

MARTIN E.
PALAMULENI

Age Reporting in North West Province, South Africa, 1996–2011

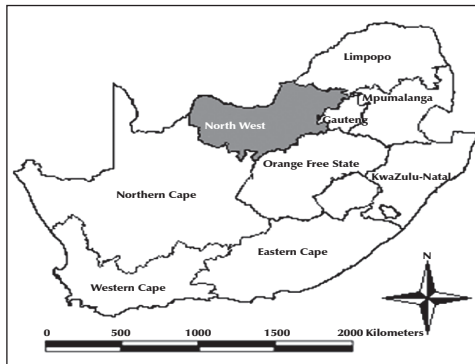


FIG. 1. MAP OF SOUTH AFRICA
SHOWING NORTH WEST PROVINCE

Martin E. Palamuleni

Population Training and Research Unit,
North-West University, Republic of South
Africa.

Introduction

AGE IS one of the most important variables in demographic analysis. It is both affected by and a determinant of fertility, mortality, migration and socio-economic variables. On the one hand, age determines entry into marriage, voting, the labor force and education. In the South African context only those aged 7 years are supposed to enroll in grade one, no children below the age of 18 is supposed to work. On the other hand, high fertility or low infant mortality implies more people below the age of 15 whereas for various reasons high levels of literacy lead to good knowledge and reporting of age statistics.

As a result of this two-way relationship, age not only forms the basis of classification for most demographic variables but also its familiarity is essential for successful social and economic planning. To underline the importance of age, the United Nations (1980) strongly recommends that LDC should include a question on age in

their censuses and demographic surveys; and Shryock and Siegel (1976) argue that no census is worth the name if it excludes a question on age.

Unfortunately, for various reasons, studies on age statistics from the statistically underdeveloped countries have revealed enormous distortions (Bailey and Makannah 1996; Caldwell 1966; Blacker 1969; Van de Walle 1966; Caldwell and Igun 1971; Nagi et al. 1973; Ewbank 1981; Byerlee and Terera 1981). Such studies on the nature and patterns of reported age statistics have attracted the attention of social commentators (Berkowitz 2012) and researchers from various countries around the globe (Ntozi 1978; Mukherjee and Mukhopadhyay 1988; Udjo 1996); Palamuleni 1996; Simelane 2002; Unisa et al. 2009). It has been found out that people tend to round off to the nearest age and some ages like those ending in 0 and 5 are preferred whereas others like 1 and 9 are avoided. The causes of these distortions are many and include people's ignorance of their true age, the instructions given to enumerators, the method used to collect age statistics and various other social, cultural and even political reasons.

There is also a tendency to under-state or over-state one's age in order to suit certain social and biological expectations. In this respect young children below the age of five, found playing with their peers at the time of the enumeration, may be reported as belonging to age group 5–9, and females in age group 10–14 who have passed puberty (menarche) may be recorded in age group 15–19 especially if it is further observed that they are married and are mothers. Likewise women above the age of 40 who are still rearing (nursing) their own children may be assigned a younger age group. Then, for prestige purposes, perhaps arising from the desire to be granted a senior citizenship status which in some cases exempts them from paying tax, there is a tendency especially among the males to exaggerate their age.

Awareness of these distortions and inaccuracies in reported age statistics have preoccupied demographers so much that the evaluation and adjustment of age statistics has become an integral part of demographic analysis. The main purpose of this study is to determine the nature and patterns of age misreporting in North West Province of South Africa.

There are three reasons for undertaking this study. First, the desire to isolate genuine distortions caused by famine and natural disasters from age misstatements. Second, since it is hoped that the characteristics of the reported age distributions will at a later stage be used to estimate levels, trends and differentials in fertility, mortality and migration, the examination of age data was felt to be of paramount importance as a means of providing an in-built mechanism of assessing the plausibility of the derived demographic estimates. In particular, to determine the probable effect of the age-sex distortions on the reported levels of

fertility, mortality and mortality. Third, it is further anticipated that age errors may provide clues to other weaknesses in the data.

Background Information

THE REPUBLIC of South Africa is divided into nine provinces, namely, Eastern Cape, Free State, Gauteng, KwaZulu-Natal, Limpopo, Mpumalanga, Northern Cape, North West and Western Cape. North West Province of South Africa is bordered by the provinces of Gauteng, Limpopo (formerly Northern) Province, the Northern Cape and the Free State, and the Republic of Botswana (Fig. 1).

North West is the sixth largest province covering a total area of 116,320 km² (approximately 9.5% of South Africa). The total population in the province increased from 3.3 million in 1996 to 3.6 million in 2001 and it is currently estimated at 3.7 million. The provincial population represents 8% of the national total and about 65% of the population in the province lives in rural areas. The province is divided into four district municipalities as follows: Bophirima, Bojanala, Southern and Central; five cross-border districts and 21 local municipalities. The available census data indicate that the largest percentage of the population in the province live in Bojanala (36%) followed by Central (23%), Southern (18%) and Bophirima (13%). The more populated industrial centers include Rustenburg, Brits and Ga-Rankuwa in the eastern region of the province. Mafikeng is the provincial capital and was the administrative center of the Bophuthatswana homeland (from 1978 to 1994). It was also the governing center of the British Bechuanaland Protectorate prior to 1960. Other major towns in the province include Potchefstroom, Klerksdorp, Lichtenburg, Ventersdorp and Vryburg.

The provincial gross geographic product (GGP) is R 3 964 per person against the national average of R 6 498 and mining forms the backbone of the provincial economy, contributing 42% to the GGP and 39% to the employment. The mining sector is dominated by large platinum mines and smelters in the Rustenburg area, as well as the gold mines of the Orkney and Klerksdorp areas. Agriculture is the second-most important sector, with 13% of the GGP and 18% of employment. Maize and sunflowers are the most important crops grown, while cattle and game farming are also well-established. Tourism is widely considered to have a major growth potential as the province is located adjacent to areas of Gauteng and Botswana.

Materials and Methods

Data Sources

THE STUDY will make use of the 1996, 2001 and 2011 South African Population Censuses and the 2007 Community Survey (Statistics South Africa 1998, 2003, 2012). The 1996 Census was the first census to be conducted in democratic South Africa (Statistics South Africa, 1998). This was followed by another census in 2001. However logistical challenges led to the postponement of the third census to 2011.

Data Analysis

IN TERMS of methods of analysis, the study will employ the most widely used measures of quality of reported statistics, namely Whipple's and Myers Indices and United Nations (UN) Joint Age-Sex Score (also known as UN Accuracy Index). These have been widely used in the analysis of the reported age statistics in developing countries (Palamuleni 1996; Poston et al. 2000; Mba 2003, 2004; Bekele 2006; Dahiru and Dikko 2013; Yazdanparast et al. 2012).

Whipple's Index is a summary measure that gives the extent of age heaping as a result of preference for ages with terminal digits 0 and 5. The index is obtained by calculating the percentage of the total reported on ages ending with 0 and 5 in the 23 to 62 age range divided by one fifth of the total population in the same age range. The index assumes any value between 0 and 500, with a value of 100 indicating no preference and a value of 500 signifying that ages of all people were reported only on ages with terminal digits 0 and 5.

