

**THE MEASUREMENT OF INTERNATIONAL AND  
INTERNAL MIGRATION IN THE 2010 GLOBAL  
ROUND OF POPULATION CENSUSES**

**TWELVE KEY RECOMMENDATIONS ON QUESTIONS,  
CONCEPTS AND PROCEDURES**

**DRS WILLIAM L J XU-DOEVE**

2006

*Paper prepared for presentation at the European Population Conference of  
the European Association for Population Studies (EAPS), Liverpool, UK,  
21-24 June 2006*

E-mail: [w.l.j.xu-doeve@uwnet.nl](mailto:w.l.j.xu-doeve@uwnet.nl)

*Up-to-date contact details can always be  
obtained from the IUSSP membership  
directory at <http://www.iussp.org/>*



## **ABSTRACT**

For most countries, a population census is the only instrument that allows demographers to obtain a comprehensive image of international and internal migration. The only exception are the minority of countries which maintain a continuous population registration system in which the moves made by individuals are systematically recorded.

This paper critically examines the current international *principles and recommendations for population censuses* of the United Nations from the perspective of the measurement of international and internal migration. These methodological guidelines are shown to be less than optimal in a number of fundamental respects.

To assist countries in significantly improving their ability to measure and analyse international and internal population movements, this paper develops *twelve specific recommendations* for the next global round of population censuses to be taken around 2010.

In formulating these recommendations, this paper first of all addresses the numerous international and national organizations and bodies involved in the development of the updated international principles and recommendations for the 2010 global round of population censuses. However, the findings of this paper are equally important for national statistical organizations which will actually be taking and processing their national 2010 population census, and for analysts of data on international and internal migration.

## **KEYWORDS**

international migration, internal migration, demography, methods of measurement, methods of estimation and adjustment; 2010 global round; population census; methodology; principles and recommendations

## TABLE OF CONTENTS

	page
Abstract	ii
Keywords	ii
Table of Contents	iii
1 Co-ordination of the 2010 Global Round of Population Censuses	1
2 The Importance of Population Censuses for the Measurement of Migration	4
3 The Measurement of International and Internal Migration	6
4 Concepts and Procedures for Measurement and Analysis	12
4.1 The Concepts of Migrant and Migratory Move	15
4.2 The Concept of Place of Usual Residence	23
4.3 The Concept of Duration of Residence	30
5 Data Processing	35
Conclusion	42
Appendix: A Brief Note on why a Question on <i>Duration of Residence</i> is preferable to a Question on the <i>Place of Residence at Some Fixed Date in the Past</i>	45
References	49
Summary of the Recommendations	50

## 1 CO-ORDINATION OF THE 2010 GLOBAL ROUND OF POPULATION CENSUSES

In this section we briefly sketch the outline of the main international co-ordinating framework designed to guide and assist individual countries in the preparation of the 2010 global round of population censuses. In so doing, our *focus* is on the *measurement of international and internal migration*.

The principal international co-ordinating structure for the organization of the next global round of population censuses is provided by the *2010 World Programme of Population and Housing Censuses* of the United Nations Statistics Division (UNSD). Within the work programme defined by UNSD, *three working groups* and *six technical subgroups* have been constituted for the review and update of the current version of the "Principles and Recommendations for Population and Housing Censuses" (United Nations, 1997). These working groups and technical subgroups comprise experts from a broad range of organizations, including international co-ordinating bodies and national statistical offices.

In particular, for the topic of migration a *Technical Subgroup on Internal and International Migration* has been established as one of the expert subgroups of the Working Group on Standards and Frameworks and a Core Set of Outputs.

In respect of *international migration*, the "Recommendations on Statistics of International Migration, Revision 1" (United Nations, 1998) are one of the key sets of guidelines for this technical subgroup. For *internal migration*, most analysts will still refer to the "Methods of Measuring Internal Migration" (United Nations, 1970), guidelines which have not been updated since their original publication.

Further, the Bureau of the Conference of European Statisticians (CES) has recommended that the *joint UNECE/EUROSTAT Seminar* (also called *Work Session*) on *Migration Statistics held at Geneva, 21-23 March 2005*, would serve as an appropriate forum to act as a focal point for information on international projects in migration statistics.

This seminar in Geneva brought together experts in the field the measurement and analysis of international and internal migration from a large number of international and national organizations actively involved from various perspectives and in various roles in the preparation of the 2010 worldwide round of population censuses. These organizations include among others:

- Conference of European Statisticians (CES)
- Council of Europe
- International Labour Office (ILO)
- International Monetary Fund (IMF)
- Organization for Economic Co-operation and Development (OECD)
- Statistical Office of the European Communities (EUROSTAT)
- United Nations Economic Commission for Europe (UNECE)
- United Nations High Commission for Refugees (UNHCR)
- United Nations Population Division (UNPD)
- United Nations Statistical Commission
- United Nations Statistics Division (UNSD)
- World Bank Group

and

- several national statistical offices (NSOs)

Particularly for countries with a relatively limited statistical capacity, the output of these expert groups will de facto constitute the leading guidelines for data collection in the areas of international and internal migration in the next population census. However, considerable efforts are also being made so as to ensure that statistically well-developed nations similarly adhere to the resulting recommendations and guidelines.

It is important, therefore, that such recommendations and guidelines be based on the most recent scientific insights in the area of the measurement and analysis of international and internal migration. *In essential areas this is not yet the case* in the currently existing recommendations and guidelines.

This paper provides a brief *explanation* of why this is the case. Further, it details *a number of specific elementary recommendations* with a view to improving this. These recommendations aim to ensure that the 2010 enumerations will result

- 1 in data on international and internal migration with the *highest informative value* on the migration processes which have actually taken place,

and

- 2 in data which will give analysts of migration the greatest *flexibility* to investigate these processes, their covariates and their consequences from angles which *suit local research objectives and needs*, while at the same time allowing *meaningful cross-national studies*.

In doing so, *this paper first of all addresses the above co-ordinating organizations, institutions and expert groups* involved in the process of developing guidelines

and recommendations in the area of the measurement of international and internal migration in the 2010 censuses.

However, the findings in this paper are equally important for *national statistical organizations* which will actually be taking and processing their own national 2010 population census, and for *analysts of data on international and internal migration*.

We begin by sketching the fundamental importance of population censuses for the measurement of migration.

## 2 THE IMPORTANCE OF POPULATION CENSUSES FOR THE MEASUREMENT OF MIGRATION

Population censuses attempt to measure wide-ranging sets of socio-economic variables on a national basis. For many of these variables, a sample survey would be a cost-effective alternative, yet yielding reliable information. However, for some topics this is not the case. Pre-eminent amongst the latter are the measurement of population stocks and flows.

As far as the measurement of *population stocks* is concerned, the reason is primarily methodological. Without an adequate record of the population stock, there would be no suitable sampling frame for any sample surveys focusing on population variables. The majority of countries do not have continuous population registration systems which could be used here instead, leaving the full census enumeration as the only alternative.

However, the rationale for measuring *population flows* through a full population enumeration is different. Population flows are multidimensional. The internal or international move made by a person has two elementary dimensions, namely the *time* of the move and the *path* from origin to destination. In addition there are at least two further fundamental dimensions, namely *age* and *sex*. This is principally because migration flows tend to be highly age and sex specific. Further, depending on the analytical framework, there may be yet more important dimensions, including covariates such as reasons for the move, educational level, employment status, marital and family status, nationality or citizenship, dwelling characteristics, and other socio-economic variables.

A comprehensive picture of population flows, even if they are disaggregated only by the elementary dimensions of time, path, age and sex, is difficult to capture reliably in a random sampling design due to the *excessive sampling variability* associated with such an intrinsically multi-dimensional phenomenon. It is not well possible to devise a stratification strategy which would substantially reduce the associated sampling error, unless there exists reliable prior information on the nature of the flows. Generally, however, such prior information is not available.

Therefore, except in the relatively limited number of countries which operate a reliable continuous population registration system in which residence is systematically recorded, a population census is the *only source of data from which a comprehensive image can be obtained on both internal and international migration*.

As pointed out, in this respect migration takes a comparatively *unique* position. While most variables measured in a population census also lend themselves to reliable measurement through sample surveys, this does not hold true for international and internal migration.

The implication of this is that the measurement of international and internal migration must be given special and careful consideration in population censuses, both in terms of the measurement instruments used and in terms of concepts and procedures.

Population censuses are usually conducted only once every ten years. Therefore, the decisions taken in respect of the measurement of migration for the 2010 round of population censuses are *critical* in the sense that they will have very long-term consequences. They will affect the ability to study migration until well into the 2020s, until the data from the subsequent 2020 round of population censuses will become available.

It is essential, therefore, that the recommendations and guidelines adopted for the measurement of international and internal migration in the 2010 global round of population censuses

- (1) be methodologically sound,
- (2) result in data whose informative value is maximal, and
- (3) result in data which allow for the accommodation of widely varying national and cross-national research objectives.

The currently existing United Nations principles and recommendations for population censuses and for the measurement of international and internal migration (United Nations, 1997, 1998 and 1970, respectively) do not meet these objectives in the best possible way.

In the remainder of this paper, we shall explore this issue in some detail. And we shall develop specific recommendations with a view to rectifying this.

### 3 THE MEASUREMENT OF INTERNATIONAL AND INTERNAL MIGRATION

In a number of important respects the standard guidelines on the measurement of international and internal migration, United Nations (1998) and United Nations (1970) respectively, reflect approaches to the *measurement* of migration which predate more recent advances in this area in terms of mathematical rigour, of demographic consistency and of analytical power.

For example, in a wide-ranging and detailed paper Xu-Doeve (2005) develops a rigorous mathematical underpinning for the demographic measurement and analysis of internal and international migration, firmly embedding this in the standard demographic paradigm of cohort and period analysis. In a logical development this work derives methods of measurement which maximize the information on both migrant stocks and migration flows. In addition, it results for the first time in valid methods of adjusting migration data for incompleteness. Such developments have important implications for the measurement of migration in the 2010 global round of population censuses.

The mathematical framework is fundamentally based on the distribution of durations of residence among cohort members, that is, among members of age groups as enumerated at the time point of the population count. This distribution describes precisely if, and if so, when cohort members experienced the event of a move into any given migration defining area (MDA) of destination, thus *fully describing the entire migration process in continuous time*.

