ebph

ORIGINAL ARTICLES

Digit preference in Nigerian censuses data of 1991 and 2006

Tukur Dahiru[1], Hussaini G. Dikko[2]

ABSTRACT

BACKGROUND: censuses in developing countries are prone to errors of age misreporting due to ignorance, low literacy levels and other social, economic and cultural factors. Ages are commonly rounded with great affinity for 0 or 5. This tendency to digit preference and/or avoidance results in age heaping or concentration of ages at certain digits. This study examined the extent of digit preference in the Nigerian census data of 1991 and 2006.

METHODS: this study utilized age data from the 1991 and 2006 Nigerian censuses reported in single years. The Whipple and Myers indices were used to determine the extent of digit preference.

RESULTS: both the 1991 and 2006 census data showed the expected pattern of errors, with Whipple and Myers indices being beyond acceptable levels. The Whipple index for 1991 and 2006 was 293 and 251 respectively, while the Myers index was 62.3 and 67.1 respectively. There was a strong preference for terminal digits 0 and 5, followed by 8 whereas terminal digits 1 and 9 were strongly avoided.

CONCLUSIONS: the quality of age data in Nigerian census data is poor as a result of misreporting and no significant improvement or difference was observed between 1991 and 2006 censuses.

(1) Department of Community Medicine, Ahmadu Bello University, Zaria, Nigeria
(2) Department of Mathematics, Ahmadu Bello University, Zaria, Nigeria

CORRESPONDING AUTHOR: Tukur Dahiru, Department of Community Medicine, Ahmadu Bello University, Zaria, Nigeria. Tel: +234 803 4501105.
e-mail: tukurdahiru2012@gmail.com
DOI: 10.2427/8843

INTRODUCTION

Nigeria has a long history of census taking. The first census, confined to the old Lagos colony, was conducted in 1866. Subsequent censuses, gradually involved other parts of the country and by 2006 12 censuses have been conducted in Nigeria [1, 2].

Censuses conducted across Africa have been documented to suffer from errors in the form of under-count or age misreporting [3].

Information about age is an important variable collected in a census because it provides basis for further demographic analyses and estimations such population estimation, fertility estimation [3]. Conceptually, age seems simple to collect in any demographic inquiry but the literature indicates that it is not as simple as it seems [4]. Several authors [3, 5, 6] have documented reasons why age reporting in censuses is not accurate, including ignorance of the exact age, conscious or subconscious preference

for certain ages (e.g. "unlucky" 13 or ages 21 and 65 in the United States), carelessness in reporting and recording, tendency to round-up ages and other social, economic and cultural factors such as literacy, urban location.

Ewbank attempted to illustrate the implication of age misreporting on parental survival technique for estimation of mortality. Results demonstrated that an age exaggeration of 2.5 years will bias the estimated life expectancy upward by approximately the same amount [5]. Age exaggeration has implications on other indices that are calculated based on age of respondents. In the assessment of data quality used in demographic and health surveys, it was found that there was intentional misrecording of information in order to exclude some women for in-depth interviews. The effect of this misrecording was an overestimation of the total fertility rate of 2-4% and underestimation of under-5 mortality rate of 2-4% [7].

To date, assessment of quality of census data has been conducted in some African countries [8, 9], but not in Nigeria. Therefore, the aim of this study was to assess the age data of the 1991 and 2006 Nigerian censuses.

METHODS

Reported age in single years for the 1991 and 2006 Nigerian census data were used and disaggregated by sex. Two methods of analyses were used, the Whipple and Myers blended indices [10, 11].

The Whipple index is a summary measure of age-heaping for ages ending in 0 or 5 [10]. It is calculated by summing persons in the age range 23 to 62 years inclusive, and calculating the ratio of reported ages ending in 0 or 5 to one-fifth of the total sample. This measure assumes linearity in the distribution of ages in each five-year age range. The index varies between a minimum of 100, indicating no concentration/preference at all, and a maximum of 500, if only 0 and 5 were reported. The United Nation's recommended standard for measuring age heaping on digits 0 and 5 classifies the quality of data from very accurate to very bad (Table 1).

The Meyer index was developed to detect preference and avoidance of all the terminal digits, from 0 to 9 [11]. It is based on the principle that in the absence of age preference and avoidance, the aggregate population of each age ending in one of the digits 0 to 9 should represent 10% of the total population. The method produces a reference value for each of the nine digits as well as a summary index for the ten terminal digits or the index of preference (IP). Theoretically, the index has a minimum value of 0 (when deviations from 10% is zero) which represents no heaping to a maximum of 90 (when heaping of all reported ages is at a single digit e.g., zero).