Other researchers have developed "Whipple-type" indices that measure the degree of heaping on other terminal digits. For instance Poston and his students introduced a "Whipple-type" index to measure the degree of heaping on age 3 for the ages between ages 23 and 53 (Poston et al. 2000) whereas Talib et al. (2007) developed indices for each terminal digit and extended the age range from 20 to 79 years. Spoorenberg (2007, 2009) has shown that the original Whipple's index is a fair and reliable measure of the quality of age reporting only when attractions on 0 and 5 digits are significant. Unfortunately, to date, the modified Whipple indices have not gained much publicity and are not widely used. As such the analysis presented in this paper make use of the original Whipple's formulae.

Another index for appraising the quality of single year age data was developed by Myers. Unlike Whipple's Index which looks at preference for ages with end digits 0 and 5, the Myers Index examines the preference (or avoidance) of report-

ing ages ending with each of the ten digits 0, 1, 2, 3, 4, 5, 6, 7, 8 and 9. The index is secured by first calculating the weighted population reported on ages ending with each of the ten digits and then expressing the blended population on each digit as a percentage of the total population, in the absence of any irregularities in the reporting of ages, the sum of the blended population on each digit is expected to be equal to 10 percent of the total blended population. A percentage in excess of 10 percent indicates preference of ages ending with such digit and vice versa.

The quality of reported age-sex distribution in five-year age groups is evaluated using the United Nations Age-Sex Accuracy Index (also known as United Nations Age-Sex Score). This index is calculated as three times the Average Sex Ratio Score plus the Age Ratio Score for males plus the Age Ratio Score for females. Age ratio for any age group is the population in the age group divided by the average of the populations in the preceding and succeeding age groups multiplied by 100 (United Nations 1952; Arriaga et al. 1994). Shryock and Siegel (1976) note “barring extreme fluctuations in past births, deaths or migration, the three age groups should form a nearly linear series. Age ratios should . . . approximate 100.0 even though actual historical variations in these factors would produce deviations from 100.0 in the age’s ratio for most ages. For each sex, an Age Ratio Score is calculated by summing the differences of each age ratio from 100, without taking into account the sign and the mean of these differences is what is known as Average Age Ratio Score.”

The UN has provided two standards for explaining these scores. The United Nations (1952) suggested that if the joint score is less than 20 then the distribution is accurate, if the score lies between 20 and 40 then the distribution is inaccurate, otherwise for scores above 40 the age distribution is highly inaccurate. The United Nations (1955) further recommended that the age ratio of 2.6 for males and 2.4 for females and the sex ratio score of 1.5 implying a joint score of 9.5 should be accepted as a minimum standard.

Lastly the reported age-sex distributions were compared with model stable populations that were found to be consistent with other demographic parameters obtained from the population censuses. The stable populations are assumed to be error free and any deviations are attributed to errors in the reported statistics.

In the absence of age misreporting or age heaping, in the absence of international emigration and immigration, in the absence of war, major epidemics or famine induced mortality, the age distribution of a country’s population is a non increasing function of that age. This is true irrespective of the country’s level of development. However the decrease in the age distribution among developing countries is higher than in the developed countries.

Although indices such as Whipple’s and Myers’ indices or United Nations joint scores are commonly used in assessing the accuracy of reported age-sex

statistics distributions these measures are just summary indices and do not reveal very much. As such researchers have often argued that detailed information regarding the distortions in the age-sex distributions may be obtained by plotting the age-sex distributions (Cleveland n.d.; Moultrie et al. 2013; Fajardo-González et al. 2014). In line with this observation, graphs have extensively been used to illustrate the extent of age errors in the study population.

Results

FIGURE 2 depicts the age-sex distributions of the population of North West Province in South Africa. The reported age-sex statistics in the province, like elsewhere in the country, are fraught with certain anomalies that are not easily explained and have been a source of contention amongst researchers working with South African data (Berkowitz 2012). Some of these anomalies are also shown in figure 2. For instance Figure 1 indicates that the population in age group 0–4 in both 1996 and 2011 censuses is more or less the same. Moreover the population in age group 0–4 is bigger than the population in age group 5–9 in 2011. This may suggest that fertility in South Africa may have increased in the recent past, a suggestion that is not only improbable but also is not supported by the available data. In the subsequent paragraphs some of the commonly used indices that measures the extent of age misreporting are used to study the quality of age-sex statistics in North West Province of South Africa.

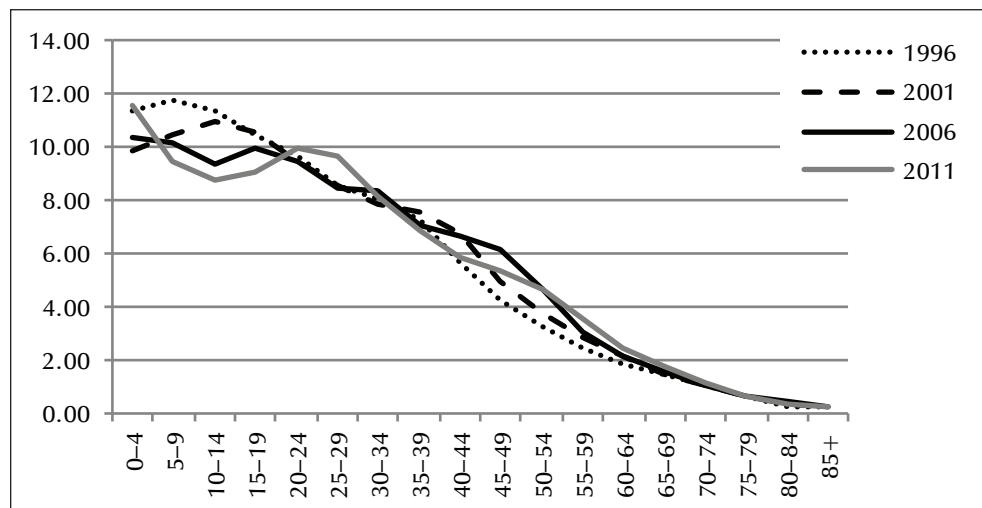


FIG. 2. POPULATION DISTRIBUTION BY AGE GROUP FOR NORTH WEST 1996, 2001, 2006 AND 2011

Whipple's Index

TABLE 1 presents Whipple's indices for males, females and both sexes for North West Province for the 1996, 2001 and 2011 Population Censuses and 2007 Community Survey. In general, the results indicate that there was less pronounced age heaping of the population on ages ending with digits 0 and 5. Whipple's index was 97 for both sexes, 97.4 for males and 96.6 for females indicating that preference of ages ending in digits 0 and 5 was slightly more common among male respondents than female respondents. Similar values for 2001 census are 95.2, 95. and 95.3, respectively. Comparisons of Whipple's indices for 1996 and 2001 censuses indicate improvement of the reported age statistics. In terms of racial groups, Whipple's Index was highest among the Asians, closely followed by Whites, then Colored and lowest amongst Africans.

The United Nations (1990) notes that if the values of Whipple's index are less than 105 then the age distribution data are deemed to be highly accurate; if the values are between 105 and 109.9, they are "fairly accurate"; if between 110 and 124.9, "approximate"; if between 125 and 174.9, "rough" and if 175 or more, "very rough" (United Nations 1990, 18–19). This means that the reported age distributions in North West Province as measured by Whipple's index can be regarded as "highly accurate."

The United Nations further states that "although Whipple's Index measures only the effects of preferences for ages ending in 0 and 5, it can be assumed that such digit preference is usually connected with other sources of inaccuracy in age statements and the indexes can be accepted as a fair measure of the general reliability of the age distributions" (United Nations 1990, 20).

