# Making use of the consistency of patterns to estimate age-specific rates of inter-provincial migration in South Africa

Rob Dorrington and Tom Moultrie
*Centre for Actuarial Research, University of Cape Town*

*This paper uses a multiplicative component approach and the related log-linear model to present and compare interprovincial migration in South Africa by province of origin, destination, age, sex and race for three periods (1991-96, 1996-2001 and 2001-07). This, inter alia, shows a consistency of pattern of migration by age, sex, and race with the exception of children under age 5 (which are prone to undercounting and data scanning errors) which may extend back some 30 years. This consistency permits the estimation of the number of migrants born in the five years prior to the census from the numbers at older ages. These numbers of migrants are then converted to annual rates of migration and Rogers-Castro multi-exponential curves fitted to these patterns. The paper ends with a discussion of the implications of these results for projecting future interprovincial migration rates and the generalisability of the findings to other environments.*

## 1. Introduction

In order to project a population one needs, inter alia, estimates of current and expected future migration (either numbers or rates). Techniques used to analyse and model migration in the developed world have rarely been applied to migration data from Africa. In addition, and in particular, South Africa has been through a socio-political transition that could mean that past patterns of migration are of limited use in projecting future patterns of migration.

This paper investigates the usefulness of applying modern models for analysing and describing migration using census data in the context of an African country which might be expected to have experienced significant changes in migration.

## 2. Methods and datasets

### a. Census and survey data

Data used in this analysis are the number of individuals by province of origin, destination, age, sex and population group (race) for the most recent move in the past five years as indicated by the 1996 and 2001 census data, and since the 2001 census for the data from the 2007 Community Survey (which involved around 2.5% of households).

### b. Analysis

First the data were disaggregated into the four components (overall, origin, destination and origin-destination interaction) using the multiplicative component approach ($n_{ijx} = (T)(O_i)(D_j)(A_x)(OD_{ij})(OA_{ix})(DA_{jx})(ODA_{ijx})$) suggested by Rogers, Willekens, Little, and Raymer (2002). These components were then analysed for similarities and differences both visually and

by making use of the associated multiplicative log-linear framework used by a number of authors (Raymer, Bonaguidi and Valentini 2006; Rogers, Willekens, Little *et al.* 2002; Rogers, Willekens and Raymer 2001, 2002, 2003).

The similarity in the age distribution of the proportion of the overall migration for ages over 4 suggests that, following an idea proposed by Raymer and Rogers[1] (2007), one can make use of this regularity to estimate the number of migrants of those born in the past five years or so, where these data are deficient due either to undercounting in the census or scanning errors during data capture, from the numbers over that age. Such problems are typical of many, if not most, censuses conducted in developing countries.

### c. Annual rates of migration by individual ages

The results produced from the application of the method described above are not ideally useful for projections: they are neither the number of migrants nor the number of migrations over a fixed period. Thus we propose a method[2] for converting those results into annual migration rates and consequently multi-exponential curves (by province of origin, province of destination, sex and population group) as proposed by Rogers and Castro (1981).

## 3. Results

While the total migration differed for each of the periods the consistency of the patterns of migration can be seen from Figure 1. This consistency was replicated in the data by sex and, to a large extent, population group.

[1] Raymer and Rogers suggested using this regularity to allow one to estimate the complete age schedule of migrants from data on only the number of migrants born between the two censuses. Because of the data deficiencies surrounding child data (undercounting, scanning, age misstatement, etc) experienced in many developing countries, particularly those in Africa, this idea is not likely to work. However, if one has data for the older ages one can reverse the idea and use those data to provide better estimates of migration of young children.

[2] Essentially, by assuming a maximum of one migration per year one can use the census data and the migrations in the year immediately preceding the census to estimate rates of migration for that year. These rates can then, in turn be used to gross up the recorded numbers of migrants in the preceding year for onward movement and the grossed up number then used, together with the backward projected census numbers to estimate the rates of migration in the second year before the census, etc.