The demographic measurement and analysis of migration thus essentially centres on acquiring knowledge of and insight into such duration distributions, and, for explanation, their relevant covariates. Formally this is identical to the fact that durations of life until death yield demographically consistent cohort death rates as a function of time and age; it is the key relationship embodied in classical life table construction. Equally, it is formally identical to the fact that durations until birth yield demographically consistent cohort birth rates as a function of time and age, by parity if so desired. Further, once such cohort rates have been obtained, then a cross-section of cohorts at a chosen time point also gives us the familiar period rates.

Here, we shall refrain from exploring the detailed mathematical framework underlying such demographic approaches to measurement. For full particulars, the reader is referred to Xu-Doeve (2005). We shall instead limit ourselves to outlining in brief and in a non-technical fashion the principal results and implications only in so far as they relate to the guidelines and recommendations for the 2010 global round of population censuses in the area of the measurement of international and internal migration.

Building on this, we derive *specific recommendations for the 2010 global round of population censuses in a number of principal areas where the current United Nations (1997) recommendations require revision.*

Let us begin by briefly reviewing the central issue, namely the elementary demographic method of measurement of international and internal migration, in some more detail.

First, on methodological grounds there is no fundamental difference between the measurement of international migration and the measurement of internal migration.

This is easy to see, since from a formal analytical standpoint a migratory move is defined by a change in place of usual residence from one migration-defining area (MDA) to another MDA. Methodologically, it is not material whether any such two MDAs are a part of one single country or not.

The implication of this is that there is *no formal difference in the measurement instruments to be used in a population census for measuring international and internal migration.*

Second, rigorous mathematical evidence shows unequivocally that the *most information-rich approach to the measurement of migration* is based on the recording of the *duration of residence*. For international migration this is the unbroken duration as a usual resident of the *country*. For internal migration, this is the unbroken duration as a resident in the current *place* of usual residence.

To obtain geographical detail, the questions on duration should be supplemented by a question on the *associated country and place, respectively, of previous usual residence*.

Together, such data on the duration of residence in the current place or country of usual residence and on the associated place or country of previous usual residence give us (partial) *migration life histories along cohort lines*.

The underlying mathematical theory of cohort change due to migration from which these results are derived is a formulation of cohort behaviour expressed in terms of stochastic Poisson processes which firmly belongs to the standard demographic paradigm (Xu-Doeve, 2005). While the mathematical theory might perhaps seem somewhat daunting to the mathematically less inclined, the rationale behind these measurement instruments is easy to understand, however.

The recording of current and previous places of residence together define the *geographical dimension* of the most recent migratory move made by a respondent. Measuring the duration of residence yields the *exact timing* of the move. Thus, such retrospective data can be rearranged *as if they were the result of a contemporaneous traffic count*: The enumerator can take imaginary position on the boundary between the migration defining areas involved at a sufficiently distant point in time in the past. (Whether or not any two MDAs are actually physically contiguous is immaterial for the sake of this argument.) Time is then allowed to progress to the present, and migrants are logged by date and path as they cross the boundary. During this entire period from the distant past up to the time point of census taking, the enumerator thus witnesses -- in virtual real time -- the actual migratory process as it takes place over time and space.

When results are adjusted for mortality in the period since arrival, such traffic count data and retrospective data on duration of residence and place of previous residence are in fact identical.

Because the full process of the making of migratory moves is observed in continuous time it will be intuitively obvious, and this is mathematically substantiated, that traffic count data allow the reconciliation

- (1) of migrant data and move data, and
- (2) of migrant stock data and migration flow (intensity) data,
- (3) of transition data over any given interval and individual move data in continuous time.

In other words, duration of residence data lead in an unambiguous manner to stocks of migrants, to flows of migrants in continuous time and to demographically consistent migration intensities (instantaneous migration rates) dynamically in continuous time. They resolve and reconcile the issues of the difference between migrants and migratory events and between transitions in discrete time and migrations (moves) in continuous time.

All results can be classified specifically by geographical flow, and by age, sex, and any other covariates allowed by the census records.

Importantly, mathematical evidence in particular *contradicts the commonly-held assumption*, held for example by UNECE (2005), §23, *that the best measurement of migration is obtained from a question on the place of usual residence at a specified date preceding the census, usually taken as a fixed number of years prior to the enumeration.*

Such transition questions in fixed time lead to irreconcilable differences between actual migrants and the information obtained on the migratory transitions.

Information on the geographical trajectory of migrants cannot be obtained, nor can demographically consistent migration intensities be derived.

In fact, while transition data can be derived from life history data on duration of residence, the reverse does not hold true: information on migratory moves cannot be derived from data on migratory transitions in discrete time.

The fact that the information value of data resulting from a question on the place of usual residence at a specified fixed date prior to the enumeration is *comparatively* limited, is also easy to see: Returning to our earlier enumerator, imagine that he or she takes one snapshot of the population in the various MDAs under consideration, that he or she then *ceases all observation* for a fixed period of say one or five years, after which one new snapshot is taken.

Matching the persons in this before-after design, and suitably adjusting for any intervening mortality, will then result in data identical to the data obtained from the retrospective census question on the place of usual residence a fixed number of years prior to the enumeration.

Clearly, and contrary to the earlier contemporaneous traffic count design, in such a before-after design *nothing of the migratory process itself* is being observed, *merely the net outcomes* after one or five years. Obviously, the true migratory behaviour of the enumerated persons cannot be ascertained with any certainty, nor -- since the process was not observed as it played out in continuous time -- can demographically sound migration intensities (instantaneous migration rates) be evaluated.

In addition, it is fundamental to note that duration of residence data are the only type of data which enable analysts systematically to *establish the degree of incompleteness of recorded migration data*. In the case of international migration, this will for instance include a major share of any illegal migrants.

Adjustment of recorded migration data for incompleteness is important. For example, Xu-Doeve (2005), who describes the estimation and adjustment procedure in detail, found that in the population census studied (Thailand, 1970) only between 25% and 50%, depending on the age group, of all male migrants who moved to Bangkok in the 12 months immediately preceding the census had actually been captured. In other words, the enumeration left between 50% and 75% of all recent male migrants to Bangkok unaccounted for. The completeness of the enumeration of migrants was worst in the most mobile age groups.

As mentioned, such estimates of and adjustments for incompleteness are possible only if duration of residence data have been collected. Data on the place of usual residence at a specified fixed date prior to the enumeration do not allow the derivation of such adjustments.

At the same time, when a question of the latter type is employed, then the smaller the number of years selected for this fixed interval prior to the census, the greater will be the degree of incompleteness of the observed migrant count. This is because migration-specific incompleteness of enumeration is strongly associated with the recentness of the migratory move.

This leads to the following recommendations:

RECOMMENDATION 1 International migration is most suitably measured by a question on the unbroken duration as a usual resident of the country, supplemented by a question on the previous country of usual residence (if the unbroken duration is less than the current exact age). These questions should receive *core topic* status.

RECOMMENDATION 2 Internal migration is most suitably measured by a question on the unbroken duration of residence in the current place of usual residence, supplemented by a question on the previous place of usual residence (if the unbroken duration is less than the current exact age). These questions should receive *core topic* status.

These two recommendations are *deeply fundamental to the demographically consistent measurement of migration*. Only such measurement instruments allow the analyst to measure and analyse the migration processes which have actually taken place in their full detail, in terms of migrant stocks, in terms of moves, in terms of migrant flows, and in terms of instantaneous migration rates.

Questions on the place of usual residence at a specified date in the past (prior to the enumeration) should not be given core status. There are in fact only two cases where a country might usefully opt to include such a question as a *supplementary topic*.

The first case is if continuity is considered important with respect to earlier censuses where migration was measured using a question on the place of usual residence at a specified date in the past. This may, for instance, be relevant if significant investments have been made earlier in terms of forecasting instruments (software, expertise building, and so on) designed around the concept of demographic accounts.

The second case is where a country wishes to explore heterogeneity in migration behaviour by distinguishing between frequent and non-frequent movers. The preferred approach here is to ask recent migrants about their one-but-last move, as we shall outline below. However, a less detailed indication of such heterogeneity can also be obtained by studying whether for recent migrants the place of previous usual residence differs from the place of usual residence at a not too distant fixed time point in the past, such as one or two years prior to the enumeration.

However, in the latter case, the choice of the width of the fixed time interval chosen places an important a priori constraint on what constitutes a frequent migrant. Methodologically, such a priori constraints built into the operational definition concepts is less than optimal. Asking recent migrants about their one-but-last move does not impose any such constraints.

Further, for the study of international migration, *additional* questions on country of birth, country of birth of parents, nationality or citizenship, and dominant language used in the household, while possibly sensitive in some contexts, remain of value. Although such questions themselves are *not* suitable for the proper measurement of *actual international migratory processes*, they add useful insights into associated phenomena. The fundamental question to be used as an instrument for the measurement of international migration itself, however, has to be the duration of residence since entry into the country.

In the next section, we shall discuss a number of detailed aspects of these two recommendations in greater depth. This in turn will lead to several further recommendations which are of major significance for the measurement and analysis of international and internal migration.

#### 4 CONCEPTS AND PROCEDURES FOR MEASUREMENT AND ANALYSIS

The recommendations of the United Nations for population censuses (United Nations, 1997) suggest a very wide range of socio-economic items for inclusion in national population censuses. As discussed above, many of the variables included as a matter of routine equally lend themselves to satisfactory measurement through *sample surveys*.

Sample surveys are generally more cost effective, and they can be repeated at shorter intervals more easily. This results in measurements at more distinct time points for the variables in question. For many socio-economic variables such higher resolution time series are of considerable informative value. Population censuses allow the study of developments over time only through suitably designed retrospective questions. Most census questions are instantaneous, however. So, while most socio-economic information obtained through a census enumeration only results in static images or at best in trends based on time points ten years apart, socio-economic time series data from repeated sample surveys additionally allow the study of dynamics over time.

Further, measuring selected items through sample surveys could help in substantially reducing the scope of the information coverage in the full census, while at the same time allowing for increased attention to those variables, including population stocks and population movements, for which a full enumeration is indispensable.

Such increased attention can manifest itself both in terms of *substantive contents* and in terms of *measurement quality*. As we shall see below, even when measuring migration at the most elementary level, limited increases in substantive contents would already provide valuable additional insights.

Regrettably, however, the use of sample surveys as an alternative to a full enumeration where this may be equally or even more appropriate is an issue which is *not systematically explored by United Nations (1997)*.