TABLE 1. WHIPPLE TYPE INDICES FOR NORTH WEST PROVINCE, SOUTH AFRICA, 1996–2011

		Original			Modified		
		Male	Female	Both	Male	Female	Both
TOTAL	1996	97.4	96.6	97.0			
	2001	95.1	95.3	95.2	87.6	73.8	80.7
	2007	96.2	92.7	94.5	89.3	86.1	87.7
	2011	96.0	95.5	95.7	80.4	98.4	76.3
African	1996	96.9	96.0	96.4			
	2001	94.5	94.6	94.5	90.6	76.3	83.4
	2007	96.2	92.9	94.6	90.8	93.3	92.0
	2011	95.4	95.0	95.2	79.8	67.9	73.9

Continued on next page

Table—Continued

		Original			Modified		
		Male	Female	Both	Male	Female	Both
Colored	1996	99.8	100.1	100.0			
	2001	100.2	101.9	101.1	103.4	93.1	98.2
	2007	78.7	80.7	79.7	138.1	107.5	82.9
	2011	99.0	97.4	98.2	80.2	76.0	76.4
Asian	1996	103.2	103.1	103.2			
	2001	109.3	99.6	105.0	78.6	69.1	68.0
	2007	75.2	105.2	87.0	260.6	210.4	143.0
	2011	105.0	102.0	104.0	64.7	59.6	53.5
White	1996	103.3	103.0	103.1			
	2001	101.6	101.8	101.7	51.0	45.9	46.7
	2007	101.5	93.0	97.2	60.0	59.1	50.0
	2011	100.8	99.9	100.4	53.9	45.2	49.5

Myers Index

TABLE 2 presents a summary of the Myers indices for males, females and both sexes and by population group for North West Province for the 1996 and 2001 Population Censuses. The preferred and avoided terminal/end digits are also displayed in figures 3 to 6. In terms of the digit preference of each of the terminal digits Table 2 show that there was over-enumeration of ages ending with digits 6, 4, 8 and 0 while ages with terminal digits, 1, 2, 3, 5, 7 and 9 were avoided. This pattern was noticed in both male and female populations in North West Province. In 2001, terminal digits 1, 9 and 8 were over-enumerated whereas the other remaining terminal digits were under-enumerated. The female population also shows over-enumeration at terminal digit 6. The 2007 community survey indicate that the digits 9, 5, 8 and 6 were preferred among the male population whereas the female population saw a preference for 5, 9, 0 and 2.

Overall the indices are close to the expected minimum value of zero suggesting very little digit preference in the population. The Myers index for both sexes declined from 3.92 in 1996 to 3.06 in 200 and 2.89 in 2007. There are differences by population group and gender. In 1996 Myers indices are highest among the African, followed by the Asians, then Colored and lowest amongst the Whites. The situation is somewhat changed for 2001 in that Myers indices are highest amongst the African, followed by the Colored, then the Asians and lowest amongst the White population. In 2007 the Myers index is highest

among the Asian population, followed by the Colored, then the African and lowest among the White population.

In term of gender, Myers indices for males are higher than similar values for females in 1996 with the exception of the Asian population. In 2001 Myers indices for males are higher than similar values for females for Africans and Whites and the opposite is the case for Coloreds and Asians. Myers indices for males declined from 4.19 in 1996 to 2.96 in 2001, whereas the Myers index for females declined from 3.7 to 3.18. In addition, Myers indices for males for Coloreds and Asians have increased from 3.29 to 4.14 and 2.64 to 4.27 respectively. In general digit preference was more pronounced amongst the male population than the female population in 1996 and the opposite is true in 2001.

TABLE 2. MYERS INDICES FOR NORTH WEST PROVINCE BY RACE, 1996 AND 2001

	1996			2001			2007			2011		
	Male	Female	Both	Male	Female	Both	Male	Female	Both	Male	Female	Both
0	0.21	0.24	0.23	-0.42	-0.55	-0.49	0.18	0.29	0.24	-0.33	0.12	-0.37
1	-0.20	-0.03	-0.11	0.89	0.58	0.74	-0.14	-0.08	-0.11	0.75	1.01	0.67
2	-0.09	0.03	-0.03	-0.22	-0.18	-0.20	-0.24	0.08	-0.08	-0.02	0.25	-0.01
3	-0.33	-0.03	-0.18	-0.06	-0.16	-0.11	0.19	-0.09	0.05	-0.04	0.01	-0.07
4	0.42	0.43	0.42	-0.31	-0.47	-0.39	-0.29	-0.57	-0.43	-0.38	-0.38	-0.38
5	-0.65	-0.79	-0.73	-0.23	-0.17	-0.20	0.41	0.84	0.62	-0.23	-0.32	-0.23
6	1.14	0.87	1.00	-0.02	0.13	0.06	0.05	-0.37	-0.15	-0.06	-0.20	-0.04
7	-0.34	-0.32	-0.33	-0.23	-0.05	-0.14	-0.51	-0.53	-0.52	-0.12	-0.25	-0.06
8	0.32	0.29	0.30	0.07	0.32	0.19	-0.09	-0.22	-0.15	0.08	-0.19	0.13
9	-0.49	-0.69	-0.59	0.52	0.56	0.54	0.43	0.64	0.53	0.34	-0.06	0.36
Me	4.19	3.72	3.92	2.96	3.18	3.06	2.53	3.71	2.89	2.34	2.78	2.31

Just like with Whipple's Index, there are variations by population groups (see Tables 3 to 6). In 1996 the Myers indices for the African population were higher for males than females indicating that the incidence of digit preference was more pronounced among African males than African females. The same is observed among the White population. The opposite is observed among the Colored and Asian populations. In 2001 the scenario presented in the preceding sentence is preserved.

TABLE 3. MYERS INDICES FOR NORTH WEST BY POPULATION GROUP AND SEX FOR NORTH WEST, 1996

	African			Colored			Asian			White		
	Male	Female	Both	Male	Female	Both	Male	Female	Both	Male	Female	Both
0	0.18	0.21	0.20	0.11	0.45	0.29	0.26	0.65	0.45	0.58	0.62	0.60
1	-0.18	0.01	-0.08	-0.23	-0.63	-0.44	-0.23	-1.20	-0.69	-0.50	-0.25	-0.37
2	-0.06	0.07	0.01	-0.55	-0.15	-0.34	0.03	0.36	0.19	-0.28	-0.42	-0.35
3	-0.32	0.00	-0.16	-0.02	-0.01	-0.02	0.06	-0.39	-0.16	-0.51	-0.40	-0.45
4	0.48	0.47	0.48	0.11	0.58	0.35	-0.47	1.15	0.31	-0.23	-0.17	-0.20
5	-0.73	-0.88	-0.81	-0.37	-0.29	-0.33	0.39	-0.44	-0.01	0.20	0.15	0.18
6	1.21	0.92	1.06	0.96	0.51	0.73	0.24	1.33	0.76	0.42	0.30	0.36
7	-0.38	-0.34	-0.36	0.14	0.09	0.12	-0.56	0.35	-0.12	0.16	-0.20	-0.02
8	0.31	0.28	0.30	0.32	0.03	0.17	0.34	-1.19	-0.39	0.49	0.42	0.46
9	-0.50	-0.75	-0.63	-0.48	-0.57	-0.53	-0.07	-0.62	-0.33	-0.34	-0.06	-0.20
	4.36	3.94	4.08	3.29	3.31	3.30	2.64	7.69	3.41	3.71	2.99	3.19

