

Using the voters' roll to build a clearer picture of migration patterns

Johan Maritz, Pieter Kok and Gerbrand Mans
25 August 2020

Using the voters' roll to build a clearer picture of migration patterns

Johan Maritz, Pieter Kok and Gerbrand Mans

The census is the only accurate source of national migration data for South Africa. However, the census is conducted once in 10 years and municipalities, confronted with rapid change in the size and composition of settlements, require reasonably recent migration analysis for the planning and development of services. In order to fill this need, it is useful to identify migration trends from alternative databases that are sufficiently large and geographically extensive.

The end of apartheid meant that people could move about the country and change their place of residence much more easily than previously. Since 1994 there have been significant changes in the national settlement space due to people's migration. Being able to track migration patterns would enable authorities to respond appropriately: to manage urbanisation, undertake economic development and job creation, and improve living conditions.

As part of the Spatial and Temporal Evidence for Planning in South Africa (StepSA) initiative – which was originated by the Department of Science and Technology – the Council for Scientific and Industrial Research (CSIR) and the Human Sciences Research Council (HSRC) explored ways of meeting this information need expressed by several municipalities.

The importance of understanding migration is self-evident in a world where globalisation, economic integration and enhanced communication have made people aware of opportunities to live and work in new environments. The strong trend to urbanisation through migration is also likely to continue for the foreseeable future. Over time migration changes not only the size of the population, but also the demographic composition of cities and towns, creating a demand for different types of services.

The national census, undertaken by Statistics South Africa (StatsSA), has been the only source of national data on migration. Other surveys have been done but they have either been limited in terms of geographic scope or focused on qualitative information.

Even census data has limited value as it is performed once every 10 years, while municipalities have shorter planning cycles. In addition, the census only records the last migration in the previous decade and fails to capture serial movements, where these have occurred. Lastly, the census provides a coarse representation of migration patterns, while municipalities require more detailed, spatially specific information.

It was in this context that the CSIR and HSRC began to explore the possibility of extracting migration information from voter registration data in the possession of the Independent Electoral Commission (IEC).

What defines migration?

Clearly migration involves the movement of people – but how far do they have to move, for how long and for what purpose, to qualify as migrants? This study favoured the definition by Kok and Collinson (2006) who proposed that migration is a “crossing of the boundary of predefined spatial units by persons involved in a change of residence”. The researchers felt it was important to adopt a definition sufficiently broad to accommodate the common South African practice of “circular migration” which involves the movement of mainly rural people to cities for the duration of their working lives and their return to their rural roots when working life ends.

The question of boundaries

Often local municipal boundaries are used to measure and reflect migration but they are not always the most suitable units. Given their size, they are coarse units that do not reveal migration occurring on a finer scale. Also, municipal areas vary considerably in size and this can lead to distortions in geo-spatial indicators and comparisons among units. A final consideration in respect of boundaries is that, where possible, they should remain unchanged to facilitate the comparison of populations over time. Where this is not the case, in order to compare data from different periods, a process of alignment of boundaries is required to create comparable areas of analysis.

Features of migration analysis

Fertility, mortality and migration are the elements that determine the growth or decline of a population. However, while fertility and mortality yield once-off *in-situ* data, migration involves movements that can occur multiple times and have both a starting point and a destination. The widespread pattern of circular migration has already been mentioned and is of critical concern to authorities in their planning of services.

Capturing data on migration can be problematic, especially where migrants are foreign and want to remain “under the radar” for various reasons, including their legal status and xenophobia. Foreign migrants are usually under-represented in censuses and surveys.

Voter registration data: a source to track migration?

Voter registration data collected by the IEC was identified as a potentially valuable source of migration information for South Africa owing to:

- The frequency with which data is updated.
- Its accuracy in terms of identifying individuals and their voting wards.
- Its coverage of the entire nation, both in terms of area and representing a high proportion of the adult population
- The very high spatial resolution, which enables detailed tracking of movements.

The major shortcoming of the data is that it only includes people aged 18 years and older and that registering to vote is voluntary.