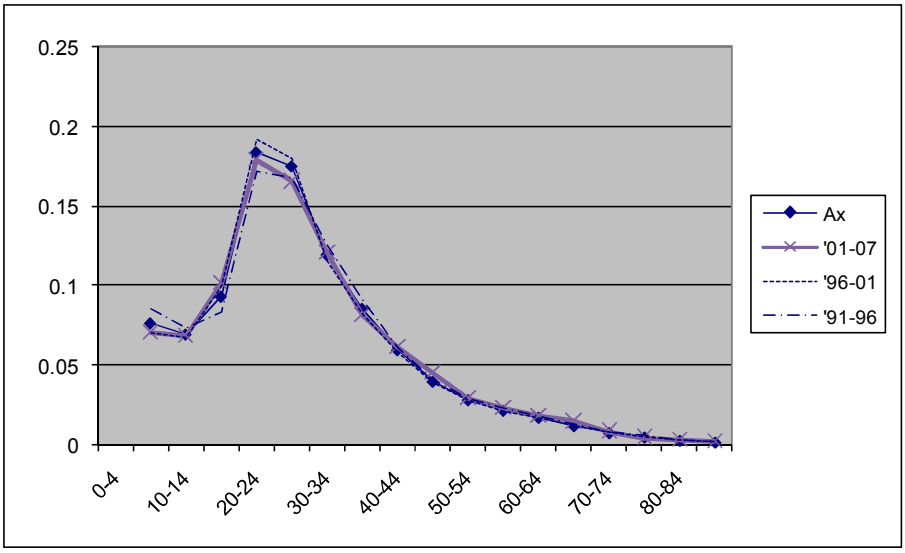


Figure 1: Proportion of the total migration of those aged 5 and older by age for the periods 1991-96, 1996-2001 and 2001-07, together with the distribution of the weighted average of all periods (Ax)

While there is a broad consistency in the patterns by province of origin and province of destination for most provinces, (as shown in the example in Figure 2) there are also some differences that require further analysis (as shown in the example in Figure 3).

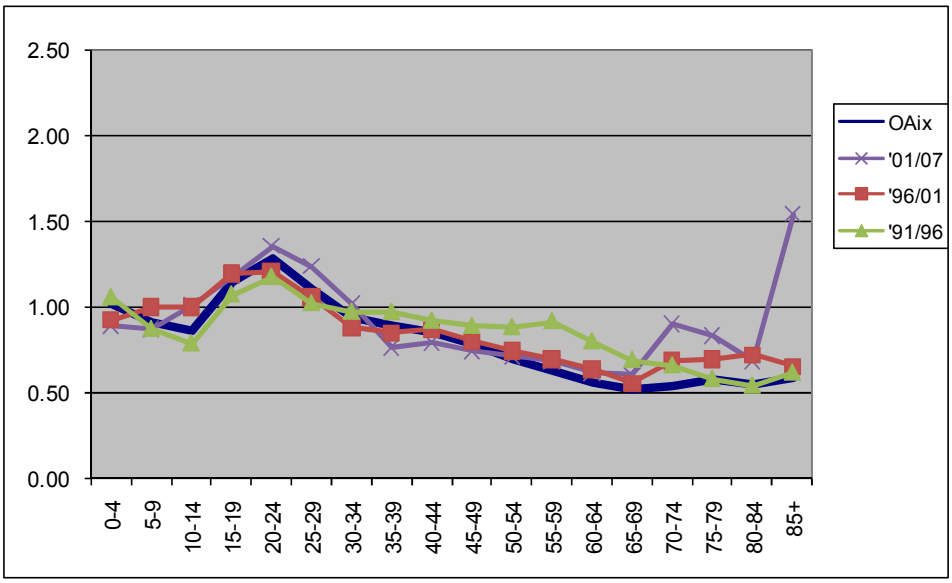


Figure 2: Deviations of the age profile of those with Eastern Cape as origin from that of all migrants