We therefore have the following recommendation:

**RECOMMENDATION 3** It should be assessed which census core topics and which census non-core topics can be adequately or even more suitably measured through periodic sample surveys. These topics should be removed from the list of topics to be considered for inclusion in population censuses.

For countries with limited experience in conducting sample surveys, appropriate training in this area is recommended.

Since, as explained above, sample surveys are not suitable if it is desired to gain a comprehensive insight into international or internal migration, we shall concentrate on population censuses in the remainder of this paper.

Methodologically, a population census takes an extraordinary position in the scientific quest for knowledge. It is a socio-economic data collection effort which is conducted without a *full and explicit* prior definition of specific research *needs, objectives and methods*. In many respects, it is a general purpose data collection process, whose *raison d'être* in terms of data coverage is at least in part explained by a desire for process continuity, formalised by international co-ordination and by national legislation and institutions.

This lack of a fully justified and explicit methodological research framework makes itself felt in the area of migration when it comes to the recommendations in United Nations (1997) in respect of *concept operationalization*, that is, in respect of the definition of more or less abstract concepts in terms of measurable variables.

In demography, for example, the use of indirect methods of estimating mortality and fertility rates is well established and mature. As a consequence, we see that the measurement of appropriate concepts through population censuses in countries with poor vital statistics registration systems is quite well operationalized in United Nations (1997).

By contrast, however, unfortunately the same cannot be said for the measurement of international and internal migration. Here mathematically rigorous and demographically consistent methods of measurement are much more recent and by no means common practice yet. Thus, in the absence at the time of writing of an awareness of one or more coherent sets of clearly defined research methods that constitute best methodological practice, we see that the operational definitions in United Nations (1997) of relevant concepts in the area of international and internal migration leave to be desired in several important respects.

The same also applies to United Nations (1998), which was taken as the principal guideline for recommendations on the measurement of international migration in United Nations (1997).

The operational definitions of several elementary concepts in the study of migration, such as the concepts of migrant, of place of usual residence and of duration of residence, are *methodologically problematic*. Mostly, this takes one of

two distinct forms, namely the unnecessary *narrowing down* of intrinsically broader concepts on the one hand, and the unnecessarily *unfocused* operational definition of intrinsically precise concepts on the other.

When there is no adequate a priori reference to a rigorous and coherent methodological framework for analysis leading to logical and compelling operational definitions and data specifications, the boundary between well-specified operational definitions and restrictive or imprecise operational definitions easily becomes blurred. In some instances this also leads to undesirable *ambiguity* in operational definitions.

As a consequence, at the same time the demarcations between the realms and responsibilities of enumerators, data processors and data analysts become fuzzy.

The ultimate effect of such methodologically problematic specification is that it *restricts the scope and flexibility of the analyst* to explore the full nature, patterns and developments over time of the international and internal migration processes which have actually taken place. In other words, poor operational specification of concepts places avoidable constraints on the analytical use which can be made of a costly census enumeration.

In the following sections we shall discuss the three most important examples of problematic operationalization which have an important bearing on the measurement and analysis of international and/or internal migration. They concern the operational definitions of the concepts of *migrant*, of *place of usual residence*, and of *duration of residence*.

On the basis of this discussion we shall derive one or more specific operational recommendations in each case with a view to cost-effectively maximizing the analytical and informative value of population census data on migration in the 2010 global round of population censuses.

#### 4.1 THE CONCEPTS OF MIGRANT AND MIGRATORY MOVE

It is common practice among national statistical organizations to specify a minimum duration of usual residence for international migrants before *de facto usual residence* in the country of enumeration is also classified *de iure* as usual residence in the country. A 12-month's term is widely accepted as the minimum residence duration here (UNECE, 2005, §14). If the actual duration is less than the required minimum, then the international move is not recognized as such. The immigrant is not then considered to belong to the enumerated population.

Minimum residence durations of up to several months are also sometimes required by national statistical organizations for an internal migratory move to be recognized as such (Xu-Doeve, 2005). If the duration is less than the specified minimum, then the enumerated move is cancelled administratively. It is a practice whose detail is often not very well documented.

It is interesting to note, however, that the United Nations (1997) recommendations do not lay down any general requirement as to any minimum usual residence duration, neither for international migration nor for internal migration.

There is, however, one paragraph, §2.66, which clearly suggests an awareness of the practice. It states: "The place of usual residence is where a person usually resides and it may or may not be the person's current residence or legal residence. The latter terms are usually defined in the laws of most countries and need not correspond to the concept of place of usual residence which, as employed in the census, is based on conventional usage. In published reports, countries should indicate whether or not household information refers to usual residents and also what the time limits are in respect of being included or excluded as a usual resident."

Note, incidentally, the operational definition given here of the concept of usual, defined tautologically in the opening line and subsequently in terms of the methodologically rather nebulous phrase "based on conventional usage". It is a definition to which United Nations (1997) makes no further reference. It is in particular remarkable that in §2.20ff where the concept of place of usual residence is first defined, no mention is made of this further operational development of the concept in this §2.66. Usual as used in usual residence is an important concept, and we shall return to this concept in detail, below.

On the issue of the use of minimum time limits in the operational definition of the concept of migrant, the recommendations provide no operational guidelines at all, with only one minor exception. This exception concerns the second of two relatively small classes of special cases where doubt may arise about the true

place of usual residence. (The first class of cases concerns those individuals who can plausibly be argued to have more than one place of usual residence.)

The United Nations (1997) recommendations define this second class of cases as follows: "Problems may also arise with persons who have been residing at the place where they are enumerated for some time but do not consider themselves to be residents of that place because they intend to return to their previous residence at some future time, and also with persons who have left the country temporarily but are expected to return after some time." (§2.22)

The recommendations then continue in the same paragraph: "In such instances, clearly stated time limits of presence in, or absence from, a particular place must be set, in accordance with the prevailing circumstances in the country, to determine whether or not the person is usually resident there."

The methodologically sound position is that a *change in location of usual residence* is both a *necessary* and a *sufficient* criterion in defining operationally what constitutes a migratory move. A migrant is operationally defined as a person who experiences such a move. Any additional criteria, be they a minimum duration of residence in the location of destination or any other conditions, should be no part of the operational definition.

From §2.66 we have to conclude that that this is not the position taken by United Nations (1997), which clearly remains at best equivocal on the issue.

The negative implications of operationally specifying additional criteria can be illustrated by examining the special case of §2.22 quoted above where United Nations (1997) explicitly suggests the imposition of minimum time limits.

Specifically, intentions and expectations to return in themselves are not valid grounds to raise doubt as to whether or not one can speak of usual residency at the location of enumeration. Migration analysts recognize the concepts of *circular* and *return migrants* as significant special subgroups amongst migrants. Setting minimum durations of presence or absence here, as suggested by United Nations (1997), would arbitrarily eliminate *frequent movers* in these categories from view.

The issue leading to this problematic §2.22 is a result of the fact that United Nations (1997) adopts a passive approach on the part of the researcher to the operational definition of the concept of *usual* within the broader concept of place of usual residence. Except for those persons who can plausibly be argued to have more than one place of usual residence, United Nations (1997) lets the enumerated individual decide: "The place of usual residence is the geographical place where the enumerated person usually resides. This may be the same as, or different from, the place where he or she was present at the time of the census or his or her legal residence. ... most persons will have no difficulty in stating their place of usual residence ..." (§2.20-2.21).

This is a form of object self-measurement. The subsequent modification of this operational definition in §2.66 which we encountered earlier does not materially alter this fact. Methodologically, self-measurement is not optimal since it is a measurement procedure which is not standardized. It follows that scientifically the resulting data cannot be compared from individual to individual. We shall return to this issue below, where we shall formulate a specific recommendation addressing this problem. However, before we do so, it is useful to explore some further aspects of operational approaches to the definition of the concepts of migratory move and of migrant.

It is important to reiterate that the widespread practice among national statistical organizations of applying blanket minimum duration of residence requirements is ill-advised on methodological grounds. Among all those who have experienced the event of a move it creates an arbitrary distinction between those who are statistically recognized as migrants and those who are not. And it puts the national statistical organization rather than the migration analyst in the position of arbiter in this matter.

Linking the two concepts of usual and of residence duration through a rule specifying an arbitrary minimum residence requirement constitutes definitional overspecification. There is *no basis in logic* to conclude that some residence duration would constitute a necessary condition for residency to be usual.

Such overspecified operational definitions of the concept of usual residence lead for example to the fact that statistically, by definition, there has not been any migration in the specified number of months leading up to the enumeration. Both *recent migrants* and *short-term migrants* are administratively reclassified as non-migrant residents of their respective previous usual place or usual country of residence. Even though properly observed as migrants during data collection, their status as having experienced a migratory move is subsequently annulled by the census administrators.

It is worth observing that, while United Nations (1997) has taken United Nations (1998) as its guideline in the area of international migration, there are major discrepancies here between these two sets of United Nations recommendations. Contrary to United Nations (1997), United Nations (1998) explicitly specifies minimum usual residency requirements for a person to qualify as a migrant.

Note that there is also a subtle difference here in the operational definition as compared to the approach described earlier. In the case of United Nations (1997, §2.22), the duration is used to determine whether residency can be qualified as usual. In the case of United Nations (1998), there are two conditions that must be

met separately and jointly, namely both a minimum duration of residence and usual residence.

Usual residence in United Nations (1998) is defined as follows: "A person's country of usual residence is that in which the person lives, that is to say, the country in which the person has a place to live where he or she normally spends the daily period of rest. Temporary travel abroad for purposes of recreation, holiday, business, medical treatment or religious pilgrimage does not entail a change in the country of usual residence." (§32)

Clearly, while a connection is specified with residency duration through the term "temporary", this is explicitly made conditional on the *purpose*. Although not truly operational yet, this definition in itself is methodologically not problematic. We note, however, that in the subsequent §33, United Nations (1998), quoting United Nations (1997), also describes the self-measurement approach which we criticized above.

The requirement of a minimum duration of residence is now made separately of the semi-operational definition of usual residence. Specifically in fact, United Nations (1998), in §36 and §37, goes one step further in predefining a difference between long-term and short-term international migrants. A usual residence duration of at least 3 months is required to qualify for short-term migrant status, and long-term migrant status is awarded instead if the person in question has been usually resident for a duration of at least 12 months in the country of immigration.

Further, short-term migrant status may be statistically denied depending on the *purpose* of the stay. The exhaustive list of purposes which are not recognized as valid for short-term migrant status comprises: "recreation, holiday, visits to friends and relatives, business, medical treatment or religious pilgrimage" (§37).