To be eligible to vote in South Africa, an individual must register his/her details in the voters’ roll for the district where s/he resides. When a voter moves to a new district, s/he must re-register in that district. This gives a geographically accurate description of voter movement because:

- An urban voting district comprises some 3 000 voters and covers an area within 7.5km of relevant voting stations.
- A rural voting district accommodates some 1 200 voters and covers an area within 10km of the voting stations.

Voting districts must align with the latest municipal boundaries but can be modified in response to settlement and population changes.

Voters’ rolls are well maintained and accurate because they are required to support some form of election every two to four years and they are scrutinised by political parties.

The above factors suggested that, despite some limitations, voter registration data could yield vital information about migration patterns.

Accessing and processing voter data

The voters' roll contains personal information on all registered voters and the IEC – in compliance with the Protection of Personal Information Act – provided StepSA with data in an anonymised format. Among other things, this meant removing names and identity numbers and replacing them with age and gender particulars of each individual. Researchers received data on the following variables:

- An anonymised person identifier.
- Gender and age.
- A four-digit birth year as supplied by the Department of Home Affairs.
- The voting district where each person was registered in 1999, 2000, 2004, 2006, 2009 and 2011.
- Details of voting districts for the above years.

The voters' roll can be considered a good dataset, as indicated by the fact that it contained 23 655 046 voters for nationwide municipal elections of 2011 and the census of 2011 recorded 33 702 589 individuals aged 18 years and older.

Alignment of voting districts

Since some voting districts had changed over the years, a process was needed to create comparable spatial units across the voters' rolls used for seven consecutive elections.

The voters' roll for 2011 was selected as the reference roll and six others were compared to it and, where necessary were adjusted to match it in terms of the spatial areas of voting districts. The next step was to allocate the pool of voters in the original voting districts to the new voting areas created to match the 2011 voters' roll. This allocation was done by randomising the voters eligible for allocation and then redistributing them among the relevant "new" districts.

Analysis of migration among voters: 1999 - 2009

The data provided by the IEC gave researchers a unique opportunity to study the migration levels of South African voters at a spatially detailed level. The analysis of the data was undertaken for two consecutive five-year periods (1999 - 2004 and 2005 - 2009) and a 10-year period (1999 - 2009). The table below reflects the proportion of the total number of registered voters who migrated between different municipalities.

Percentage of total registered voters migrating between municipalities*

Province	1999 - 2004	2005 - 2009	1999 - 2009
Free State	7.5%	7.3%	14.9%
Eastern Cape	10.5%	9.5%	20.1%
Gauteng	10.7%	11.2%	21.4%
KwaZulu-Natal	9.9%	11.1%	21.1%
Limpopo	9.5%	7.3%	17.2%
Mpumalanga	9.7%	9.7%	19.6%
Northern Cape	11.1%	8.8%	19.8%
North West	12.0%	10.5%	22.5%
Western Cape	8.7%	9.6%	18.6%
South Africa	10.0%	9.9%	20.0%