TABLE 4. MYERS INDICES FOR NORTH WEST BY RACE AND SEX, 2001

	African			Colored			Asian			White		
	Male	Female	Both	Male	Female	Both	Male	Female	Both	Male	Female	Both
0	-0.49	-0.66	-0.57	0.54	0.41	0.48	1.40	-0.22	0.65	0.16	0.49	0.33
1	0.97	0.64	0.80	0.72	0.50	0.61	-0.45	0.90	0.17	0.02	-0.01	0.01
2	-0.20	-0.18	-0.19	-0.23	-0.34	-0.29	-0.10	0.10	-0.01	-0.47	-0.13	-0.30
3	-0.04	-0.13	-0.08	-0.47	-0.24	-0.35	-0.74	-0.85	-0.79	-0.20	-0.44	-0.33
4	-0.32	-0.50	-0.41	-0.38	-0.48	-0.43	0.51	-0.54	0.03	-0.17	-0.15	-0.16
5	-0.25	-0.18	-0.22	-0.43	0.22	-0.10	0.02	0.41	0.20	0.08	-0.16	-0.04
6	-0.02	0.11	0.05	-0.02	-0.14	-0.08	-0.33	-0.19	-0.27	-0.03	0.36	0.17
7	-0.25	-0.06	-0.15	-0.55	0.01	-0.27	-0.26	0.05	-0.12	0.06	-0.03	0.01
8	0.05	0.35	0.20	0.27	-0.05	0.11	0.21	-0.14	0.05	0.31	0.08	0.19
9	0.54	0.61	0.58	0.53	0.12	0.32	-0.26	0.49	0.08	0.23	0.00	0.11
	3.12	3.42	3.26	4.14	2.51	3.04	4.27	3.89	2.36	1.72	1.87	1.65

TABLE 5. MYERS INDICES FOR NORTH WEST BY RACE AND SEX, 2007

	African			Colored			Asian			White		
	Male	Female	Both	Male	Female	Both	Male	Female	Both	Male	Female	Both
0	0.20	0.26	0.23	-0.22	1.83	0.83	-0.36	1.85	0.55	0.10	0.30	0.20
1	-0.12	-0.11	-0.11	-0.89	-0.91	-0.90	-1.43	-2.89	-2.02	-0.14	0.56	0.21
2	-0.24	0.12	-0.07	-0.47	1.29	0.43	-1.21	-1.36	-1.27	-0.08	-0.57	-0.33
3	0.14	-0.11	0.02	0.61	0.32	0.46	-3.13	4.35	-0.07	0.91	-0.18	0.36
4	-0.30	-0.70	-0.50	0.05	0.34	0.20	-2.18	3.53	0.16	-0.07	0.45	0.19
5	0.44	0.86	0.65	2.89	0.59	1.72	2.01	-1.21	0.69	-0.48	0.81	0.17
6	0.05	-0.33	-0.14	-0.62	0.75	0.07	2.93	-1.31	1.20	0.05	-0.87	-0.42

	African			Colored			Asian			White		
	Male	Female	Both	Male	Female	Both	Male	Female	Both	Male	Female	Both
7	-0.57	-0.54	-0.55	0.43	-1.31	-0.46	4.32	0.91	2.93	-0.37	-0.39	-0.38
8	-0.16	-0.20	-0.18	0.01	-2.79	-1.41	-0.78	-2.76	-1.59	0.72	0.24	0.48
9	0.57	0.76	0.66	-1.79	-0.11	-0.94	-0.18	-1.12	-0.56	-0.64	-0.35	-0.49
	2.78	3.99	3.10	7.97	10.24	7.41	18.51	21.28	11.04	3.56	4.71	3.22

