

Unpublished paper ^{delivered} presented at the Demsa Conference on
"Demographic Data, Definition and Dissemination",
Univ. of the Western Cape, Bellville 26-27 Sept 2003.

Migration data from South African censuses: evaluation and suggestions

by

Pieter Kok (HSRC)

Census '96 generated a very useful data source for the study of migration patterns in South Africa. Despite a few flaws, the data provided a first-ever opportunity to undertake fairly detailed analyses of internal migration in respect of the entire country. The migration data generated by Census 2001 will be somewhat more limited in scope but still very useful for the analysis of internal migration. At the same time the place-of-birth data from Census 2001 will be more useful than those generated by Census '96. This paper also provides a few suggestions regarding migration questions in future censuses.

Introduction

Migration data generated by censuses will probably always be problematic. The first problem relates to the need in a census questionnaire to restrict the number of questions to an absolute minimum. This clearly limits the scope of the migration information that can be generated. The second problem relates to the coding of the origin data for the millions of individuals enumerated in a census (the correct names, localities and types of the places where household members have lived before). The third problem relates to memory lapse, notably with regard to the dates at which moves occurred (especially in the more distant past).

How then does one deal with these problems while at the same time generating useful census data for appropriate migration research? The answer is rather complex, but practical solutions are available. The last two censuses dealt with these issues rather effectively. Census '96 contained a rich battery of migration questions (see Box 1), which is a solid foundation for the design of future census questionnaires. Census 2001 built upon this solid foundation to potentially provide very useful migration data.

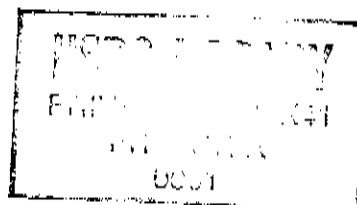
Since the number of migration questions included in census questionnaires must be restricted to an absolute minimum to save costs and elicit a high and accurate response rate, some important choices need to be made. These choices will therefore always represent trade-offs between what is absolutely essential and what is useful for studies of spatial mobility, based on the costs associated with every question.

A *de jure* count of the population was obtained by means of the questions on place of usual residence (see items 3.1 and 3.2 in Box 1). The rationale for the *de jure* count is that people should be counted where they normally live, not where they happened to be on census night. These *de jure* localities are the places where people spend most of their time, and where and for whom planning should be done. Although the notion of 'usual place of residence' is potentially flawed (see Kok 1999), it helps to place populations in a proper spatial perspective.

The data obtained from Question 3.2 in Box 1 have (so far) been made widely available only to the level of magisterial district (rather than the specific place) of origin. This was probably the correct decision. In the census of 1980 there was a question on place of residence five years before. These responses were then coded at the suburb level, which theoretically made it possible to undertake studies of (local) residential mobility. The problem was, however, that this was so costly an exercise that only 5 per cent of the questionnaires could be fully coded down to this level of spatial detail.

Evaluating South Africa's census-based migration data

In the past, South Africa lacked suitable data on internal migration. This means that historical trends could not be analysed to the extent required in a country that underwent such notable political, social and economic changes during the last decade of the 20th century as South Africa. The only useful available historical data on internal



BOX 1

MIGRATION QUESTIONS ASKED DURING CENSUS '96

SECTION A: IN RESPECT OF EACH HOUSEHOLD MEMBER:

- 1.1 "Was (the person) born in South Africa? (Include the former Transkei, Bophuthatswana, Venda, Ciskei—TBVC states)
1 = Yes
2 = No"
- 1.2 "(If 'No') In what country was the person born?
Write in the name of the country"
- 2 "Is (the person) a migrant worker? (Someone who is absent from home FOR MORE THAN A MONTH each year to work or to seek work.)
1 = Yes
2 = No"
- 3.1 "Is this DWELLING (e.g. house, room, shack, flat) the place where (the person) usually lives, i.e. where (the person) spends at least four nights per week?
1 = Yes
2 = No"
- 3.2 "(If 'No') Where does (this person) usually live?
Name of suburb/village/settlement:
Name of city/town/farm/tribal authority:
Name of magisterial district:
If not South Africa, please state name of country:"
- If no usual address, circle "3" 3"
- 4.1 "In which year did (the person) move to the DWELLING (e.g. house, room, shack, flat) where he/she usually lives?
Write in the year that he/she moved 19.....
OR
1 = The person has never moved. 1
(Lived in the dwelling since birth)"
- 4.2 "(For the person who has moved)
From where did (the person) move? (Before moving into the dwelling where he/she usually lives)
Name of suburb/village/settlement:
Name of city/town/farm/tribal authority:
Name of magisterial district:
If not South Africa, please state name of country:"

SECTION B: IN RESPECT OF THE ENTIRE HOUSEHOLD:

5. "Are there any persons who are usually members of this household, but who are away for a month or more because they are migrant workers? (A migrant worker is someone who is absent from home for more than a month each year to work or to seek work).
1 = Yes
2 = No
(If "Yes") Indicate the person's particulars:
Age in years
Gender
Relationship to head of household
Where is (the person) living:
Name of suburb/village/settlement:
Name of city/town/farm/tribal authority:
Name of magisterial district:
If not South Africa, state name of country:"

migration were those for the period 1975-1980, but these were flawed by the exclusion of data in respect of the former homelands of Transkei, Bophuthatswana and Venda. Census '96 provided a very welcome change by making available, for the first time in South Africa's history, data on internal migration for the entire country and population. This new source of data has invited the undertaking of appropriate analyses of the data. However, the prior absence of suitable data has obstructed the development of experience in the analysis of migration data over time. Subsequent censuses needed to provide these comparative data, and fortunately Census 2001 has done exactly that.

In Kok et al. (forthcoming) a comprehensive evaluation of Census '96 data is provided. It was found that the 1996 census provided, for the first time in South Africa's history, a very valuable source of data in respect of *internal* migration for the entire country and its population. The validity of the data from Census '96 on *international* migration was not comprehensively tested, but an initial evaluation of the latter data showed up some serious inconsistencies, not least of which were the country codes used.

Suggested scope of census-based migration data

Information on residential mobility can be obtained cost-effectively by means of localised sample surveys, if and when required. *A country's central statistics office (Statistics South Africa in our case) has the responsibility to concentrate its resources on the more general migratory moves in the country as a whole, rather than localised moves.* The (full) census should therefore provide data that will allow comprehensive analyses of migration patterns (across political and administrative boundaries), because these data cannot be obtained in any meaningful way from sample surveys.

Censuses cannot be expected to provide data for the analysis of the causes and consequences of migration. These are dealt with more efficiently in other types of surveys. Censuses can only be expected to provide data for the analysis of the (a) volume, (b) direction, (c) pattern, and (d) time of migration.

(a) Volume. To determine the volume of migration, the researcher must at least be able to distinguish non-migrants from (former) migrants. In its simplest form, this requires a question such as 'Have you ever lived elsewhere?' The issue at stake is not so straightforward, however. As showed by Kok (1999), it may be of crucial importance to restrict one's definition of migration to only those moves that took place from one migration-defining area to another. This requires a predetermination of what constitutes a migration-defining area. If, for example, a decision is made to use (magisterial) districts as migration-defining areas, the question should then have been phrased as: 'Have you ever lived in another district?'

Despite some major demarcation flaws in South Africa's magisterial districts during the apartheid era (see Geyer et al. 2000), magisterial districts are still potentially the most useful migration-defining spatial units for this purpose. This is because they do not vary in geographical size as much as the other possible spatial units (such as the areas of jurisdiction of local governments), and are sufficiently large to allow the identification of spatial patterns and trends. One major problem, though, is that only a minute proportion of the population can be expected to know the boundaries of magisterial districts, especially for places where people lived a long time ago. They may also be unaware of recent district boundary/name changes. A similar argument may apply in respect of local government boundaries. The best solution would be to use the migration-defining area 'restriction' only in the question relating to the *current* (usual) place of residence, and to record in as much detail as possible the location of the *previous* place of residence – see (b) and (c) below.