*2009 municipal boundaries were the basis of spatial analysis

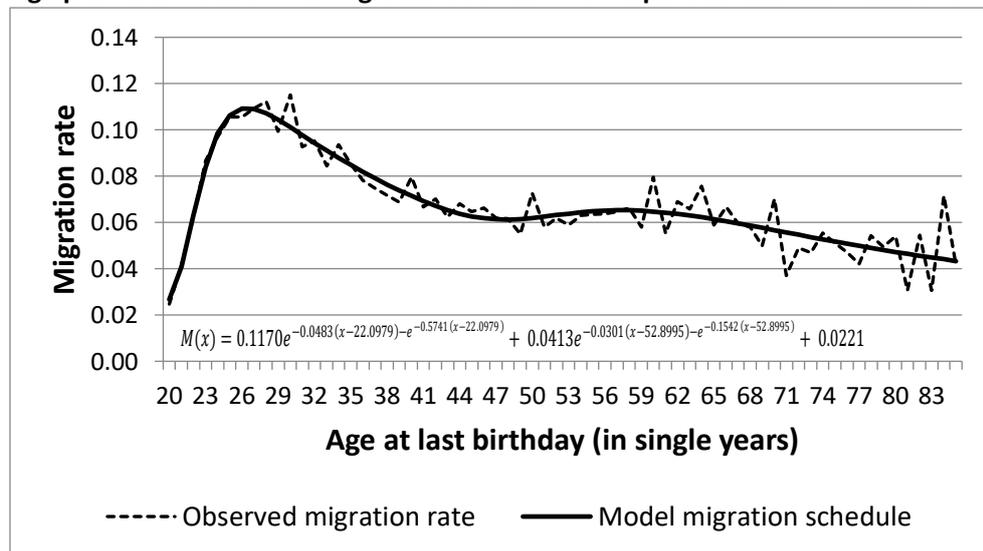
The figures in the table confirm a general trend established as far back as the 1970s: migration levels in South Africa have remained remarkably constant over time. Kok, O'Donovan, Bouare and Van Zyl (2003) observed that “despite dramatic political, social and economic changes in the country (including the abolition of apartheid’s migration-related measures such as influx control and group-areas demarcations), the overall level of migration between the late 1970s and early 1990s did not change significantly”. A further study spanning 1996 - 2001 and the voters’ roll analysis for periods from 1999 to 2009 show continuity in levels of migration.

Migration and age

The proportion of migrants present in various age groups is known as the migration rate for age. Age-specific figures show that migration is much more prevalent at certain ages. Migration studies in various countries have revealed a typical pattern, with peaks in migration occurring in infancy, among young adults and at retirement.

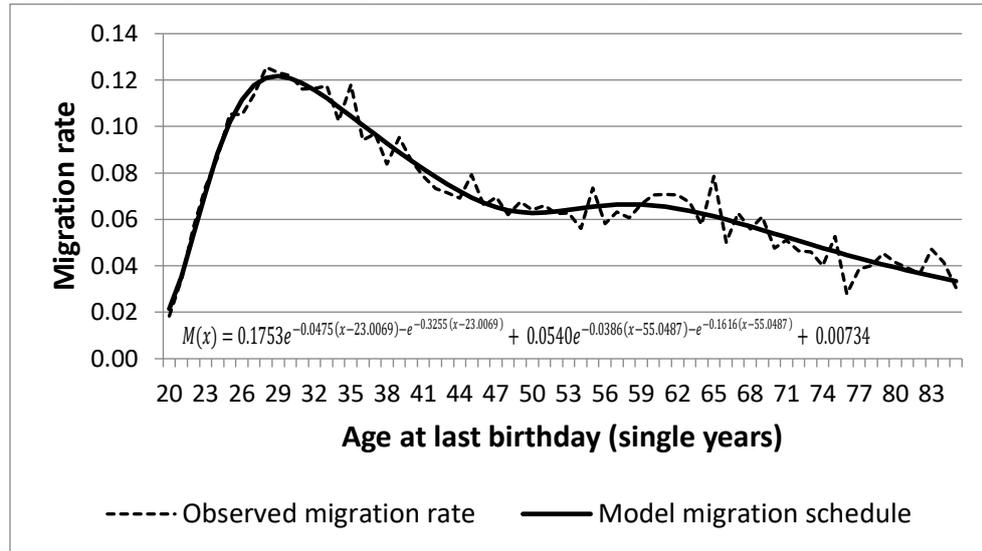
It is possible to create statistical models of the migration rate for age. Modelled estimates, based on the 1996 census data, have been calculated for the period 1992 - 1996. Graphs representing this modelling and the observed (actual) data derived from the IEC data are presented below. The graphs are unusual in that they only start at the age of 20 years. This is because the IEC data only covered people aged 18 and older and the first complete age cohort of census data within this range is for 20 - 24 years.

Age profile of voters who migrated between municipalities between 1999 and 2004



Source: Derived from IEC voters’ roll of 2009.

Age profile of voters who migrated between municipalities between 2004 and 2009



Source: Derived from IEC voters' roll of 2009.