RESULTS

The Whipple indices for the sexes and total sample indicates how the quality of age data for the two censuses were very bad (Table 2). Digit preference and avoidance as measured using the Myers index shows there was general preference for digits 0 and 5 while others were avoided with digits 1 and 9 being most avoided (Table 3). There was general improvement in the quality of age reporting using Myers index; avoidance of certain digits declined and preference for certain digits also declined. For instance, Myers index for digit 1 declined from 20.2 to 16.8 while avoidance for digit 1 declined from -5.6 to 5.2. Table 4 shows the rank order of digit preference/avoidance for the 10 digits. It clearly shows the declines (or improvements) in the quality of the age data. Table 4 is a summary of rank order of digit preference and avoidance. Some reversal of the digit preference and avoidance was found when two numbers were combined: (2, 8), (4, 6) and (3, 7).

DISCUSSION

This study has brought to the fore the level of quality Nigeria's censuses data, at least as far age reporting is concerned. From the analyses of Nigerian censuses data of 1991 and 2006, it can be seen that the quality is poor with respect to Myers' and Whipple's indices.

Results indicate a concentration on terminal digits 0 and 5 and very poor data quality. However, looking at sex disaggregation, the index is slightly better for males compared to that of females. The quality of Whipple's index has been linked to the level of literacy of the respondent or the individual declaring the age i.e. the better the literacy, the better

TABLE 1

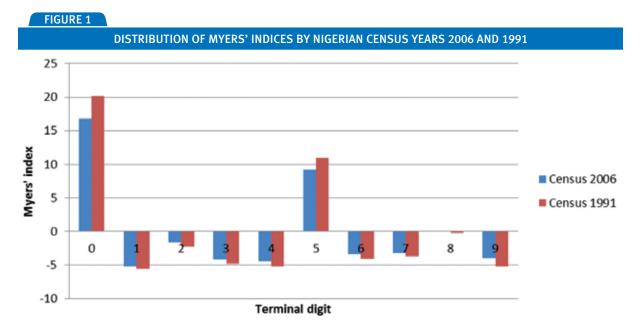| UNITED NATIONS RECOMMENDATION FOR AGE HEAPING AS IDENTIFIED BY WHIPPLE'S INDEX		
WHIPPLE'S INDEX	QUALITY OF DATA	DEVIATION FROM PERFECT
<105	very accurate	<5%
105-110	relatively accurate	5-9.99%
110-125	OK	10-24.99%
125-175	Bad	25-74.99%
>175	very bad	>=75%

TABLE 2

WHIPPLE'S INDICES FOR NIGERIAN CENSUSES, 1991 AND 2006			
YEAR OF CENSUS	WHIPPLE'S INDEX		
	MALE	FEMALE	TOTAL
1991	250	307	293
2006	247	255	251

TABLE 3

PATTERN OF DIGIT PREFERENCE AND MYERS' INDICES FOR NIGERIAN CENSUSES, 2006 AND 1991						
TERMINAL DIGIT	CENSUS YEAR					
	2006			1991		
	MALE	FEMALE	TOTAL	MALE	FEMALE	TOTAL
0	16.4	17.3	16.8	18.4	21.9	20.2
1	-5.1	-5.4	-5.2	-5.2	-6.0	-5.6
2	-1.6	-1.7	-1.6	-2.0	-2.6	-2.3
3	-4.1	-4.3	-4.2	-4.4	-5.1	-4.8
4	-4.2	-4.5	-4.4	-5.0	-5.5	-5.2
5	9.0	9.4	9.2	10.7	11.3	11.0
6	-3.3	-3.6	-3.4	-3.9	-4.3	-4.1
7	3.2	-3.2	-3.2	-3.3	-4.1	-3.7
8	-0.2	0.2	0.0	-0.4	0.0	-0.2
9	-3.9	-4.1	-4.0	-4.9	-5.5	-5.2
Myers' Index	50.7	67.4	67.1	58.3	66.3	62.3
Index of Preference	25.4	33.7	33.5	29.2	33.1	31.2