Note that, contrary to §32, visits to friends and relatives are now also included as a disqualifying purpose, as they are in the definition of country of usual residence when this definition is repeated in "Box 1" on p 18. For long-term migrant status there are no such disqualifying purposes.

It is useful to observe that, following this operational approach, the measurement of short-term migrants necessarily requires *closure* of the interval between migratory moves. That is, at least two successive moves must have been observed so as to be able to differentiate between short-term migrants on the one hand and long-term migrants who have arrived only recently on the other.

The standard procedure in a population census is to measure durations of residence in connection with a single migratory move only, namely the most recent one. So this allows the determination of recent migrants, but not of short-

term migrants. Thus, the proper measurement of the latter requires fuller information on the migration life history of individuals.

Further, methodologically, there is another important issue at stake when using such operational definitions of short-term versus long-term migrants. By classifying the data into an unnecessary predefined analytical format they violate the principle to "let the data speak". The data should inform the analyst, first, whether there exists any heterogeneity in a cohort which can be qualified as long-term versus short-term migration, and if so then, second, where the cut-off point actually lies -- at three months or at some other duration.

In other words, it should be left to the migration analyst, using methodologically appropriate approaches to data analysis, to make out if any such heterogeneity is present in the data and, if so, what its precise characteristics are.

In summary, given a proper mathematical-analytical framework for the study of migration, such overspecified operational definitions are unnecessarily restrictive. They a priori deny the migration analyst who uses the resulting data essential *scope* and *flexibility* in the investigation of the full and true nature and characteristics of the migration processes which have actually taken place over time.

By operationally defining a minimum duration of residence requirement, recent migration and any developments in recent migration are ignored by definition. Mobile sections of the population remain disproportionately hidden from view. Any existing heterogeneity in terms of frequent and non-frequent movers in the population is thereby also obscured.

Further, as a consequence, the study of relevant covariates to assess the impact of frequent movers on the migration defining areas of origin and destination is rendered impossible.

Finally, the omission of United Nations (1997) to be specific in this area, and the lack of uniformity between United Nations (1997) and United Nations (1998) lead to unnecessary ambiguity which hinders cross-national comparability of data and findings.

A further fundamental methodological issue is this. As will be clear from our earlier discussion, in its essence the measurement of international and internal migration fundamentally resolves to the measurement and analysis of durations of residence distributions. A priori *restrictively* defining concepts operationally in terms of measurements yet to be taken mortgages and prejudices the migratory process as it will be observed. This is methodologically problematic.

The above arguments therefore lead to the following recommendation:

RECOMMENDATION 4 The operational definitions of the concepts of migratory move and of migrant in terms of a required minimum duration of residence or usual residence at the destination should be rejected. Instead, the operational definitions of these concepts should be based exclusively on a change of usual residence.

Obviously, this recommendation calls for an explicit operational definition of the concept of usual within the broader concept of usual residence. Clearly, further, we need a definition which does not avoidably prejudice the work of the migration analyst. In particular also it should allow for recent and short-term migrants while excluding short-stay non-migrants, such as holiday makers, whose usual residence is elsewhere.

Considering the above discussion, the recommendation here is as follows:

RECOMMENDATION 5 An exhaustive international standard list should be drawn up of purposes of stay which disqualify a place of residence as the place of usual residence. The purpose of stay should be ascertained and the list should be applied exclusively in respect of those persons whose uninterrupted duration of residence is less than one month.

By implication, an uninterrupted duration of residence of one month or over automatically qualifies residency as usual. *If -- and only if --* the duration is less, then the *purpose* of stay is the discriminating criterion in determining whether the actual residence is the usual residence.

This reflects the position that a person with an uninterrupted duration of residence of one month or more contributes significantly to the local economy in terms of expenditure for accommodation, living expenses, the purchase of goods and services, and so on, moneys which are permanently withdrawn from disbursement in the previous place of residence.

Purposes which, for example, represent some form of (search for) an economic basis of existence, education, military service, retirement, and family reunion should, of course, qualify.

The exhaustive international standard list of disqualifying purposes, reasonably indicating that one's usual residence is elsewhere, could comprise the following items: recreation, tourism and holidays; business travel ; temporary visits to friends and relatives; temporary medical treatment; and religious pilgrimage.

By employing a list which is a standard internationally, the collection of data which are also comparable cross-nationally is facilitated.

Further, rather than ignoring recent migration by definition and rather than predefining short-term and long-term migration, it is valuable to allow for the exploration of just these categories in more detail.

Often, for example, recent migrants take specific positions on the labour and housing markets and in terms of services required, such as education and health care. Recent migrants also tend to include disproportionate numbers of illegal international migrants. In addition, methodologically, data on recent migrants are essential in the study of the incompleteness of enumeration of migrants.

Also, shorter-term migrants if present are likely to constitute one or more subgroups with specific characteristics which may well be significant. The investigation of such groups deserves to be facilitated, as well. In particular, we therefore need the data which allow us to measure if there are any shorter-term migrants in an enumerated cohort, and if so, what the characteristics of their migration behaviour are.

Therefore we have the following recommendation:

**RECOMMENDATION 6** To capture the essence of high-mobility sections of the population, the life history measurement of recent arrivals, defined as all those who have arrived in the current place of usual residence within the past 5 years, should be extended to include not just the last move but also the last-but-one move. This should be done using similar measurement instruments, that is, duration of residence in combination with associated previous usual residence.

Analytically, capturing all earlier migratory moves for all cohorts would clearly be preferable, since this allows a full migration analysis from birth to the time point of the enumeration for the entire population. However, this would place a heavy burden on the enumeration.

In order to strike a balance between what is possible and what is reasonably feasible at very little overall additional effort, this recommendation limits the observation of migration life histories in two ways.

First, the measurement is limited to recent migrants only. Second, within this subgroup, the measurement is limited to data on one additional earlier move only. This will be sufficient to explore some of the most pronounced features of the issue addressed by this recommendation.

For the purpose of this recommendation only, recent is defined as having arrived in the current place of usual residence within the 5 years leading up to the time point of the enumeration. A shorter interval than 5 years is not recommended here. This is, because the more recent the date of arrival, the higher the migration-specific underenumeration rate, resulting in migrant data which may be highly skewed in terms of covariates, including but not limited to residence duration, geographical trajectory taken, age and sex.

This completes the review of operational definitions of the concepts of migrant and migratory move. Next, we shall focus on the concept of place in place of usual residence.

## 4.2 THE CONCEPT OF PLACE OF USUAL RESIDENCE

In the previous section, we discussed the concept of usual within the broader concept of place of usual residence. It led to a specific recommendation. Subsequently, residence is always understood as usual residence as defined operationally in that section. In the present section we shall concentrate of the concept of place. As we shall see, here too, the United Nations (1997) recommendations lead to operationalizations which are methodologically problematic.

Clearly, place of residence is a fundamental concept in the measurement of migration. The set of places together constitute the elementary building blocks of the migration defining areas (MDAs) used by the analyst of migration to specify origins and destinations.

Here, the methodological issue is somewhat more complex, as the ultimate operational definition is a result of a priori elementary definitions as well as of elements of the census process, in particular of enumeration, of coding and of tabulation.

United Nations (1997, §1.98ff) provides some elementary guidelines as regards the identification of the smallest area by which data are classified. In most of the discussion in those paragraphs, the *enumeration district* (enumeration area) is used. However, there is some discussion of the application of geo-referencing (geocoding) based on the use of GPS (Global Positioning System) data as an alternative. Yet, here too, such data are used to reference larger areas, either a collection of units defined as segments or block faces allocated to a single nodal geographical co-ordinate, or areas whose definition is based on a uniform square grid system in which the national territory has been subdivided.

The ability to aggregate such smallest areas into, or match such smallest areas with, existing administrative divisions (minor civil divisions, villages in rural areas, and so on) is stressed.

Each of these two basic approaches raises some minor methodological issues. For example, when the enumeration district is used, the operational census process, rather than specific research needs and objectives, is taken as the defining criterion in the identification of the smallest areal units. Further, although this is desirable, in practice enumeration districts are not necessarily stable from census to census, hampering intercensal comparison.

Allocation of units to geographical nodes assumes some form of a priori classification which too may not agree with specific research needs and objectives. In the recommendations, square grids are left undefined in respect of their size, and particularly larger grid squares may be difficult to reconcile both with administrative divisions and with specific research needs and objectives.

Further, administrative divisions are not necessarily stable, especially in urban and rural areas which are characterized by dynamic economic and demographic conditions. If the (sets of) smallest areal units, be they enumeration districts or GPS based areas, are designed to match administrative divisions, then the instability of administrative divisions may also affect intercensal comparability.

Another issue is that respondents will not be able to identify places of residence in terms of enumeration areas or GPS based areas, since generally the use of such areas is for internal purposes of the national statistical organization only. In practice, this may in particular be problematic in respect of the accurate measurement of places of previous residence.

Finally, the delineation of enumeration areas is not subject to any form of international standardization, so that any meaningful cross-national comparison of data classified by such areas is impossible.

In addition to these elementary areal units of enumeration, the United Nations (1997) recommendations use two further classes of smallest areal units. They are the locality and the minor civil division.

A *locality* is defined as follows: "For census purposes, a locality should be defined as a distinct population cluster (also designated as inhabited place, populated centre, settlement and so forth) in which the inhabitants live in neighbouring sets of living quarters and that has a name or a locally recognized status. It thus includes fishing hamlets, mining camps, ranches, farms, market towns, villages, towns, cities and many other population clusters that meet the criteria specified above." (§2.49)

The term *minor civil division* is generally used to describe the lowest-level official areal unit recognized in a country for administrative purposes. Confusingly, it is also sometimes referred to as the *smallest civil division*. United Nations (1997, §2.50) notes: "Localities as defined above should not be confused with the smallest civil divisions of a country. In some cases, the two may coincide. In others, however, even the smallest civil division may contain two or more localities. On the other hand, some large cities or towns may contain two or more civil divisions, which should be considered as segments of a single locality rather than separate localities."

These areal unit concepts also raise methodological problems. The operational definition of locality excludes any size criterion. By definition localities may range from farms and hamlets up to towns and cities. Clearly, using the concept of locality may well inadvertently result in inappropriate apple-pear type, or rather cherry-pumpkin type, comparisons.