It is clear that the age-related migration profiles for the periods 1999 - 2004 and 2005 - 2009 are quite similar, indicating that age selectivity remains consistent over time. Migration is most common in the early working years (24 – 35 years) and there is a slight peak at retirement, around 60 years.

Visualising migration flows

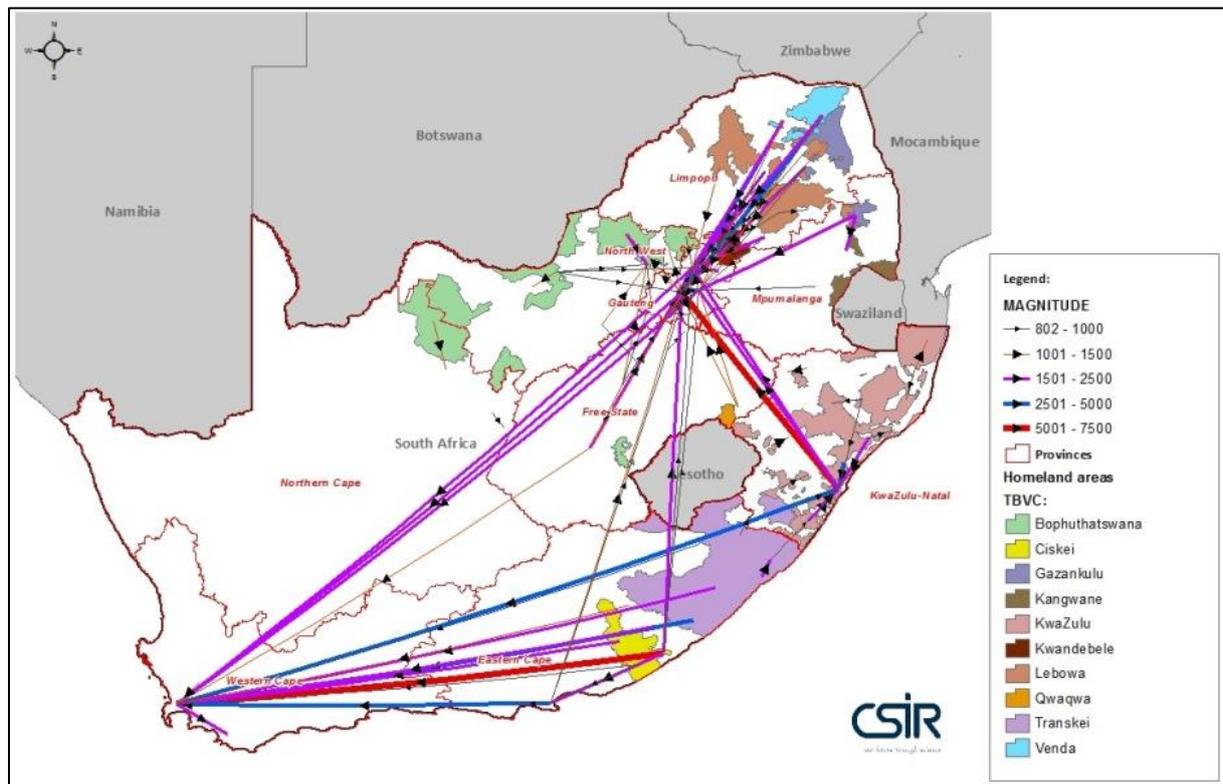
Once the voting district data was adjusted and randomised, data tables for each election period were prepared and summarised, enabling researchers to sum up the flow of voters from an origin zone to a destination zone, using ArcMap GIS. This made it clear which were the *from* voting districts and which the *to* voting districts. The process was repeated to specify flows to and from local municipal areas. At this point it was possible to link data to the 2009 local municipality file.

Users find it extremely useful to see a map of spatial flows which facilitates the immediate grasp of patterns of origins and destinations. This is particularly true for decision-makers who are often overloaded with documents and reports. The research team therefore decided to illustrate major migration movements by inserting flow lines on a map, creating a graphic much like a spider diagram.