(b) Direction: The direction of migration can be determined either by a single question such as 'Where did you live before moving here?' or by a battery of questions dealing with multiple earlier moves. With a battery of questions the data can be restricted to the last (e.g. two or three) moves or even to all the moves during a person's adult life (for example), or be expanded to cover all moves since birth (and including place of birth) for comprehensive life-course migration analyses. In a census such a battery of questions is not viable, which means that a single question (dealing with the last move or the place of residence on a specific date prior to the census)

is usually opted for. As will be seen later, it may be advisable to ask two questions – one question covering place of residence on a predetermined date prior to the census (e.g. at the time of the previous census), and another covering the origin of the last migratory move.

A crucial need in migration research is for information generated by a question on previous place of residence, such as 'Where did you live before moving here?' Another, preferably complementary, approach could be to ask a question such as 'Where did you live five years ago?' Despite the problems associated with fixed migration intervals, it is sometimes necessary for comparative and analytical purposes to restrict one's analyses to a specific time period. A five-year period is probably the most appropriate because it is long enough to allow for a sufficient number of moves to have taken place, and not too long for people to have forgotten the details of where they lived. Moreover, significant boundary and name changes are unlikely to have taken place. Of course the period does not have to be five years or a multiple of five years. Any interval can be used as long as it is carefully chosen to maximise accuracy and utility.

By *combining* the origin of the last move with the place of residence five years ago, the analyst can arrive at unambiguous measures of the level of migration and the frequency of moves. *It is suggested, therefore, that future censuses should record both the origin of the last move and the place of residence five years ago.* However, if a choice between these two questions **must** be made, the *origin (and date) of the last move* is preferred to a question on place of residence at a fixed prior date (for the reasons mentioned by Kok 1999).

(c) *Pattern:* Migration patterns can be described in various ways. The most common of these are related to *type of locality* before and after the move (e.g. rural and urban, urban and urban, rural and rural, or urban and rural) and *stepwise* moves (e.g. hierarchical moves such as rural-town, town-city, city-metro).ⁱ Other, less familiar, movement patterns include those related to 'primary metropolisation' (rural-metro) or 'secondary metropolisation' (e.g. town-metro or city-metro). To obtain that kind of information it is necessary to determine the urban (town/city) or rural (farm/ village) status of the places of origin and destination.

Once again, the issue is much more complex than may be expected at first sight. There have been lively debates concerning the classification of places as 'urban', 'rural', 'semi-urban' or 'peri-urban' in the South African urbanisation literature (see, for example, Graaff 1986 and Gelderblom & Kok 1994). Furthermore, the status of places changes over time. Leaving such a complex classification entirely to the respondents would not only be unfair to them but would also probably yield results that cannot be compared.

It is suggested, therefore, that not only the information on the rural or urban status of previous and current (usual) places of residence should be asked, but the corresponding farm/village/town names should also be obtained. This should not be done for detailed coding purposes (which is probably too costly), but purely with a view to classifying the places concerned into one of at least two categories, namely 'rural' and 'urban'. Coding staff should perhaps further divide the broad category 'rural' into (i) 'rural', (ii) 'semi-urban' and (iii) 'peri-urban'.ⁱⁱ

The suggested classification, which should preferably (although it is not ideal) be based on the *current* situation (i.e. at the time of the census; **not** the situation at the time of a person's birth or previous move), be coded as follows:

- (1) 'rural', i.e. areas where the majority of people are dependent on income derived from agricultural pursuits, and not forming part of the area of jurisdiction of a metropolitan local government;
- (2) 'semi-urban', i.e. those (predominantly informal), non-urban settlements/villages that are altogether more than, say, 50 km from the nearest formal town and do not qualify as 'urban' (i.e. as a town or city);
- (3) 'peri-urban', i.e. non-urban smallholdings or settlements/villages within, say, 50 km of a formal town;
- (4) 'town/city', i.e. an area that is (unambiguously) classified as 'urban';

Another important matter that should be raised here is the need to obtain as much information as possible on place of birth. In Census '96 the question on place of birth was restricted to the country level, which made

BOX 2

SUGGESTED MIGRATION QUESTIONS FOR FUTURE CENSUSES

SECTION A: IN RESPECT OF EACH HOUSEHOLD MEMBER:

1. Where did the mother of (the person) live when (the person) was born?

If abroad: Name of country:

If in South Africa:

If in urban area: Name of suburb/town:

If in rural area: Name of farm/village/area:

For all areas:

Name of tribal/local authority:

Name of nearest (large) town:

Magisterial district (if known):

FOR OFFICE USE ONLY: 1 = Rural, 2 = Semi-urban, 3 = Peri-urban, 4 = Town/City

District/Country code:

2. Is (the person) a migrant worker? (Someone who is absent from home FOR MORE THAN A MONTH each year to work or to seek work.)

1 = Yes

2 = No

3.1 Where did (the person) live on (the day of the previous census)?

At this address:

1 = Yes

2 = No

3.2 (If 'No') Where did (the person) usually live on that day?

IF "NO":

If abroad: Name of country:

If in South Africa:

If in urban area: Name of suburb/town:

If in rural area: Name of farm/village/area:

For all areas:

Name of tribal/local authority:

Name of nearest (large) town:

Magisterial district (if known):

FOR OFFICE USE ONLY: 1 = Rural, 2 = Semi-urban, 3 = Peri-urban, 4 = Town/City

District/Country code:

4.1 Is this DWELLING (e.g. house, room, shack, flat) the place where (the person) usually lives, i.e. where (the person) spends at least four nights per week?

1 = Yes

2 = No

If no usual address, circle '3'

3

4.2 (If 'No') Where does (the person) usually live?

If abroad: Name of country:

If in South Africa:

If in urban area: Name of suburb/town:

If in rural area: Name of farm/village/area:

For all areas:

Name of tribal/local authority:

Name of nearest (large) town:

Magisterial district (if known):

FOR OFFICE USE ONLY: 1 = Rural, 2 = Semi-urban, 3 = Peri-urban, 4 = Town/City

District/Country code:

5.1 In which year did (the person) move to this DISTRICT (or the district where he/she usually lives)?

Write in the year that he/she moved

OR

1 = The person has never moved. (Has lived in this dwelling since birth.)

1

BOX 2 (continued)

5.2 (For the person who has moved)
From where did (the person) move? (Before moving to this DISTRICT
(or the district where he/she usually lives.)

If abroad: Name of country:

If in South Africa:

If in urban area: Name of suburb/town:

If in rural area: Name of farm/village/area:

For all areas:

Name of tribal/local authority:

Name of nearest (large) town:

Magisterial district (if known):

FOR OFFICE USE ONLY: 1 = Rural, 2 = Semi-urban, 3 = Peri-urban, 4 = Town/City

District/Country code:

SECTION B: IN RESPECT OF THE ENTIRE HOUSEHOLD:

6. Are there any persons who are members of this household, but who are now away for a month or more because they are migrant workers? (A migrant worker is someone who is absent from home for more than a month each year to work or to seek work.)

1 = Yes

2 = No

6.1 (If 'Yes') Indicate the person's particulars:

Age in years

Gender

Relationship to head of household

6.2 (If 'Yes') Where is (the person) currently living?