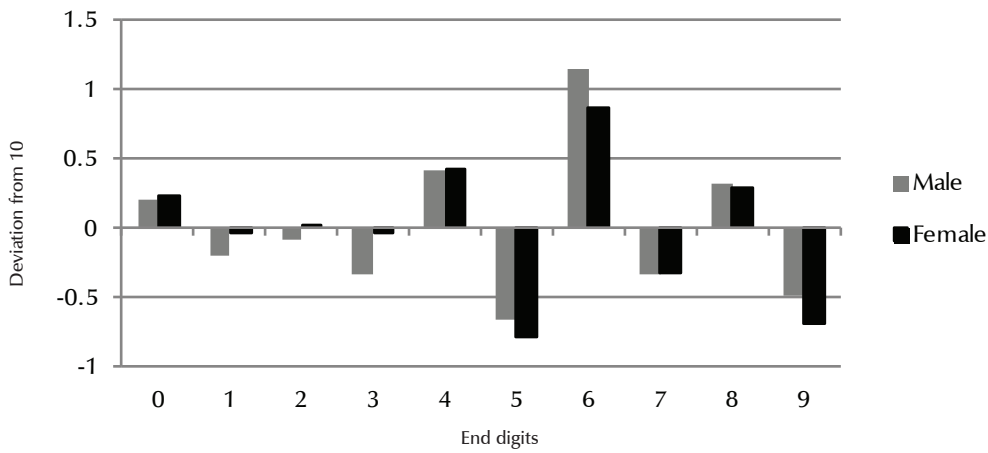


FIG. 3. DEVIATION FROM 10 OF END DIGITS, NORTH WEST, 1996

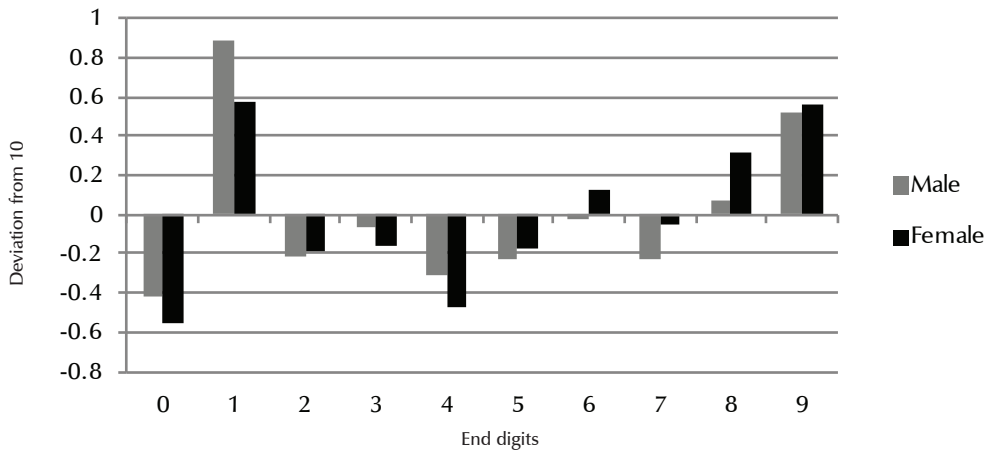


FIG. 4. DEVIATION FROM 10 OF END DIGITS, NORTH WEST, 2001

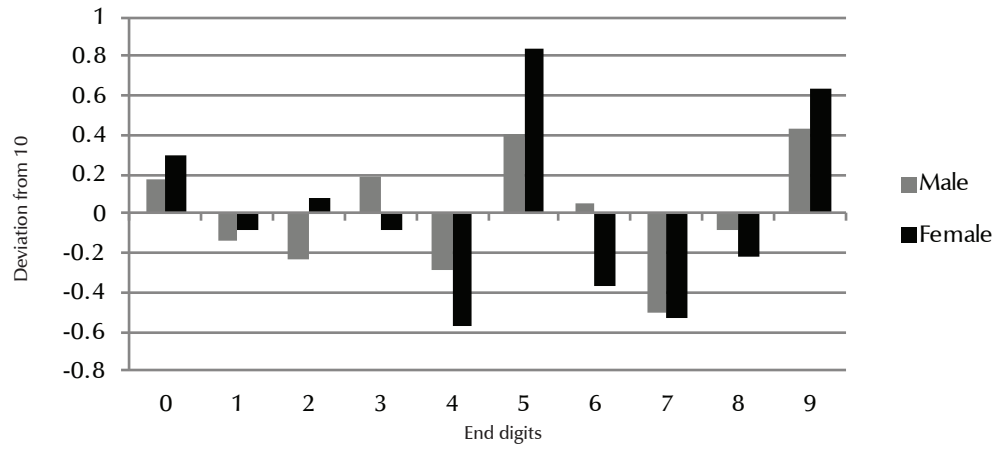


FIG. 5. DEVIATION FROM 10 OF END DIGITS, NORTH WEST, 2007

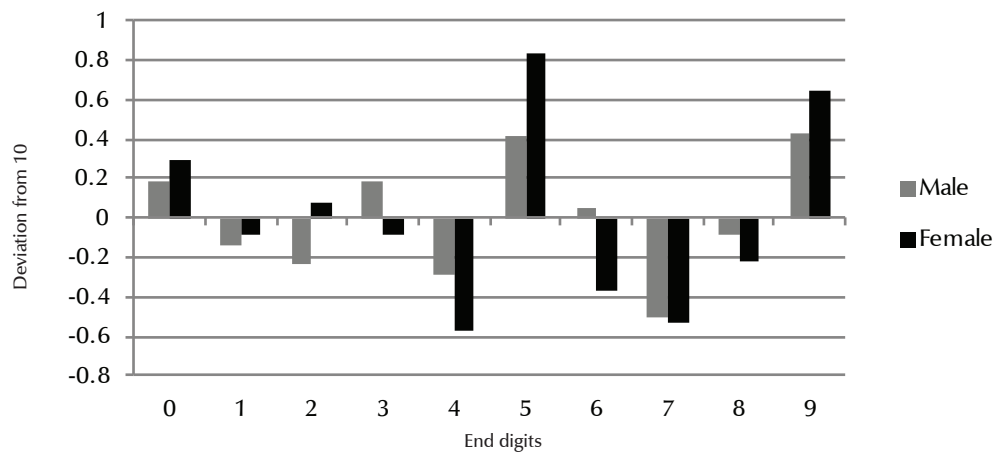


FIG. 6. DEVIATION FROM 10 OF END DIGITS, NORTH WEST, 2011

TABLE 6. MYERS INDICES FOR NORTH WEST BY POPULATION GROUP AND SEX, 2011

	African			Colored			Asian			White		
	Male	Female	Both	Male	Female	Both	Male	Female	Both	Male	Female	Both
0	-0.42	-0.01	-0.46	0.10	0.83	0.17	-0.40	0.23	-0.40	0.59	1.33	0.52
1	0.82	1.02	0.72	0.83	1.03	0.71	0.19	0.86	0.24	0.04	0.88	0.14
2	-0.04	0.23	-0.03	-0.22	0.15	-0.14	0.35	-0.01	0.09	0.21	0.51	0.14
3	-0.02	0.02	-0.05	-0.40	-0.42	-0.47	0.03	-0.97	-0.32	-0.19	0.10	-0.17
4	-0.38	-0.41	-0.40	-0.33	-0.11	-0.21	-0.57	0.53	-0.19	-0.38	-0.17	-0.29
5	-0.25	-0.30	-0.23	-0.21	-0.57	-0.36	1.35	0.56	1.10	-0.24	-0.51	-0.28

	African			Colored			Asian			White		
	Male	Female	Both	Male	Female	Both	Male	Female	Both	Male	Female	Both
6	-0.09	-0.17	-0.04	0.12	-0.54	-0.10	0.27	-0.01	0.28	0.13	-0.41	0.04
7	-0.09	-0.20	-0.03	-0.59	-0.63	-0.47	-0.12	0.07	0.02	-0.21	-0.71	-0.24
8	0.12	-0.15	0.16	0.33	0.06	0.37	-0.54	-0.94	-0.60	-0.27	-0.57	-0.15
9	0.35	-0.03	0.36	0.37	0.20	0.48	-0.54	-0.34	-0.22	0.32	-0.46	0.28
	2.57	2.53	2.49	3.50	4.54	3.48	4.36	4.51	3.48	2.58	5.66	2.24

United Nations Joint Age-Sex Score

THE UNITED Nations Joint Age-Sex Accuracy Index allows the examination of the sex ratios and age ratios for each sex.

Sex Ratio Analyses

TABLE 7 and figure 2 presents the age specific sex ratios for North West Province obtained from the 1996 and 2001 censuses and 2007 Community Survey. Variations in sex provide an opportunity for assessing the extent of age misreporting in a population by five-year age groups. Under normal circumstances, it is expected to have a slightly higher number of males at young ages and a consistently greater number of females at older ages since females live longer than males (United Nations 1973). The age-specific sex ratios reveal an excess of females over males in all age groups with exception to age range 25–54. The deficit of females in the age range 25–54 may be attributed to migration of a substantially higher proportion of males in these prime working age groups in search of jobs and other opportunities in life. The 1996 and 2001 census data also show a decreasing trend in the sex ratio with the advancement of age, reaching its lowest level in the oldest age group, 85 years and older (see Fig. 2). The increasing excess of females with the advancement of age is also due to a lower risk of death among females than males with the advancement of age (see Fig. 6), as there is no evidence to show that males more often than females tend to move out of the province in the older age groups. The table, however, indicates some fluctuating sex ratios that can be attributed to errors in age data. The sex ratio for the province increased from 96.8 in 1996 to 98.6 in 2001 and to 101.2 in 2007. The increase in the sex ratio could be due to the interplay of an improvement in male mortality, an increase in female mortality, over-reporting of males, underreporting of females, in-migration of males and out migration of females or a combination of any of these.