Mapped results and initial applications

The map below is one of a set of maps on the theme of migration which was generated from voter registration data and is available on the StepSA portal (<http://stepsa.org>). It depicts migration flows between local municipalities with specific consideration of the location of former homeland territories. Recording all flows would result in a cluttered and unreadable map, so only flows exceeding 800 people have been indicated on the map, capturing the more substantial migration patterns.

Main migration flows of South Africa voters in relation to former homeland areas



Source: Maritz, 2014.

By including the locations of the former homelands on the map, it is possible to relate major migration flows to these areas. Given the lack of development and high unemployment in these areas, outmigration to the larger centres was expected – and it is clearly shown. There are definite net flows from areas in the Eastern Cape to Cape Town and Gauteng, especially southern Johannesburg. Flows from Vhembe District in Limpopo to Ekurhuleni and Johannesburg in Gauteng are also shown. At the same time, there are substantial flows between metropolitan areas, especially the Gauteng metros and Cape Town.

Explanations for the migration patterns revealed by this data spatialisation exercise will be sought through follow-up studies.

Value and limitations of IEC data

The limitations of the IEC data relate mainly to the fact that it only includes individuals aged 18 years and older who opt to register. Large sections of the population are excluded. Therefore, analysis of voters' rolls does not replace migration information captured in the 10-yearly census.

A lesser concern is the fact that some people who move from a voting district do not re-register at their new place of residence and simply travel to vote where they previously lived. However, the numbers are probably quite small and do not impact on the reliability of data when working at an aggregate level. This could become an issue when working at a fine, localised level.

The benefits of using IEC data to track migration have been identified by Maritz and Kok (2014) as follows:

- Unlike census data, IEC data enables researchers to identify successive migratory moves by individuals and gain a longitudinal perspective on migration.
- Migration surveys depend on the memory of respondents who may inaccurately recall dates and places, as well as their ages. IEC data is free of recall errors.
- It is also free of coding errors that may occur in survey data.

Maritz and Kok (2014) compared interprovincial migrant flows derived from census data for the period 1001 – 1011 with those derived from IEC data. A good correlation was found and this provides a measure of confidence in the IEC-derived data – at least at a high level.

Concluding comments

Although South Africa has better data than many developing countries it still requires more frequent, finer-scale information, especially for local planning and policy development. District and local municipalities have identified migration as a key issue in planning.

Given the low frequency of national data collection on migration, the IEC district-based voter data has considerable potential as a proxy dataset on migration. A key advantage is that the data relates to relatively small spatial units. In addition, comparisons of IEC data and other migration information show that the former is a viable source of complementary data.

The limitations of IEC data relate largely to human behaviour and the effects of this can mostly be overcome by the aggregation of data or using additional socio-demographic information.

There are empirical regularities that characterise modelled migration schedules and these may be considered as important as the well-established regularities observed in fertility or mortality schedules. As Rogers and Castro (1981) have highlighted, schedules may be used to graduate or smooth out irregularities in observed data; to generate data for each year of age; and assist in assessing the reliability of empirical data or resolving problems caused by missing data.

When applying the IEC-derived data to the developmental issue of the former homeland territories, the spatialised results indicated major migration streams to the larger metropolitan centres in the Western Cape and Gauteng.

The IEC migration data still has much more to reveal. At the time of writing only a preliminary age analysis had been conducted at a national level. Further age analyses as well as analyses pertaining to gender and other population characteristics had yet to be undertaken. The data will be updated, where possible, as the voters' roll is amended for future elections and will be used more intensively to understand migration behaviour and yield value for national, regional and local planning and policy making.

References:

Kok P and Collinson M (2006). Migration and urbanisation in South Africa. Pretoria: Statistics South Africa

Kok P, O'Donovan M, Bouare O and Van Zyl J (2003). Post-apartheid patterns in internal migration in South Africa. Cape Town: Human Sciences Research Council

Maritz J and Kok P (2014). Using election registration data to measure migration trends in South Africa. *Town and Regional Planning*, 31 - 42

Rogers A and Castro L (1981). Model schedules in multistate demographic analysis: the case of migration. WP-81-22. International Institute for Applied Systems Analysis: Laxenburg, Austria