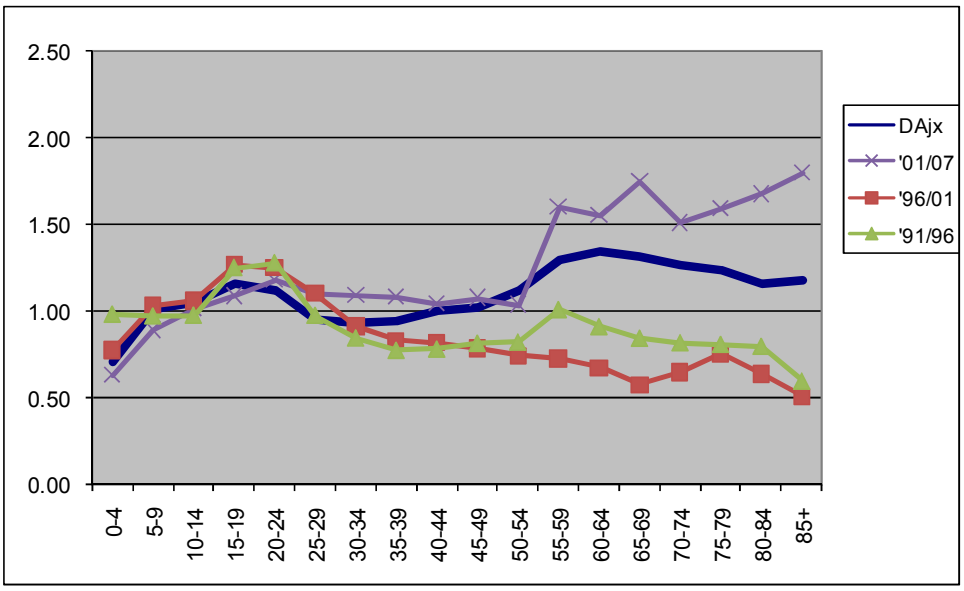


Figure 3: Deviations of the age profile of those with Western Cape as destination from that of all migrants

## 4. Discussion

Somewhat surprisingly the proportion of total internal migration by age has changed very little over the past 15 years, with very little difference between the males and females, not only in the pattern but also in the overall magnitude. This holds for each of the four population groups separately, with the major difference between the years simply being the total number of migrants.

This pattern by age seems to be fairly consistent with that produced from the migration over the period 1975-1980 (Kok and Collinson 2006; Kok, O'Donovan, Bouare *et al.* 2003) of those recorded in the 1980 census[3]. The major difference is the much smaller proportion of migrants that are female that is suggested by these estimates. Significantly, the patterns of provincial deviation of the age distribution of the proportion of migrants from that of all migrants are consistent with well established patterns for the major 'sending' provinces (Eastern Cape and Limpopo) and the major 'receiving' provinces (Gauteng and Western Cape) but differ over time for some of the other provinces.

The results presented here offer some potential for estimating patterns and levels of migration in developing countries from limited and deficient data. An important consequence of the possible generalisability of the approach is that it does not merely regard migration as the residual item in a cohort-component projection, but makes use of data collected on migration

[3] The 1980 census excluded people living in three areas designated by the Apartheid government as being 'independent', namely, Transkei, Bophutastwana and Venda.

and observed regularities in the age distribution of migration. In this regard, then, the paper is a contribution to our ability to estimate migration from limited and defective data.

## References


Kok, P. and Collinson, M. 2006. *Migration and urbanization in South Africa*. Report 03-04-02. Pretoria, South Africa: Statistics South Africa. Available:

Kok, P., O'Donovan, M., Bouare, O. and Van Zyl, J. 2003. *Post apartheid patterns of internal migration in South Africa.* Pretoria, South Africa: HSRC Press.

Raymer, J., Bonaguidi, A. and Valentini, J. A. 2006. "Describing and Projecting the Age and Spatial Structures of Interregional Migration in Italy", *Population, Space and Place* **12**:371-388.

Raymer, J. and Rogers, A. 2007. "Using age and spatial flow structures in the indirect estimation of migration streams", *Demography* **44**(2):199–223.

Rogers, A. and Castro, L. J. 1981. *Model Migration Schedules*. RR-81-030. Laxenbeurg, Austria: International Institute for Applied Systems Analysis. Available:

Rogers, A., Willekens, F. J., Little, J. S. and Raymer, J. 2002. "Describing migration spatial stucture", *Papers in Regional Science* **81**:29-48.

Rogers, A., Willekens, F. J. and Raymer, J. 2001. "Modeling interregional migration flows: Continuity and change", *Mathematical Population Studies* **9**:231-263.

Rogers, A., Willekens, F. J. and Raymer, J. 2002. "Capturing the age and spatial structures of migration", *Environment and Planning A* **34**:341-359.

Rogers, A., Willekens, F. J. and Raymer, J. 2003. "Imposing age and spatial structures on inadequate migration-flow datasets", *The Professional Geographer* **55**(1):56-69.