TABLE 7. AGE SPECIFIC SEX RATIOS FOR NORTH WEST, SOUTH AFRICA, 1996–2011

Age groups	1996	2001	2007	2011
0–4	98.8	99.4	101.3	103.0
5–9	99.5	100.2	100.6	102.4
10–14	98.2	98.4	93.9	105.9
15–19	96.4	98.3	102.3	104.5
20–24	91.8	100.7	103.6	107.0
25–29	94.6	97.3	104.3	110.3
30–34	101.4	102.7	114.9	115.1
35–39	109.6	107.1	109.8	107.3
40–44	106.6	113.6	110.3	103.6
45–49	106.7	103.2	118.9	104.0
50–54	97.3	103.7	106.2	105.8
55–59	93.0	96.1	96.2	98.3
60–64	77.6	82.9	82.1	90.2
65–69	78.8	72.9	75.0	81.2
70–74	71.6	70.4	64.4	72.6
75–79	65.5	61.9	59.7	61.7
80–84	56.2	55.0	52.1	55.7
85+	48.8	42.7	49.2	46.4

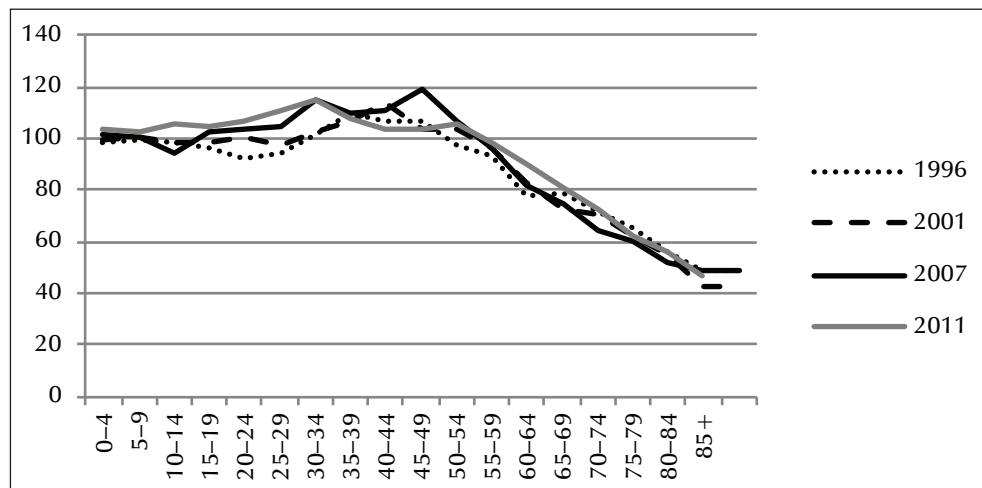


FIG. 7. AGE SPECIFIC SEX RATIOS FOR NORTH WEST, SOUTH AFRICA, 1996–2011

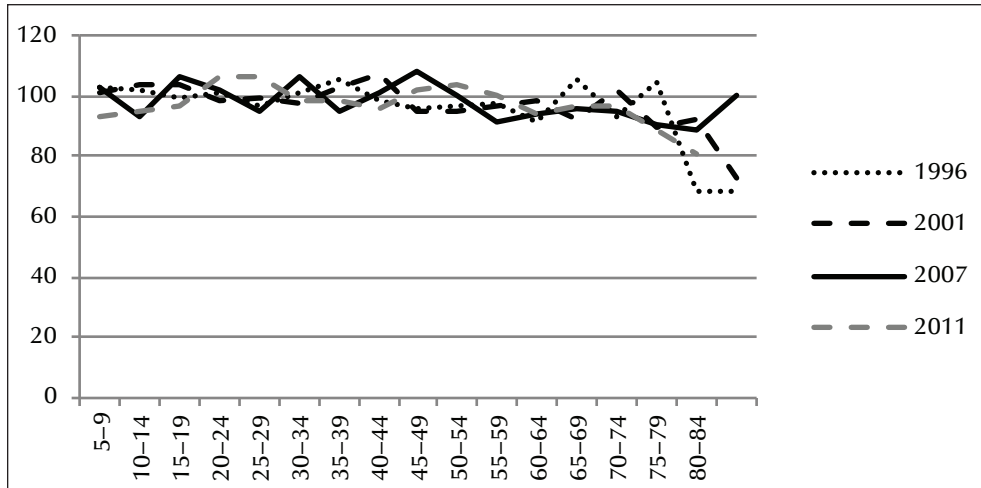


FIG. 8. AGE RATIOS FOR MALES, NORTH WEST, 1996-2011

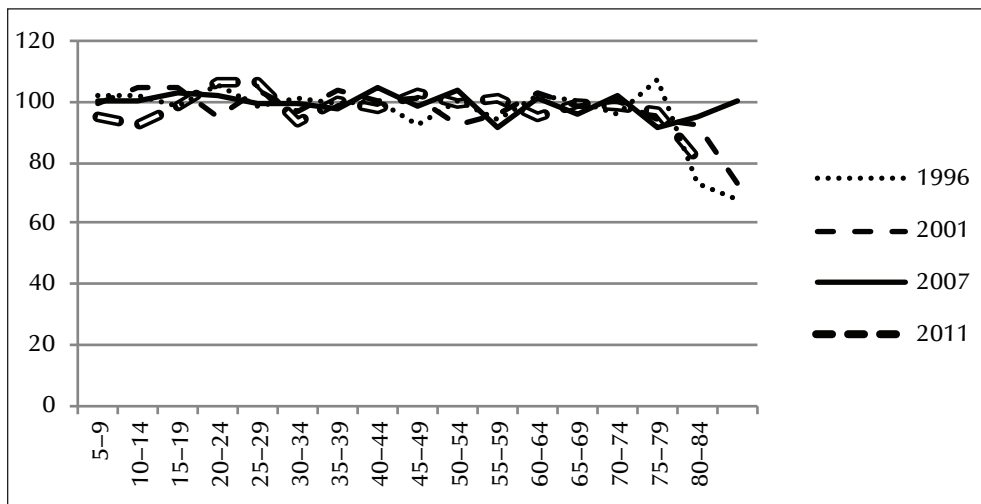


FIG. 9. AGE RATIOS FOR FEMALES, NORTH WEST, 1996-2011

Age Ratios and Age Accuracy Indexes

AGE RATIO is computed in this study as the percentage of the population in a given five-year age group to the average of the preceding and following age groups (Shryock and Siegel 1976). According to the scheme of evaluation 100 is the point of balance; age ratios higher than 100 indicate an

over-enumeration of the particular age group, while an age group is deemed under-enumerated if its age ratio goes below 100. The results presented in Table 6 and illustrated in figures 7 and 8 show that there are distortions in the reported age data of North West Province. In 1996 over-enumeration took place at age groups 5–9, 10–14, 20–24, 30–34, 35–39, 65–69, 75–79 among the male population and age groups 5–9, 10–14, 20–24, 30–34, 75–79 among the female population. In 2001 over-enumeration was observed in age groups 10–14, 15–19, 35–39, 40–44, 70–74 among the male population and age groups 10–14, 15–19, 25–29, 35–39, 40–44, 45–49, 60–64 among the female population. In 2007 the following age groups were over-enumerated: 5–9, 15–19, 20–24, 30–34, 40–44, 45–49, 50–54 in the case of males and 15–19, 20–24, 40–44, 50–54, 60–64, 70–74 in the case of females. In 2011 over-enumeration is observed in age groups 20–24, 25–29, 45–49 and 55–59 in both male and female populations. The male population in 2011 also shows over-enumeration in age group 50–54.