Further, the operational criterion of having a name is problematic. For example, the use of names frequently differs between locals and non-locals. Often, locals use more finely disaggregated, informal or historical names which will not be universally recognized by outsiders. Also, some locals may refer to formal administrative names, while others might refer to such locally common names. The areas thus referred to may in fact well be different. Educational level and social status may play a role in such usage of names, leading to undesired extraneous heterogeneity in the measured data.

The alternative operational criterion of a locally recognized status is even more problematic, since in a census this is practically immeasurable unambiguously. In fact, allowing for two alternative, equivocal and possibly partially overlapping operational criteria, namely that of having a name and that of having a locally recognized status, makes for a definition which has little or no methodological value in research into migration.

Further, as observed in §2.50 there is no unambiguous relationship between the concepts of locality and of minor civil division. Using the two concepts side by side is therefore problematic.

Methodologically, administrative units may well be unsuitable for the study of migration. Generally, the subdivision of a nation's territory in civil divisions is primarily based on historical, political and administrative considerations, more so than on criteria which are relevant to the study of migration. From case to case, this may be convenient or inconvenient, depending on the specific research needs and objectives.

As mentioned, administrative divisions are not necessarily stable over time, hampering intercensal comparison. And as there is no international standard for the administrative subdivision of a country, cross-national comparisons using such areal units are not generally feasible.

In addition to the above three classes of areal unit, many countries use some form of operational criterion to classify the national territory using the concepts of *urban and rural*. It is a classification frequently used by analysts of migration. However, here, too, there is no internationally agreed standard for such classification (United Nations, 1997, §2.52). This renders cross-national comparisons using this classification methodologically problematical.

Also, the United Nations recommendations make no systematic attempt to reconcile this classification with enumeration areas, localities or civil divisions (United Nations, 1997, §2.52ff).

Apparently, the battery of mutually overlapping definitions of elementary areal units confounds even the authors of the recommendations when in §2.35 they define the duration of residence for the purpose of the measurement of internal migration as follows: "The duration of residence is the interval of time up to the date of the census, expressed in complete years, during which each person has lived in (a) the locality that is his or her usual residence at the time of the census and (b) the major or smaller civil division in which that locality is situated."

Assuming that both criteria must be satisfied, then the geographical criterion would appear to be simply the civil division encompassing the locality of usual residence. Ambiguity remains, however, particularly in the phrase "major or smaller civil division". Possibly the intention is the "lowest-level civil division" which comprises the entire locality in question, but this is by no means clear.

Apart from the imprecise and equivocal nature of such operational definitions as in §2.35, they also restrict the scope and flexibility of the migration analyst. This happens for a number of reasons.

First, an unnecessarily crude regional subdivision is adopted for the measurement of migration. Administrative divisions, even the lowest level ones, may harbour considerable internal heterogeneity. This is the case both in rural areas and in urban areas.

Further, any internal migration within such administrative divisions remains hidden from view. Consider for example the case where a locality is a city. Internal migration within major cities can be significant, both in terms of size and in terms of socio-economic and cultural implications. Yet, by the above operational definition, the administrative division covers at least the entire city, so that any intra-city migration remains unobserved.

More generally, it is well-known that most internal migration consists of migratory moves over relatively short-distances. The use of large areal units in the measurement process may thus arbitrarily obscure an important part of migratory movements in the population. And it does so selectively by disproportionately filtering out the movements over the shortest distances.

Also, as mentioned, the boundaries of administrative divisions are not necessarily stable over time, hindering comparison between successive censuses. And by using administrative divisions as the criterion of geographical delineation,

international comparison is severely impeded, since such divisions are notoriously difficult to compare from country to country.

Similar problematic issues arise in respect of the operational definition of the previous place of usual residence. United Nations (1997) defines in §2.38: "The place of previous residence is the major or smaller civil division, or the foreign country, in which the individual resided immediately prior to migrating into his or her present civil division of usual residence."

For internal migration, the choice between "major" and "smaller" civil division is left indeterminate. Selecting either, however, can potentially lead to very significant uncertainty as to the precise origin, depending on the administrative organisation of a country.

The destination is defined operationally yet more ambiguously here as merely a civil division, without being specific. Potentially, this leaves a national statistical organization the option to select large areal units as destination areas, masking any internal heterogeneity and obscuring any intra-divisional migration.

Clearly, methodologically, using any prior regional classifications is fraught with difficulties. Therefore, and with a view to obtaining data with the highest information value, measurement of places of usual residence should to the maximum extent possible be *independent of any prior regional classifications*, be they enumeration areas, localities, any level of civil division, or pre-defined urban-rural dichotomies.

This can be achieved easily and highly cost-effectively by *geo-referencing* individual places of residence. For the current place of usual residence, geo-referencing can be done by taking a GPS (Global Positioning System) field measurement at the time of the enumeration.

This is, of course, not directly possible in cases where the enumeration does not take place at the current place of usual residence, if the census questionnaires are mailed, or if the census is administered through the internet. Equally, this is not directly possible for any previous places of usual residence.

However, in most countries, there are well-developed and commonly-used systems of marking places of residence. The best known and most widely used is the *address system* as used for postal purposes. Where this is the case, it is straightforward to obtain the GPS data, namely by cross-referencing addresses and GPS co-ordinates.

A second-best alternative is the postcode system which many countries use, provided it is sufficiently precise. The latter may, however, not always be the

case. While in towns and cities precision is common down to street-segment or block level, in rural areas, the postal zones may be considerably cruder. Postcodes may, however, be useful in their own right, since they allow for automated validation checks on the address data.

Thus, in the absence of GPS field readings, measuring *full address data* including postcodes (where they exist), both *for current and previous places of usual residence*, is the preferred approach.

In rural areas in developing countries and in parts of urban areas, such as squatter settlements, the exact place of residence (living quarters, housing unit, and so on) may not have any formal address. However, there will almost always be a nearby regular *postal delivery point*. Using the co-ordinates of that point will generally be more precise than any of the regional units suggested by the United Nations (1997) recommendations.

Only in cases where both the own residential address approach and the postal delivery point approach fail, then recourse can be taken to recording the place of current usual residence by enumeration area, and place of previous usual residence by either named settlement or lowest-level civil division, *whichever is geographically more precise*.

By 2010, the use of GPS readings and automated cross-referencing of address details with GPS data must be assumed a standard approach in census operations globally.

This allows census information systems to include properly *geo-referenced data on the places of current and previous usual residence for all enumerated persons*.

Such information once and for all removes the methodological restrictions and rigidity imposed by the use of unnecessarily crude, ambiguous and partially overlapping sets of areal units by which persons are classified in the census recording process, areal units, moreover, which have been predefined for reasons other than the investigation of migration in the population.

Only such a method which aims to record current and previous places of usual residence independently of any prior regional classifications allows the analyst to study the *full detail* of the migration processes which have actually taken place.

And it enables the data user flexibly to compile migration defining areas (regional classifications) which are *tailored* in a dedicated fashion to suit the specific research needs and objectives of the investigation in question.

In addition, only such data enable the researcher to define tailor-made migration defining areas will allow both proper *comparability between successive censuses*

of observed migration behaviour within the population and meaningful *cross-national comparisons* of the migration behaviour experienced by different populations.

Only such information would allow, for instance, the study of immigrant ghettoization of specific neighbourhoods within cities, or the cross-national study of rural-urban migration using a uniform standard for the definitions of urban and rural across countries rather than the often incomparable national definitions.

Therefore we have the following recommendation:

RECOMMENDATION 7 For every enumerated person both the place of current usual residence and any places of previous usual residence measured, if the latter are within the country, should be properly geo-referenced in the census database using GPS (Global Positioning System) co-ordinate data.

Clearly, such data should not be made available to researchers of migration at the level of the individual person. It is an essential task of the census administration to protect the privacy of the enumerated individuals. Researchers should, however, be in a position to specify migration defining areas as a function of specific research needs and objectives. The national statistical organization should then aggregate the data in agreement with such specifications, such that the privacy of the enumerated individuals is safeguarded.

In terms of *international migration*, the current United Nations recommendations regarding the place of previous usual residence (United Nations, 1997) merely suggest the recording of the *country of origin*. Consideration could be given to be *more specific* as to this origin.

This may be particularly useful in respect of countries which are known to be important sources of international migrants, either in terms of numbers and/or in economic or socio-cultural terms. This would give a valuable additional and deeper insight into the nature of the most important international flows of migrants.

This completes our discussion of the concept of place of usual residence. In the next section, we shall focus on the important concept of duration of residence.

### 4.3 DURATION OF RESIDENCE

The concept of duration of residence is fundamental to the measurement of international and of internal migration. If the operational definition of this concept is methodologically flawed, then this can not only impede the work of the migration analyst, it can in fact render the measurement of migration itself quite impossible.

The operational definition of this concept given by United Nations (1997) for the measurement of international migration differs from the one given for the measurement of internal migration. Earlier, we already discussed aspects of the operational definition of duration of residence for the purpose of measuring internal migration. We repeat the definition here for convenience:

"The duration of residence is the interval of time up to the date of the census, expressed in complete years, during which each person has lived in (a) the locality that is his or her usual residence at the time of the census and (b) the major or smaller civil division in which that locality is situated." (United Nations, 1997, §2.35)

The important point to note now is the clause "expressed in complete years".

For the purpose of measuring international migration, the operational definition is different:

"Recording the calendar year and month of arrival of a foreign-born person to the country of enumeration permits the calculation of the number of completed years between the time of arrival in the country and the time of inquiry, usually the census date." (§2.255)

The use of the qualifier "expressed in complete years" for internal migration as against "the number of completed years" for international migration leads to unnecessary ambiguity. Demographers use the term "in completed years" more usually when expressing ages. Here it is understood that the magnitude of the variable is given ignoring any fractional part; that is the value is always rounded down. The term "complete years" does not have such a connotation, however. Rounding to the nearest natural number then is the more obvious approach. A rationale behind the difference in approaches to rounding in United Nations (1997) is not given, however.

Next, United Nations (1997) continues, exclusively for the purpose of measuring international migration, that: "Information on time since arrival can also be collected by asking how many years have elapsed since the time of arrival, instead of in what calendar year the person arrived. However, use of such a question is not recommended because it is likely to yield less accurate information." (§2.257)

There is no such recommendation for the purpose of measuring internal migration, however.

Further, clearly, the assumption underlying §2.257 is that an international migrant can recall the month of arrival. It seems only reasonable then to assume that internal migrants possess equal powers of recollection, certainly in terms of the most recent move.

