per cent, or 1.6 at the highest in recent years. In general, the situation for the immigration figures can be deemed acceptable.

Emigration is different. In most of the years in the 1990s, the percentage unknown in Norwegian official statistics was around 5 per cent. Then in 2003, the figure suddenly jumped to 13 per cent, staying at slightly above this until 2009, when it shot up to 24 and then 27 per cent. In 2011, the percentage was a little lower with 21 per cent, but this is still not satisfactory.

**Table 2. Percentage of unknown country of next residence in official emigration statistics. 2000-2011**

Year	Norway – official statistics	Norway revised *)	Sweden	Denmark	Finland	Iceland	Faroe Islands	Netherlands
2000	6.7	6.7	5.9	5.3	0.2	0.1	0.0	22.5
2001	9.4	9.4	5.0	5.7	0.0	0.0	0.1	23.3
2002	5.9	5.9	8.2	7.0	0.2	0.0	0.1	31.2
2003	13.0	13.1	9.8	7.8	0.0	0.1	0.0	34.3
2004	15.9	16.0	9.6	6.8	0.4	0.1	0.0	31.9
2005	15.6	16.7	8.4	7.4	0.6	0.5	0.1	30.3
2006	15.2	16.3	11.1	6.4	0.2	0.5	0.1	31.3
2007	17.7	18.4	12.6	4.5	0.1	0.8	0.2	25.5
2008	18.0	19.9	8.4	3.9	6.8	0.6	0.1	23.5
2009	24.1	26.6	4.8	5.1	0.6	0.6	0.2	23.7
2010	26.8	29.5	13.5	4.5	0.9	0.8	0.1	26.0
2011	20.5	22.7	7.3	5.1	0.4	0.8	0.8	21.5

Sources: National statistical databases

\*) Percentages after excluding some possibly imputed transactions. See more in chapter 'The percentage unknown is probably higher in reality'.

In the comparison group, Finland, Iceland, Greenland and Faroe Islands are all very low, with figures typically below 1 per cent. Denmark is stable at 5 per cent, while Sweden has had two-digit figures recently, but is not as high as Norway. Only the Netherlands can compete with Norway. In Dutch emigration statistics, the percentage unknown has always been 20-30 per cent. In summary, the table has two NSIs with substantial problems with the coverage, one of which is Norway. At the same time, it appears that some countries have almost no problems. The huge differences between the countries are hard to believe.

Some countries have a relatively low percentage in table 2, but have high figures for the size of the correction factors (table 1). It can safely be assumed that many of the emigration moves with unknown destination country are found in the correction factor. However, Finland is impressive with low figures in both tables. At this current time, the reason for this is not known.

Based on the way the Nordic registration system functions, we can assume that the coverage of country of next residence is almost 100 per cent for these migrants, and hence that the unknown cases apply to non-Nordic destination countries only. A limitation of the denominator to include only the non-Nordic emigration does not change the problem in the absolute sense, but the percentages will be more realistic for the non-Nordic migration. For Norwegian statistics, this new measure increases the percentage unknown by more than 10 percentage points in each of the recent years. Consequently, the figure in 2010 for unknown non-Nordic country of next residence reaches 40 per cent (and 32 per cent in 2011). Even for Sweden, this method of calculation means a noticeably higher level of unknown variables (10-20 per cent).

The next step is to include 'partly unknown' and see if that increases the percentages for some of the countries. For the most part, it does not have any noticeable consequences.

Only the Netherlands can see an increase of around 5 per cent, mainly due to the migration to and from countries that do not exist any longer in the Caribbean. Cases with Denmark used as a grouping were not included.

## **B. The percentage unknown is probably higher in reality**

While exploring the new variable for source of emigration registration in the spring, we discovered that there were several cases where the destination country was registered even on administrative emigration events. It transpired that most cases related to one of the regional offices, where citizenship had been used as a substitute value for missing 'country of next residence'. As a result of the discovery, the central Tax Administration instructed the local offices to record a valid country value only where real information exists.

The implication of this is that the quality of the variable has looked better than it is. If all decision-based emigration events that had a country of next residence in the examination period were false, the percentage unknown in this period in reality would be 30 instead of 20. A small number of the entries were probably based on facts, but not enough to change the overall picture to any great extent.

Another possible way of finding the effect of imputed cases is to exclude notifications registered by single employees that have practised the imputation method. The result is that percentages increase by 2-3 percentage points in a few recent years (see column 'Norway revised' in table 2).

## **C. Which destination countries have the lowest coverage?**

It is natural to ask which of the different countries of next residence are most affected by the missing information. It is of course difficult to comment on the countries themselves, but the citizenship of the migrants may give us some indications, and is nevertheless interesting in itself.