TABLE 8. AGE RATIOS FOR NORTH WEST PROVINCE, SOUTH AFRICA, 1996, 2001 AND 2007

Age group	1996		2001		2006		2011	
	Male	Female	Male	Female	Male	Female	Male	Female
0 – 4								
5 – 9	103.0	102.0	100.7	99.3	103.0	100.0	93.1	94.7
10 – 14	102.4	102.2	104.2	105.1	92.7	100.1	94.7	92.5
15 – 19	99.6	98.4	103.7	105.0	106.5	102.6	97.0	98.8
20 – 24	101.4	105.6	98.3	95.4	102.4	102.1	106.4	106.8
25 – 29	97.0	98.4	99.3	103.7	95.3	99.2	106.5	106.6
30 – 34	101.5	101.0	97.5	96.5	106.7	99.2	98.7	93.6
35 – 39	105.6	99.8	103.3	103.7	94.7	97.3	98.2	100.7
40 – 44	98.5	100.3	107.7	100.1	101.1	104.4	95.7	97.7
45 – 49	95.4	92.1	95.2	101.4	108.4	99.0	102.1	102.7
50 – 54	96.9	100.7	95.1	92.2	100.2	103.9	103.8	99.7
55 – 59	98.0	94.0	96.9	95.8	91.1	91.9	100.0	101.6
60 – 64	91.3	102.5	98.6	103.0	94.3	101.0	93.6	95.2
65 – 69	105.1	100.4	91.9	98.6	95.5	95.9	96.5	99.3
70 – 74	93.3	96.4	101.9	100.0	94.5	102.2	96.3	98.7
75 – 79	104.6	107.2	89.7	94.6	90.4	91.2	88.7	96.7
80 – 84	68.1	72.8	91.9	92.7	88.5	94.9	80.9	81.2
85+								

Table 7 presents the average sex ratio score, the age ratio score for males and females and the UN Joint Age-Sex Score by population group for North West Province for the 1996 and 2001 Population Censuses. In terms of race the UN

Joint Age-Sex Score indicate that the reporting is better for the Whites, followed by Africans, then Colored and lastly Asians. The high score for the Asian population is probably due to the high sex ratio score which is in turn influenced by the dominance of males in the Asian population. The age ratio score is slightly higher for the male populations than female population suggesting that age reporting is less satisfactory for the males than the females. The sex ratio score is three times higher than the suggested “minimum” indicating that available data on sex is also unsatisfactory.

Overall the UN Joint Age-Sex Score in the NW Province was 25 in 1996 declining to 23 in 2001 and increasing to 29.6 in 2007. The UN Joint Age-Sex Score for 2011 is 24.7. According to the UN, this indicates that the reported age distribution in the province could be described as “inaccurate.” The results also indicate that fluctuations in the score and the reporting of age-sex data between 1996 and 2011 have largely remained unchanged. The overall quality however, is better than that found in neighboring countries and is expected for a developing country. The values are significantly lower than what is found in many African populations but are comparable with those from Southern Africa (Bekele 2006; Dahiru and Dikko 2013; Mba 2003, 2004; Yazdanparast et al. 2012). For instance, the UN Joint Age-Sex Score for Malawi in 2008 was 30.9 (Malawi Government 2011) and for Nigeria in 2006 was 50 (Dahiru and Dikko 2013).

TABLE 9. UN JOINT AGE-SEX SCORE FOR NORTH WEST PROVINCE FOR SOUTH AFRICA, 1996, 2001 AND 2007

Province / population group	Year	ASRS	AARSM	AARsf	UN Score
TOTAL	1996	5.11	5.32	4.42	25.05
	2001	5.28	4.19	3.54	23.58
	2007	6.8	5.7	3.0	29.2
	2011	5.0	5.3	4.4	24.7
African	1996	5.35	5.84	4.92	26.83
	2001	5.57	4.17	3.54	24.43
	2007	7.057	5.733	3.326	30.2
	2011	5.2	5.7	4.7	26.1
Coloured	1996	8.00	8.52	5.20	37.72
	2001	4.90	6.85	6.18	27.73
	2007	32.099	42.062	23.096	161.5
		5.7	6.3	4.6	28.0

Continued on next page

Table—*Continued*

Province / population group	Year	ASRS	AARSM	AARSF	UN Score
Asian	1996	11.21	5.27	9.90	48.81
	2001	14.12	10.96	7.22	60.54
	2007				
	2011	30.9	15.0	6.4	114.2
White	1996	4.92	5.42	3.82	24.00
	2001	4.77	6.04	5.90	26.26
	2007	12.998	15.177	9.623	63.8
	2011	3.5	4.6	4.6	19.8

These so-called “inaccuracies” in the age and sex data are based on the assumption that “accurate” age data are rectangularly distributed, and that age-specific sex ratios decline over the life cycle in an even manner. Departures from these patterns will result in “inaccurate” data patterns.

Moreover, the UN Joint Age-Sex Score has other limitations as a summary measure of the accuracy of age sex data. Among these are “the failure to take into account the expected decline in the sex ratio with increasing age, and of real irregularities in age distribution due to migration, wars, and epidemics, as well as normal fluctuations births and deaths . . . also, the considerable weight given to the sex ratio component in the formulae. . .” (Shryock and Siegel 1976, 126). Despite these and other limitations, the measure is useful for making rough comparisons and distinctions between and among populations regarding the accuracy in the censuses of reporting age by sex (*ibid.*).

Discussion

THIS PAPER has explored the quality of the reported age statistics for North West Province in South Africa using data collected in 1996, 2001 and 2011 censuses and 2006 survey. The accuracy of the single-year age distributions was determined by Whipple’s and Myers indices of digit preference. For five year age distributions, the United Nations Joint Age and Sex Score and stable population analyses were used to assess the quality of age data. The analysis was carried out for the total provincial population, and the four population groups.

The results confirm that the quality of the age-sex population distribution in South Africa as a whole and North West Province in particular is better than in most African countries. This is expected given the level of social and economic development in the country. However the reported age-sex statistics are dis-

torted and the nature and pattern of age misreporting are similar to those found in other developing countries.

The reported age structure closely resembles what is expected in a developing country going through fertility transition. The reported age distribution reveals a young age structure and further indicates that the population of North West Province is getting older with the passage of time. The elderly population shows a modest increase overtime. The proportion of the population below age 15 is also on the decrease, whereas the economically active population (population in age range 15–64) appears to be increasing.

The consequences of the above-described population structure on the social and economic development of North West Province are gloomy. At a family or household level, given the high proportion of population below age 15, the age structure suggests that the heads of the family or household have the enormous responsibility to see to it that there is enough food, clothing and shelter for everyone. At the national level the high and increasing dependency burden implies a lot of pressure on the government to make available such social services as schools and health facilities and create employment for the young population. All this indicates that at both levels more effort will be on consumption rather than on investment.

Lastly, the examination of the reported age distribution further revealed the preponderance of females in the population. This was reflected by the low overall sex ratios and low ASSRS, and raises the question “where do our men go?” A number of factors were suggested to account for this anomaly and these included such variables as sex differentials in enumeration, labor migration and high male mortality.

The government of South Africa should pay more attention to improving the vital registration system in the country. This is because the records produced by this system serve two essential purposes. First, the individual records establish a person’s civil status and the facts on which it is based. Proofs of these facts such as age, place of birth or death are essential documents required for many official purposes in the society in which a person lives or has lived. Second, the information collected through the vital registration system provide the basis for the computation of indices of fertility, mortality and migration, among others, showing changes in population size and composition and important features regarding the health of a population. Third, information on the registration records provides an independent database on which census results can be matched and assessed.

Limitation of the Study

The study was based on demographic data published by Statistics South Africa on its website. This data is available through SUPERWEB software. It was later discovered that the reported age distributions do not include the “not stated” category. Further inquiries revealed that census data that is made available to users, including researchers, have been adjusted in one way or the other. Unfortunately the methods used to adjust the data are not provided.