This is important, because as we saw it is the distribution of cohort members by durations of residence which holds the fundamental information which allows the analyst to recover migrant stocks, migrant moves, migration flows and instantaneous migration rates as a function of continuous time. It is also this distribution which lies at the heart of procedures for the adjustment of migration data for incompleteness.

The greater the *precision* with which true empirical durations of residence are observed and recorded, the greater the power of analysis.

In addition, the more precise the analyst's knowledge is of the detailed form of the true empirical distribution of residence durations, the better the ability to detect any heterogeneity within cohorts in respect of migration behaviour. This, too, is of considerable significance, since it has, for example, often been observed that frequent movers constitute a distinct subgroup among populations with their own specific migration behaviour.

Thus, apart from the usual issues of validity and reliability, the key issue in operationalizing the concept of duration of residence is the *precision* of the *measurement instruments* and of the *measurement scale* used.

This therefore leads to the following recommendation which is fundamentally important to the measurement of international and of internal migration:

**RECOMMENDATION 8** It is imperative that true durations of residence be measured with as much precision as possible during the actual enumeration process through the use of appropriate measurement instruments and techniques. A measurement scale which is precise down to year and month of arrival is satisfactory.

The selection of measurement instruments and techniques which are most *appropriate* in terms of *validity*, *reliability* and *accuracy (precision)* is likely to be -- at least in part -- culturally determined. This is a key matter which requires careful assessment based on solid comparative methodological research.

Thus, it is recommended that countries systematically explore which structure and phrasing of the questions used to measure durations of residence will result in measurements with optimal validity and reliability while at the same time yielding the highest possible degree of accuracy (precision), and how to administer those questions with best effect.

What is found to be most appropriate in one cultural context may well prove to be inferior in other cultures and countries. For example, it is likely that the optimal phrasing will vary from country to country and from culture to culture.

As far as the measurement scale is concerned, if detail by month is considered excessive, given for example limited calendar awareness amongst a population, then alternative, second-best, measurement scales more appropriate to the local conditions should be sought.

Examples could be seasons in countries where they markedly occur; or demarcations of time by important events such as spring, moon and other festivals in China and elsewhere. Since the latter are based on the lunar calendar, conversion would be needed. It is, however not necessary that any such demarcations are stable from year to year in the demographic calendar, so long as they are clearly defined for each individual year.

A related issue is that of making the observed data on residence durations available for research. Here, too, avoidable methodological bottlenecks may be created in the census process.

With its particular focus on short-term and long-term migrants discussed earlier, United Nations (1998) stresses classification of durations of residence, expressed in months, as follows: [3, 12), [12, ∞). The second class is often broken down into one with an unspecified but limited upper boundary and one with an unlimited upper boundary.

United Nations (1997) makes no specific recommendations on the classification of durations of residence. However, it does suggest recommended tabulations, a recommendation to which many countries in practice adhere.

In these tables, the following classification is used, expressed in years: [0, 1), [1, 5), [5, 10), [10,  $\infty$ ) for internal migration; and [1, 5), [5, 10), [10,  $\infty$ ) for international migration. Thus, recent international migrants are ignored altogether in the suggested tables.

Clearly, in this respect, both United Nations (1998) and United Nations (1997), fall far short of our recommendation on residence durations, above, which specifies precision down to calendar month level.

The simple but essential fact of the matter is that countries which limit the publication of duration of residence data to such crude classifications will not be able to analyse actual migration processes which have taken place with any degree of precision. In section 3 of this paper, above, we describe in detail the reasons why this is so.

Obviously, there are reliability issues in the use of retrospective data such as those on durations of residence. For example, in cases where the most recent event of a move took place several years ago, a person may well have difficulties recalling the timing with the required precision. For a cohort as a whole, this will usually result in measurements which exhibit elements of shifting and/or heaping in the reporting of event timings.

The possible occurrence of such shifting and/or heaping, particularly for less recent moves, must not, however, be taken as grounds to group reported duration of residence data in broad duration classes, neither at the stage of data collection and coding nor at the stage of publication whether in digital or in paper form.

It should be left to the migration analyst to explore whether there are any such errors of reporting in the data set, and, if applicable, to judge from cohort to cohort which instruments are the most appropriate to assess their characteristics so as to be able to adjust for any such errors.

Further, it is important that durations of residence for all cohort members, whether they be recent migrants, non-recent migrants or non-migrants, be made available *in full*, and *for all cohorts*.

Making retrospective data on residence durations available only up to a maximum of, for example, five years prior to the enumeration is less informative than the full set of retrospective data.

Equally, there are no good reasons for not making available the data on youngest or on oldest cohorts.

In summary, the ability to measure and analyse the processes of international and of internal migration which have actually taken place with maximum scope, depth and flexibility is crucially dependent on the methodologically sound operationalization of concepts such as those of migrant and migratory move, of current and previous place of usual residence, and of duration of residence.

As we have seen from the discussion above, however, once concepts have been operationally defined and measured accordingly in the field, subsequently valuable detail may be lost and further constraints may be placed on the analyst in the processes of coding, data entry and tabulation. We shall briefly turn to these issues from a slightly more general perspective in our next and final section.

## 5 DATA PROCESSING

The taking of a population census is not a goal in itself. It is a relatively costly data collection exercise whose principal value lies in subsequent data use in research and analysis.

So, the census process does not end with the routine standard census publication programme. More appropriately, this is only the beginning of the actual use of the census. It is important that national statistical organizations are well prepared and well equipped to deal with the needs of researchers in terms of the ability to make available data sets on demand, data sets tailored to meet the advancing theoretical and methodological needs of research in disciplines including demography.

From this perspective we shall next discuss several important issues regarding census data processing. These are issues which are fundamental to the ability to study many facets of the population, including the international and internal migration processes which have actually taken place. As we have seen, population censuses are of particular, and in most countries unique, value for the study of international migration and of internal migration.

Even if every care has been taken to operationalize elementary concepts with due consideration of the requirements of the analyst of migration and if the subsequent fieldwork is of high quality, the scope and flexibility of the analyst may be restricted *ex post* in the data processing process. It is important therefore that this process too is designed with the requirements of subsequent data use in analytical research in mind.

The first issue here is that of *measurement coding*. The coding of the field observations may take place after taking the field measurements, the traditional approach, or as part of the measurement process itself. The latter will rapidly become more and more common as field measurement and data capture -- the conversion of the information obtained in the enumeration to a format that can be interpreted by a computer -- become integrated into a single step through the use of modern information technology.

Clearly, considering the above recommendations, it is elementary that coding schemes are designed from the outset such that *no information is lost in the coding process*, however the latter process may be operationalized technically.

For example, current and previous places of usual residence should be coded using the full geo-referenced data as outlined above. Similarly, duration of

residence data should be coded such that the full precision of the measurements is preserved. The use of any schemes to group measured data in classes in the coding stage prejudices the task of the migration analyst and must be rejected.

Thus, we have the following recommendation:

RECOMMENDATION 9 Coding schemes used for the digital entry of the field data should be designed such that the full precision of the measurements is preserved.

A second important issue is that of the methods of *storing, managing and retrieving the digital data*.

United Nations (1997) still describes data capture as a separate step in the census process after the actual enumeration. In §1.201 it is recommended that the edited field data be entered, one record per person, in a so-called master file (or micro-data file) for later tabulations.

Specifically United Nations (1997) states in this paragraph: "This master file ... can have a simple rectangular sequential format. There is usually no need for a database structure with index files."

Yet, the principles and recommendations continue in §1.209 to state that "In order to expand the life and usability of the data, and as a complement to the standard production of tables, national statistical offices are encouraged to store the census data in various computerized database forms so as to better satisfy the full range of needs of internal and external data users. ... Needs vary widely from user to user according to specific interests and circumstances. There is therefore no preferred approach to setting up a census or population database."

In §1.211 it is even recommended as one approach to consider the feasibility to develop the necessary database software in-house.

Also, the recommendations perceive a clear separation between the software environment for data capture and storage on the one hand, and for data production on the other, as is borne out by §1.205: "The use of software packages specifically designed to produce census tabulations is highly recommended."

Considering the overall approach to the use of information technology in the census process as recommended by United Nations (1997), we note several important drawbacks.

Clearly, considering the recommended "master file" and any database as separate entities introduces a substantial but unnecessary element of double work. It also brings with it an avoidable additional risk of information corruption and loss.

From an information technology point of view, furthermore, it is unclear why a census bureau would even wish to consider to maintain "various computerized database forms", and how this could contribute efficiently and cost-effectively to meeting the specific needs of different data users.

However, more importantly, the recommended approach to data management places a very heavy burden on the training and skills of census bureau staff, in particular when non-industry-standard software is used and when non-standard digital tabulations are required.

Further, even if a database is used, then the lack of clear guidance in §1.209ff seriously hinders the exchange of both expertise -- such as on information definition, management and retrieval -- and data within national statistical organizations over time and between national statistical organizations and researchers.

In fact, the analysis of migration places comparatively heavy demands on the retrieval of tailor-made information.

For example, even for the elementary measurement of migration, it is necessary for each migration defining area (MDA) of destination under study to produce digital *organized data sets* ("tabulations") of durations of residence by all relevant MDAs of origin, by age and by sex. In this process, MDAs defined to match the specific research needs and objectives will have to be assembled from the recorded detailed data on places of usual residence.

Any attempts at explanation of the observed migration phenomena require that the data sets are also specific by relevant explanatory covariates. These must be time-referenced, so that their relationship with the retrospective information on migratory moves can be set and analysed in the correct time perspective and order of precedence.

Further, a deeper insight into the dynamics of the international and internal migratory processes can be obtained if in addition similar data sets are produced for selected time points prior to the enumeration. This is valuable in particular for the study of frequent migrants and of short-term migrants.

However, more generally, such additional data sets allow the migration analyst to explore more deeply the information on the dynamics of the relationships between cohorts and migration behaviour over time and age which is contained in the retrospective data.

Such deeper studies require the reallocation of persons to the appropriate places of usual residence at those earlier time points using the migration life history data recorded in the enumeration.

None of the standard tables defined by United Nations (1997) as basic examples for a census publishing programme meet the minimum requirements for even the elementary measurement of migration as just described, let alone the associated explanatory analysis or the various forms of deeper analysis.