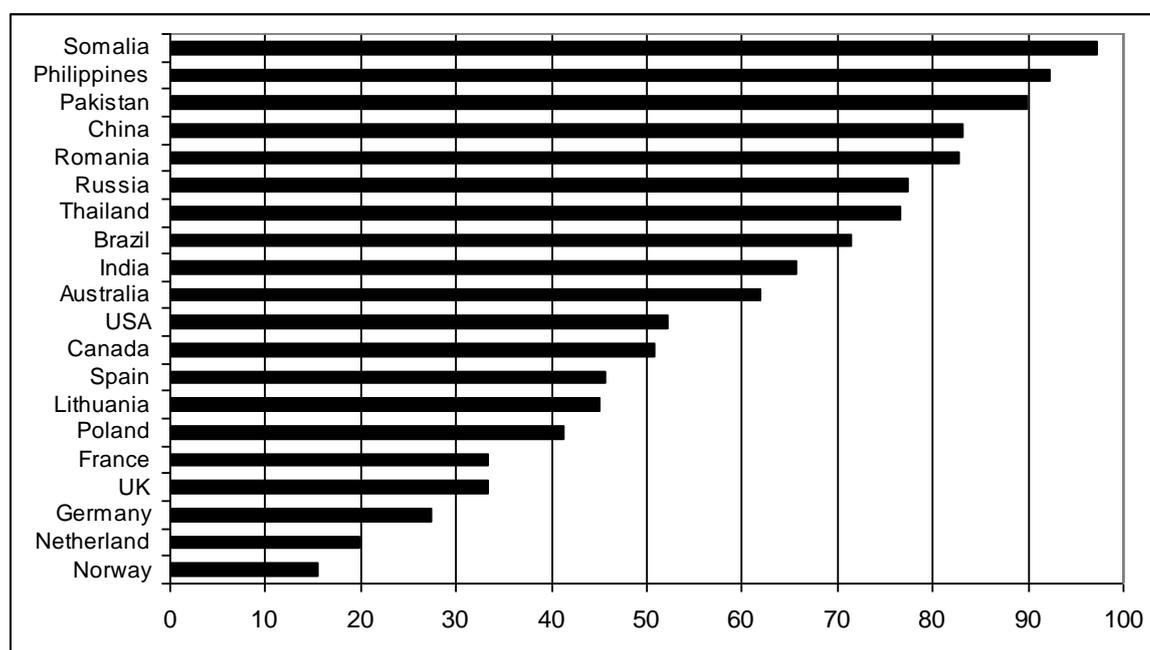
The data we use are transactions received after January 2011 when the new variable 'source of emigration registration' was included. If the emigration is registered according to an administrative decision, the chance of knowing the destination country is rather slim.

As expected, extremely few of the emigrating Nordic citizens are registered on the basis of a decision – only 0.6 per cent! The main reason is that they tend to emigrate to one of the Nordic countries, where their emigration from the sending country is taken care of by the population registration systems in the two countries involved. If these migration events are excluded, 17 per cent of the Nordic citizens emigrating were registered on the basis of a decision.

Among Norwegian citizens, 9 per cent are administratively emigrated (16 per cent without the Nordic emigration). For this citizenship group, the main problem for the statisticians is not this percentage, but that too many of them do not have their emigration registered at all because of the strict rules. The next group is citizens from Switzerland and the EEA outside the Nordic countries, where 38 per cent were registered according to a decision. For all other citizens combined, the figure was as high as 73 per cent.

A breakdown of the most numerous single citizenships (excluding emigration events to the other Nordic countries) shows percentages from 16 (Norway) to 97 (Somalia). For the Somalians, it would be interesting to know if they emigrate back to Somalia or to the UK, for instance, but if they do not fill in an emigration form at the population registry it is very difficult to know their destination country.

**Figure 2. Percentage of emigration events registered by administrative decisions, by citizenship. Excluding emigration to the other Nordic countries. The most numerous citizenships. 17 January 2011–22 May 2012**



As one would expect, the proportion of non-reported country of next residence is very high in the lag records, varying annually from 50 to 75 per cent. For the other transactions (that at least do not look as if they are delayed), the percentage unknown has remained very stable at 14 per cent over the last four years. The policy to include lags in the annual statistics means that the percentage unknown is higher than it otherwise could have been. On the other hand, even 14 per cent is high, and higher than the level in some countries of comparison.

Safe imputation methods (including even information on postal addresses) are used in the production of Norwegian migration statistics according to the EU regulation, but the effort is of little help.

## VII. Summary and conclusions

Demands from new concepts and users represent a challenge to the producers of migration statistics. Some of the concepts have a large range of factors that need to work out; much like a steeplechase where every obstacle must be passed without failure in order to reach the finishing line.

We have asked ourselves if the Norwegian data are good enough to be used for more advanced statistics on international migration. Many answers are found outside the limited field of migration statistics. Notwithstanding, Norway seems to pass all the opening tests. Its data and statistics appear to have good prerequisites for responding to the demands on different kind of advanced migration statistics. One of the strengths it has together with a few other countries is the possibility to add a broad range of person information.

On the last, decisive leg, however, Norway appears to encounter problems. The test is about the percentage unknown in the geographical variables. It is true that the first part of this –

‘intra-Nordic migration’ – is passed with flying colours, but when it comes to emigration to non-Nordic countries Norway hits a brick wall.

It should be noted that the results are not always very bad when we limit the countries of interest to the closest non-Nordic area. However, in the concept of circular migration, the point is the migration between, to and from different kinds of ‘developing countries’. It is at this point that the results start to be disappointing.

Can anything be done to reduce the deficiencies? If there are solutions, they are to be found in the population registration system. It is probably possible to get more migrants to notify their migration themselves. Another possibility lies in the potential for greater internationalisation of administrative systems, such as the exchange of national insurance information within the EU/EEA.

---