Conclusion

IT CAN be concluded that the quality of age statistics in South Africa is good and it has improved since the emergence of new political dispensation in the 1990s that culminated in the election of the first democratic government in 1994. □

References

- Arriaga, E. E. et al. 1994. *Population Analysis with Microcomputers: Presentation of Techniques*. Vol. One, Bureau of the Census, Washington.
- Bailey, M. and T. J. Makannah. 1996. An evaluation of the age and sex data of the Population Censuses of Sierra Leone, 1963–1985. *Genus* 52 (1–2): 191–199.
- Bekele, B. 2006. Analysis on the quality of age and sex data collected in the two population and housing censuses of Ethiopia. *Ethiopian Journal of Science* 29 (2): 123–132.
- Berkowitz, P. 2012. Census 2011: The (incomplete) (probably inaccurate) sum of us. *Daily Maverick*, 31 October 2012.
- Blacker, J. G. C. 1969. Some Unsolved Problems in Data Collection in Africa. *IUSSP International Population Conference London 1969*, 227–285.
- Byerlee, D. and G. Terera. 1981. Factors Affecting Reliability in Age Estimation in Rural West Africa: A Statistical Approach. *Population Studies* 35 (3): 477–90.
- Caldwell, J. C. 1966. A Study of Age Misstatement among Young Children in Ghana. *Demography* 3 (2): 477–490.
- Caldwell, J. C. and A. A. Igun. 1971. An Experiment with Census-type Age Enumeration in Nigeria. *Population Studies* 25 (2): 287–302.
- Cleveland, L. n.d. Cross-Census Assessment of Age-Sex Ratios: An application of newly updated UN assessment guidelines to micro data census samples from the IPUMS International. www.statistics.gov.hk/wsc/cps103-p11-S.pdf.

- Dahiru, T. and H. G. Dikko. 2013. Digit preference in Nigerian censuses data of 1991 and 2006. *Epidemiology Biostatistics and Public Health* 10, 2.
- Ewbank, D. C. 1981. *Age Misreporting and Age-Selective Underenumeration: Sources, Patterns, and Consequences for Demographic Analysis*. Committee on Population and Demography, National Academy Press, Washington, D.C.
- Fajardo-González, J., L. Attanasio, and J. T. Ha. 2014. An Assessment of the Age Reporting in the IPUMS-I Microdata. Paper submitted for presentation at the 2014 Annual Meeting of the Population Association of America.
- Malawi Government. 2011. *2008 Malawi Population and Housing Census*. National Statistical Office, Zomba.
- Mba, C. J. 2003. Assessing the reliability of the 1986 and 1996 Lesotho census data. *Journal of Social Development in Africa* 18 (1): 111–126.
- . 2004. Challenges of population census enumeration in Africa: an illustration with the age-sex data of the Gambia. *Research Review*, n.s., 20.1: 9–19.
- Moultrie, T., R. Dorrington, A. Hill, K. Hill, I. Timaeus, and B. Zaba, eds. 2013. *Tools for Demographic Estimation*. IUSSP, Paris.
- Mukherjee, B. N. and B. K. Mukhopadhyay. 1988. A Study of Digit Preference and Quality of age data in Turkish Census. *GENUS* 44 (1/2): 201–227.
- Nagi, M. H. et al. 1973. Digit Preference and Avoidance in the Age Statistics of Some Recent African Censuses: Some Patterns and Correlates. *International Statistical Review* 41 (2): 165–174.
- Ntozi, J. P. M. 1978. Patterns of age and coverage errors in tropical African surveys. *East African Journal of Rural Development* 11 (1 and 2): 116–144.
- Palamuleni, M. E. 1996. Age Mis-reporting in Malawian Censuses and Sample. *Southern African Journal of Demography* 5 (1): 11–17.
- Poston, D. L. et al. 2000. An Analysis of the Quality of the Age and Sex Data of the Republic of Korea and its Provinces, 1970 and 1995. Unpublished mimeograph, Department of Sociology, Texas A&M University.
- Shryock, H. S. and J. S. Siegel. 1976. *Methods and Materials in Demography*. Condensed edition. New York: Academic Press.
- Simelane, S. E. 2002. An overall and demographic description of the South African Population based on the Census 1996. Occasional Paper Series, 2002/1, Statistics South Africa, Pretoria.
- Spoorenberg, T. 2007. Quality of age reporting: Extension and application of the modified Whipple's index. *Population-E* 62 (4): 729–742.
- . 2009. Is the Whipple's index really a fair and reliable measure of the quality of age reporting? An analysis of 234 censuses from 145 countries. Paper presented at the IUSSP Conference, Morocco.
- Statistics South Africa. 1998. The Population of South Africa, Population Census 1996: Age Tables of South Africa and its Provinces. Report No. 1, Pretoria.
- Statistics South Africa. 2003. *Census 2001: How the count was done*. Pretoria. Statistics South Africa.

- Statistics South Africa. 2012. *Census 2011: How the count was done*. Pretoria. Statistics South Africa.
- Talib, Ali M. S., M. S. Hamid, and K. M. Zin. 2007. Age Reporting Behaviour : A Case Study of 1991 and 2000 Population and Housing Censuses, Malaysia. Department of Statistics, Malaysia.
- Udjo, E. O. 1996. An evaluation of the reported age-sex distribution of Botswana in the 1996 Census. *Southern African Journal of Demography* 6 (1): 1–8.
- Unisa, S., L. K. Dwivedi, R. S. Reshmi, and K. Kumar. 2009. Age reporting in Indian Census: An insight. Paper presented at the IUSSP Conference, Morocco.
- United Nations. 1952. Accuracy Test for Census age distributions tabulated in five year and ten year groups. *Population Bulletin of the UN* (New York) 2: 59–79.
- United Nations. 1955. Methods of Appraisal of Quality of Basic Data for Population Estimates. Manuals on Methods of Estimating Population, Manual II. Series A, Population Studies, No. 23, 31–54.
- United Nations. 1973. The Determinants and Consequences of Population Trends: New Sample of Findings on Interaction of Demographic, & Economic and Social Factors. Vol. 1. New York.
- United Nations. 1980. Principles and Recommendations for Population and Housing Census. Statistical Papers. New York.
- United Nations. 1990. *1988 Demographic Yearbook*. United Nations. New York.
- Van de Walle, E. 1966. Some Characteristic Features of Census Age Distributions in Illiterate Populations. *American Journal of Sociology* 71 (5): 549–557.
- Yazdanparast, A., M. A. Pourhoseingholi, and A. Abadi. 2012. Digit preference in Iranian age data. *Italian Journal of Public Health* 9 (1): 64–70.

Abstract

Age Reporting in the North West Province, South Africa, 1996–2011

The aim this study is to examine the nature and patterns of the reported age statistics in the North West Province using the 1996, 2001 and 2011 South African population censuses. First, the enumerated age-sex population distributions from the censuses are evaluated using such indices as Whipple's, Myers and United Nations Age-Sex Score to highlight the nature and patterns of age misreporting in the province. Second, essential age-sex features of the population and their implications for development planning are highlighted. The study indicates that the quality of the reported age statistics in North West Province, like in the other provinces in South Africa, is quite good and in contrast to findings from most other Sub-Saharan populations. Furthermore, the reported age statistics show a preference of terminal digits 6, 4, 8, 0 and 2, implying a preference of ages ending in even numbers. Further research should establish the factors responsible for this pattern.

Keywords

age, census, North West Province, South Africa