United Nations (1997) does indeed highlight the need, for example in §1.204, for national statistical organizations to be prepared to produce special census tabulations on demand. However, experience shows that in practice to date many countries find it difficult to accommodate such demands flexibly, timely and cost-effectively, if they can do so at all. Adherence to the United Nations (1997) recommendations for the design of the information technology component within the census process, recommendations which were already obsolete at the time of writing in the mid-1990s, clearly contributes to this.

Recalling that the analysis of migration is one of the key rationales for taking population censuses in the first place, it is quite fundamental that approaches to the use of modern information technology in the census process are brought up to date.

The information collected in population censuses is by no means unique in terms of the volume, complexity or transaction intensity to the extent that they would warrant specially-designed information input, storage, and search and retrieval tools. The obvious approach is the use of industry-standard relational database management systems (RDBMSs) and tools.

The cost of modern off-the-shelf RDBMS software is negligible when put in the context of the overall costs of a census enumeration, even in the economically least developed nations. Similarly, the ongoing dramatic drop in prices of servers, workstations, storage and network capacity has brought a state-of-the-art professional hardware infrastructure well within reach of even these countries. This low cost also makes the relatively high demand on storage space of RDBMSs a practical irrelevance. Further, when used to store, manage and retrieve census data, the transaction intensities are very low in relation to system design specifications; and given the handling capacity of modern RDBMSs, the underlying processing intensity of such systems is easily compensated for.

The use of such industry-standard systems and tools to manage the census data will directly and significantly contribute in four essential areas, namely

- (1) in relieving the overall burden on the census organization,
- (2) in increasing effectiveness and efficiency of the census organization,
- (3) in lowering the level of special skills and expertise required of census IT staff while at the same time placing a greater emphasis on routine industry-standard skills and expertise, and
- (4) in standardizing and streamlining the entire census process.

As a result, it will also facilitate the exchange and transfer of experience and expertise within the census bureau over time, and as well as between statistical organizations.

Further, RDBMSs allow an efficient and cost-effective integration of the data collection and the data capture processes, integration which, as mentioned, will become standard in census taking in the same way that it has been in survey research and in other areas for many years now.

RDBMSs are designed with a special emphasis on procedures to control and safeguard the consistency, integrity and security of the database, elements which are of considerable importance in the entire census process, from data capture and data management through to making data available for use in research.

They allow for the easy and consistent incorporation of data from successive population censuses, from other censuses such as on housing, and from incidental or periodic sample surveys, without disruption of any already existing data structures. The census data in the database may in particular also conveniently serve as sampling frames for such sample surveys.

But most importantly, the use of such industry-standard systems and tools to store, manage and retrieve the census data will enhance the flexible, timely and cost-effective accessibility and usability of the census information for research. It will facilitate comparative and time series analyses using successive censuses, as well as cross-national analyses.

As mentioned, population censuses are no goal in themselves. The ultimate purpose of census taking is *data use* in the quest to acquire timely and relevant knowledge on selected aspects of a nation, knowledge which can stand the test of scientific scrutiny. Managing the data using modern off-the-shelf RDBMSs is the most effective and efficient way in which truly lasting value can be obtained from the capital-intensive census data collection operation. For analysts of migration data it is the best approach available with a view to being able to access and exploit the full richness of the data collected.

Therefore we have the following recommendation:

RECOMMENDATION 10 Industry-standard relational database management systems (RDBMSs) and tools should be used to input, store and manage census data, and to retrieve organized data sets on demand and according to specification by the data user.

Given the deployment of industry-standard relational database management systems (RDBMSs) and tools, a *sound structural database design* that can stand the test of changing needs over time should be developed. The development of this design is a key and strategic area of decision making in the deployment process which requires expertise in the area of relational database systems.

Especially national statistical organizations which have no experience yet with such systems could benefit from assistance here with a view to maximizing the lasting benefit obtainable from their deployment, and with a view to avoiding common pitfalls.

The United Nations could usefully assist here by developing *elementary principles* and *operational guidelines* for best practice in relational database design for population censuses. This could in particular be useful in helping the database systems analysts who will implement the system to understand precisely what the characteristics of the census information are and what substantive and procedural demands will be made on this information.

Thus, we have the following recommendation:

RECOMMENDATION 11 The United Nations should develop elementary principles and operational guidelines for best practice in relational database design for population censuses.

Finally, as we have seen, even the full set of basic migration tables specified in United Nations (1997) is inadequate for the analysis of international or internal migration. In fact, the usefulness of a large quantity of detailed printed tables for general purpose use by analysts has long been a thing of the past.

While a small number of standard summary tables may remain interesting to the general public, the value of the pre-defined highly subclassified detailed tables suggested by United Nations (1997) for production as standard but without any reference to specific research needs and objectives is questionable.

The building of skills and experience in the production of digital *organized data sets* ("tabulations") on demand as specified by researchers and to be used for

further specialist analysis, on the other hand, is much more important. These are skills which are not difficult to acquire if industry-standard information systems query tools are used. However, they do require some hands-on training.

Here, too, the United Nations could usefully assist, namely by developing census database information retrieval *examples and training exercises*. These could in fact suitably replace the lay-out specifications in United Nations (1997) of the more detailed tables suggested for printing.

This could best be achieved by employing a given and well-designed (partial) synthetic or anonymized real population census database, embodying the above principles and guidelines for best design practice. An illustrative and educational set of clearly-explained search queries using an ANSI/ISO standard database query language such as SQL could then be supplied.

This could help census bureau staff in producing both the standard summary tables for general public information purposes, and any on-demand digital organized data sets subject to specification by the actual data user.

Clearly, it would be most useful if this material be made available *in digital format*, allowing users to practice their skills in standard and on-demand information search and retrieval hands-on.

Thus, we have the following recommendation:

<p><u>RECOMMENDATION 12</u> The United Nations should replace the specifications of the suggested more detailed printed census tabulations by operational advice and instruction on database queries for standard and on-demand information retrieval using a standard database query language such as SQL.</p>
---

With this recommendation, we bring our discussion of a number of key issues in the processing and the making available of population census data to a close.

In the concluding section we shall briefly review our principal findings and assess the scope for the implementation of our recommendations.

## CONCLUSION

For most countries, the regular population census is the only data collection instrument which will enable analysts to derive a comprehensive image of the processes of international and internal migration affecting the country. Only the relatively small number of countries which maintain a continuous population registration system in which all individual migratory moves are recorded, are in a more favourable position.

In practice most national statistical organizations will refer to publications of the United Nations as a principal source of guidance when organizing a population census. The current principles and recommendations for population and housing censuses (United Nations, 1997) are the basic source of reference. For topics on international migration, the current recommendations on statistics of international migration (United Nations, 1998) will be the principal reference. And for topics on internal migration, usually the basic reference will still be the manual on methods of measuring internal migration (United Nations, 1970).

However, both for the measurement of international migration and for the measurement of internal migration, these guidelines and recommendations are not optimal in a number of quite fundamental respects

Specifically, as detailed in this paper, adherence to these guidelines and recommendations in their current form will *not produce the most informative data on migration possible*. Adhering to these guidelines will also unnecessarily limit the scope and flexibility for the migration analyst to use census data on migration to their best advantage. As a consequence, the ability to investigate the international and internal migration processes which have actually taken place is needlessly impaired. Both research focusing on specific local needs and comparative research at the international level face serious and avoidable obstacles.

In twelve specific key recommendations, this paper sets out where changes in these United Nations guidelines and recommendations are required so as to allow the study of migration to the best effect.

When implemented, these recommendations will give a country the richest information on international and internal migration processes reasonably obtainable through a population census. In addition, implementation will provide the national census bureau with the capacity to facilitate that the maximum benefit

is derived from the information contained in the raw field data in both a timely and a cost-effective manner.

For ease of reference, these twelve key recommendations are reproduced in the "Summary of the Recommendations" at the end of this paper. However, this summary is presented without any explanation or motivation. In order to understand the rationale underlying each specific recommendation, it is important to refer to the appropriate sections of this paper.

The two fundamental recommendations, numbered 1 and 2, are that the *measurement of durations of residence* be given *core topic* status, both for international migration and for internal migration.

In addition, the recommendations on the *operational definitions* of the concepts of *migratory move* and of *migrant* (recommendation 4), of the concept of *place of usual residence* (recommendations 5 and 7), and of the concept of *duration of residence* (recommendation 8) are equally of major importance. If the current recommendations of the United Nations are followed here, then the value of the information on international and on internal migration collected through a population census will be much reduced.

Recommendation 6 deals specifically with the important issue of *high-mobility* sections of the population, including frequent movers and short-term migrants.

Recommendation 3 addresses the key question of the *balance* between the *number of variables* measured in population censuses on the one hand and the *quality of measurement* on the other.

Finally, in particular -- but not exclusively -- for the study of migration, it is essential that *no information* is *lost* in the data processing phase of the census process, and that the information technology used to input, to store and manage, and to retrieve census data is brought up to *modern information systems standards*. This is reflected in recommendations 9 to 12.

The next global round of population censuses will be around the year 2010. The United Nations and numerous other international and national bodies are currently actively involved in the process of reviewing and updating the recommendations for this new round of censuses.

From the discussion in this paper, it will be clear that countries which do not take advantage of the advances in the measurement and analysis of international and internal migration which are reflected in the key recommendations in this paper, will be at a disadvantage as far as the study of migration is concerned until the data from the 2020 census become available.

Generally, the implementation of these twelve recommendations by the United Nations and by national statistical organizations will be simple and substantively and procedurally beneficial, both in the short term and in the longer term.

Specifically, most of the recommendations directly addressing the measurement of international and of internal migration, namely 1, 2, 4, 5, and 6, should not pose any difficulties for national statistical organizations during the 2010 census. Equally, the recommendations on data processing (9, 10, 11, and 12) will be straightforward and cost-effective to implement. Recommendation 3 on the balance between quantity and quality will pay off as soon as it is implemented.

As outlined in the corresponding sections of this paper, recommendations 7 and 8, on the operational definitions of the concepts of place of usual residence and of duration of residence, might not be *fully* realizable yet in some poorer countries by 2010. The detailed measurement of place of previous usual residence as recommended will clearly be easiest to achieve in administratively well-organized countries. And the precise measurement of durations of residence requires elementary calendar awareness which might be less than perfect, particularly amongst the poorly educated.

All the same, a less than perfect implementation of these recommendations 7 and 8, using the approximating approaches described in this paper, is preferable to not implementing them at all. For, the closer a national statistical organization approaches the full implementation of these two recommendations as well, then the more benefit will be derived from the census enumeration for the measurement and analysis of international and internal migration.