TABLE 4

RANK ORDER OF DIGIT PREFERENCE AND AVOIDANCE FOR NIGERIAN 2006 AND 1991 CENSUSES			
TERMINAL DIGIT	CENSUS YEAR		LEVEL OF PREFERENCE/ AVOIDANCE
	2006	1991	
0	+16.8	+20.2	Most preferred
5	+9.2	+11.0	
2, 8	-1.6	-2.5	
4, 6	-7.8	-8.9	
3, 7	-7.7	-8.5	
1, 9	-10.1	-10.8	Most avoided

FIGURE 1

DISTRIBUTION OF MYERS' INDICES BY NIGERIAN CENSUS YEARS 2006 AND 1991

the index [3]. The two Nigerian censuses, have documented poor literacy levels among the females compared to males, and this difference may explain the difference in Whipple's indices between the two sexes [12, 13]. Overall, there was slight improvement in the Whipple's index between the censuses; decreasing from 293 in 1991 census to 251 in 2006 census. This decline was noticeable in both males and females. This observation could be attributed to several factors such as literacy level of respondents, use of good, diligent interviewers, use of calendar of local events to gauge respondent's age in cases of unknown age. It is assumed here that between 1991 and 2006 there was an improvement in literacy levels of Nigerians and people became more aware of their exact ages. This can be deduced from the fact that more males than females are literate in Nigeria and the Whipple's index also shows this sex differentials as females had higher Whipple's than males for the two censuses. Also, the programme launched by the National Population Commission to improve birth registration coverage must have had some impact albeit not estimated here. Furthermore,

during the 2006 census calendar of local events were developed as aid to exact age estimation and this also would have had some effect in age reporting.

The Myers index results show a general pattern of terminal digit preference and avoidance. A strong preference for terminal digits 0 and 5, and 1 and 9 was noted. Digits 2 and 8 showed the least deviations; digit 8 showed the least followed by digit 2. There were declines in negative deviations between the two census data for the two digits; digit 8 showed no deviation in 2006 census from -0.2 in 1991 to 0.0 in 2006 while digit 2 declined from -2.3 in 1991 to -1.6 in 2006. For these two digits, this observation conforms to what was observed by Nagi, Stockwell and Snavely [6].

CONCLUSIONS

It can be concluded that the quality age data is poor as a result of misreporting and it seems that there was no significant improvement of the indices with the passage of time between 1991 and 2006 censuses.

References

[1] Ohadike PO. Development in Africa: A Socioeconomic Demographic Perspective, UNRIPS, 1998, Accra

[2] Dahiru T. National Census in Nigeria: Past Experiences

and Future Considerations. Abia Journal of Population Studies 2010; 1: 106-18

[3] Kpedekpo GMK. Essentials of Demographic Analysis for

Africa, London Heinemann Publication, 1982

[4]   Yazdanparast A, Amin MP, Abadi A. Digit preference in Iranian age data. Italian Journal of Public Health 2012; 1: 64-8

[5]   Ewbank D. Age Misreporting and Age Selective Under-enumeration. Patterns and consequences for Demographic Analysis. (Report of Committee on Population and Demography). Washington DC: National academy Press. 1981

[6]   Nagi MN, Stockwell EG, Snavely LM. Digit preference and avoidance in the age statistics of some African censuses: some patterns and correlates. International Statistical Review 1973; 2: 165-74

[7]   Institute for Resource Development. 1990. An Assessment of DHS-I Data Quality. DHS Methodological Report No.1. Columbia, Maryland. Institute for Resource development/Macro System, Inc

[8]   Bailey M, Makannah TJ. Patterns of digit preference and avoidance in the age statistics of some recent African censuses:

1970-1986. Journal of Official Statistics 1993; 3: 705- 15

[9]   Bailey M, Makannah TJ. An evaluation of age and sex data of the population censuses of Sierra Leone: 1963-1985. Genus 1996; 1-2: 191-9

[10] Hobbs FB. Age and sex Composition. In: Jacob A. Siegel and David A. Swanson. (Eds). The Methods and Materials of Demography (pp 125-173). USA: Elsevier Academic Press, 2004

[11] Myers RJ. Errors and bias in the reporting of ages in census data. Transaction of the Actuarial Society of America 1940 (October); 41: 104

[12] National Population Commission, Nigeria 1991 Census Data: Analytic Report at National Level, Abuja. 1998

[13] National Population Commission, 2006 Population and Housing Census of the Federal Republic of Nigeria. National and State Population and Housing Tables. Priority Tables Volume I. Abuja, 2009

*