If abroad: Name of country:

If in South Africa:

If in urban area: Name of suburb/town:

If in rural area: Name of farm/village/area:

For all areas:

Name of tribal/local authority:

Name of nearest (large) town:

Magisterial district (if known):

FOR OFFICE USE ONLY: 1 = Rural, 2 = Semi-urban, 3 = Peri-urban, 4 = Town/City

District/Country code:

it impossible to determine lifetime (let alone life-course) patterns of migration for the vast majority of the population. It is suggested that the place of birth should be treated in the same way as the other 'previous places of residence' in terms of locality status and detailed place names. The 2001 census improved significantly on this question by not only asking the country of birth but (for the South African born) also the province of birth. Unfortunately, this is still not quite enough for the purposes of migration research. It is essential to know the specific area/place (e.g. magisterial district).

(d) *Time of the move:* It should be remembered that people's responses to questions about past dates are significantly affected by memory lapse. It is therefore important not to rely too much on data that refer to events that happened a long time ago. Although migration intervals have many potential flaws they are still very useful in terms of census requirements.

Five years is probably a good compromise for the purpose of obtaining information for fixed migration intervals. Therefore, as has been the case in 1980 and 2001, if the next South African census is to be held in the year 2006 (for example), it may be a good idea to use a five-year interval (2001–2006) for the question on place of residence on a fixed prior date. There is no compelling reason why census day in 2001 (10 October 2001) would be less appropriate than any other date. It would also add much value to the analyses if data on inter-census migration were directly available.

In view of all these considerations we recommend that a battery of migration questions similar to the one suggested in Box 2 be used during future population censuses. The suggested battery of questions reflects an 'ideal', yet realistic, approach to the type of migration data that a census can be expected to generate. Not only is it sufficiently comprehensive, but also likely to be quite viable. The only inherently complicated elements are the different settlement types, but it should be worth the additional effort to avoid having to revert back to the traditional (and unsatisfactory, oversimplified) rural/urban dichotomy.

Conclusion

Migration is a complex component of the empirical study of population dynamics. Not only is it difficult to record and classify previous places of residence, but it is also a major headache to define 'migration' and distinguish it from other forms of spatial mobility. No wonder, therefore, that migration research has been lagging behind fertility and mortality research. In South Africa, unlike most countries, the absence of appropriate data contributed to the relative underdevelopment of migration research.

Census '96 opened up many opportunities for meaningful national analyses of internal migration patterns in South Africa, and comparative census data have since been generated by Census 2001. The responsibility rests with South African migration researchers to utilise the opportunity that this would create, not only to analyse the census data but also to take these analyses into appropriate models, especially with a view to informing planning and policy making.

Notes

ⁱ See Conway (1980) for a description of the different varieties of stepwise moves.

ⁱⁱ To qualify as either 'semi-urban' or 'peri-urban', informal, non-urban settlements/villages should contain a large number of people (e.g. more than, say, 5 000), the majority of whom depend on incomes obtained from non-agricultural pursuits or urban-based sources. (See Graaff 1986 for a more detailed description of these two settlement types.)

References

- Conway, D. 1980. Step-wise migration: toward a clarification of the mechanism. *International Migration Review*, 14(1):3-14.
- Gelderblom, D. & Kok, P. 1994. *Urbanisation: South Africa's challenge (Volume 1: Dynamics)*. Pretoria: Human Sciences Research Council.
- Geyer, M., Orkin, M., Lehohla, P. & Kahimbaara, J. 2000. Revisiting the South African magisterial districts of 1999. *Development Southern Africa*, 17(2):263-276.
- Graaff, J.F. de V. 1986. *The present state of urbanisation in the South African homelands and some future scenarios*. Occasional Paper No. 11. Stellenbosch: Research Unit for Sociology of Development.
- Kok, P. 1999. The definition of migration and its application: Making sense of recent South African census and survey data. *Southern African Journal of Demography*, 7(1):19-30.
- Kok, P., O'Donovan, M., Bouare, O. & Van Zyl, J. (forthcoming). *Internal migration in South Africa: An analysis of census and other secondary data*. Pretoria: Human Sciences Research Council.