## Appendix: A Brief Note on

### why a Question on *Duration of Residence* is preferable to a Question on the *Place of Residence at Some Fixed Date in the Past*

In this paper we developed a set of practical operational recommendations on the measurement of international and internal migration, specifically aimed at the 2010 worldwide programme of population and housing censuses.

This note serves to reinforce one of the fundamental notions underlying this paper, namely the superiority of a question on the duration of usual residence over a question on the place of usual residence at some fixed date in the past.

For the direct measurement of *internal migration* through population censuses and surveys, four questions have traditionally received most attention:

- 1 place of birth
- 2 place of usual residence at some fixed date in the past
- 3 duration of usual residence
- 4 place of previous usual residence

For the direct measurement of *international migration*, these same questions can be used. The usual approach is simply by substituting "country" for "place".

Additionally, more recently further questions have been suggested for the measurement of international migration, such as, for example, country of citizenship or nationality, ethnicity, and mother tongue of the respondent and/or of the respondent's parents. However, such questions are quite sensitive, and may well be politically and culturally charged and open to abuse.

The key question yielding the highest information value and allowing for the most powerful analysis in the study of both international and internal migration is the question on *duration of usual residence*. For geographic detail, it must be combined with a question on place or country of *previous usual residence*. For more insightful socio-economic analysis, it can be combined with a range of common other questions, including, of course, questions specifically designed to measure aspects of international migration.

The singular and unique merit of a question on duration of usual residence is underpinned by rigorous mathematical demographic arguments (Doeve, 1986, Xu-Doeve, 2005). However, it is also easy to see intuitively.

Recording the duration of residence is equivalent to recording the exact date of the transfer (move) of a person from one migration-defining area to another. Hence, the measurement of such durations *allows the complete reconstruction of migration processes as they have actually taken place in real time.*

In fact, when the data are adjusted for the effect of mortality through standard reverse survival, the information yielded by this question is in many ways similar to the information which one could obtain when a continuous civil registration system recording changes in usual residence would have been available.

Consequently and specifically, recording duration of residence in population censuses and surveys allows for the measurement of

- 1 *demographically consistent instantaneous migration rates,*
- 2 *along cohort lines,*
- 3 *as a function of continuous time and age.*
  
- 4 Further, *period rates* can be derived directly.
  
- 5 In addition, probabilities to make 0, 1, 2, ... moves within arbitrarily specified time intervals can easily be derived as well (the *multiple move* issue).
  
- 6 Also, using standard increment-decrement life table techniques, *transitions in discrete time* can be derived directly.

Consequently, *migrants* and *migrations (individual moves)*, *moves* in continuous time and *transitions* in discrete time, and *migrant stocks* and *flows* (expressed in absolute numbers and in intensities) can be distinguished and evaluated unambiguously and exactly.

However, this question on duration of residence offers one further and unique benefit:

- 7 *It is the only measurement instrument which allows for the adjustment of migration data for incompleteness of enumeration.*

Among mobile cohorts, such incompleteness can be extremely high: Xu-Doeve (2005), for example, found for the male cohort aged [20, 25) in Bangkok that

nearly three-quarters of all recent migrants had been missed in the Thai population census studied.

Such demographic information cannot be derived from mere questions on place or country of birth or on citizenship/nationality. However, equally, *such information cannot be produced when a question on the place of usual residence at some fixed date in the past is employed.*

The latter type of question has been strongly advocated within the past 20 years or so, in particular by Anglo-Saxon geographers working within a demographic accounting framework. Demographic accounting is a descriptive framework originally suggested as an alternative to the classical demographic paradigm (Rees and Wilson, 1977). An early detailed statement of how demographic accounting leads to promoting the use of a question on the place of usual residence at some fixed date in the past can be found in Rees (1984). This view has since received widespread following, particularly in developed countries but also in many developing nations (Bell, 2005).

However, the essential methodological issue is that, unlike a question on duration of usual residence, a question on the place of usual residence at some fixed date in the past *does not measure individual migratory moves.*

Consequently, some of the many drawbacks and limitations of a question on the place of usual residence at some fixed date in the past are the following:

- 1 *No instantaneous migration rates can be obtained.*
- 2 *The migration rates obtained are not consistent with the standard definition of occurrence/exposure rates.*
- 3 *The analysis is in discrete time, and so cohort migration rates as a function of continuous time and age are unavailable.*
- 4 *Multiple moves and migration trajectories, including staged migration, circular migration and return migration, are not or not properly dealt with. Consequently, also, migrants and migrations (individual moves), and migrant stocks and flows are not correctly identified.*

The question on the place of usual residence at some fixed date in the past only allows the measurement of *net migration transitions in discrete time and age* specific by place of usual residence at the fixed date in the past and by place of usual residence at the time of enumeration.

While information on migratory moves in continuous time and age allows the precise evaluation of such transitions over arbitrarily specified discrete intervals

of time and age, the reverse -- reconstructing moves from transition data -- is impossible.

Clearly, in the construction of, for example, multistate life tables for the analysis of migration behaviour, as well as for insightful and demographically consistent subnational population projections, such as urban-rural projections, detail as specified here is nevertheless required.

- 5 Consequently, in the study of migration, users of information derived from a question on the place of usual residence at some fixed date in the past necessarily have to *rely heavily on* -- methodologically often dubious -- *approximation and estimation*.

It is interesting to note in this context that, in defining a set of 15 core indicators of internal migration, Bell (2005) fails even to mention the most fundamental of demographic measures, the cohort *instantaneous migration rate* as a function of age and time. It is evidence of the extent to which a predominant focus on data on the place of usual residence at some fixed date in the past leads to indicators and instruments of unnecessarily limited analytical power and value.

- 6 Importantly, finally, information derived from a question on the place of usual residence at some fixed date in the past *does not allow for the adjustment of migration data for incompleteness of enumeration*.

Given the magnitude of such incompleteness, this will in fact usually *invalidate any migration statistics* derived on the basis of such data.

## REFERENCES

- Bell, M (2005) *Towards Rigorous Cross-National Comparison of Internal Migration: Who Collects What?* Paper presented at the XXV International Population Conference of the International Union for the Scientific Study of Population (IUSSP), 18-23 July 2005, Tours, France. (Copies are available from the author and from IUSSP)
- Doeve, W L J (1986) *How Do We Measure Migration? The Preferred Migration Questions for the Global 1990 Round of Population Censuses.* Paper presented at the International Conference on Urbanization and Urban Population Problems (ICUUPP), Tianjin, People's Republic of China, October 1987. (Copies are available in the conference proceedings)
- Rees, P (1984) *Does It Really Matter Which Migration Data You Use in a Population Model?* Working Paper 383, School of Geography, University of Leeds
- Rees, P and A G Wilson (1977) *Spatial Population Analysis.* Edward Arnold
- UNECE [United Nations Economic Commission for Europe] (2005) *Report of the March 2005 Joint UNECE/EUROSTAT Seminar on Migration Statistics.* United Nations
- United Nations (1970) *Manual VI. Methods of Measuring Internal Migration.* United Nations
- United Nations (1997) *Principles and Recommendations for Population and Housing Censuses, Revision 1.* United Nations
- United Nations (1998) *Recommendations on Statistics of International Migration, Revision 1.* United Nations
- Xu-Doeve, W L J (2005) *The Demographic Measurement of Migration and its Adjustment for Underenumeration.* Paper presented at the XXV International Population Conference of the International Union for the Scientific Study of Population (IUSSP), 18-23 July 2005, Tours, France. Revised Edition, 14 August 2005. (Copies are available from the author and from IUSSP)

## SUMMARY OF THE RECOMMENDATIONS

For the convenience of the reader, below we repeat the *recommendations* presented in this paper for the *measurement of international migration and of internal migration* in the *2010 global round of population censuses*.

This summary is presented without any explanation or motivation. In order to understand the rationale underlying each specific recommendation, it is important to refer to the appropriate sections of this paper. The relevant section is indicated in square brackets after each recommendation.

- 1 International migration is most suitably measured by a question on the unbroken duration as a usual resident of the country, supplemented by a question on the previous country of usual residence (if the unbroken duration is less than the current exact age). These questions should receive *core topic* status. [3]
- 2 Internal migration is most suitably measured by a question on the unbroken duration of residence in the current place of usual residence, supplemented by a question on the previous place of usual residence (if the unbroken duration is less than the current exact age). These questions should receive *core topic* status. [3]
- 3 It should be assessed which census core topics and which census non-core topics can be adequately or even more suitably measured through periodic sample surveys. These topics should be removed from the list of topics to be considered for inclusion in population censuses. [4]
- 4 The operational definitions of the concepts of migratory move and of migrant in terms of a required minimum duration of residence or usual residence at the destination should be rejected. Instead, the operational definitions of these concepts should be based exclusively on a change of usual residence. [4.1]
- 5 An exhaustive international standard list should be drawn up of purposes of stay which disqualify a place of residence as the place of usual residence. The purpose of stay should be ascertained and the list should be applied

exclusively in respect of those persons whose uninterrupted duration of residence is less than one month. [4.1]

- 6 To capture the essence of high-mobility sections of the population, the life history measurement of recent arrivals, defined as all those who have arrived in the current place of usual residence within the past 5 years, should be extended to include not just the last move but also the last-but-one move. This should be done using similar measurement instruments, that is, duration of residence in combination with associated previous usual residence. [4.1]
- 7 For every enumerated person both the place of current usual residence and any places of previous usual residence measured, if the latter are within the country, should be properly geo-referenced in the census database using GPS (Global Positioning System) co-ordinate data. [4.2]
- 8 It is imperative that true durations of residence be measured with as much precision as possible during the actual enumeration process through the use of appropriate measurement instruments and techniques. A measurement scale which is precise down to year and month of arrival is satisfactory. [4.3]
- 9 Coding schemes used for the digital entry of the field data should be designed such that the full precision of the measurements is preserved. [5]
- 10 Industry-standard relational database management systems (RDBMSs) and tools should be used to input, store and manage census data, and to retrieve organized data sets on demand and according to specification by the data user. [5]
- 11 The United Nations should develop elementary principles and operational guidelines for best practice in relational database design for population censuses. [5]
- 12 The United Nations should replace the specifications of the suggested more detailed printed census tabulations by operational advice and instruction on database queries for standard and on-demand information retrieval using a standard database query language such as SQL. [5]