Trends in DHS data quality in Sub-Saharan Africa: An analysis of age heaping over time in 34 countries between 1987 and 2015

Tara Stones and Mark Lyons-Amos, School of Health Sciences and Social Works, University of Portsmouth

Working Paper of Portsmouth-Brawijaya Centre for Global Health, Population, and Policy

Number 4

Abstract

This paper evaluates one aspect of data quality within DHS surveys, the accuracy of age reporting as measured by age heaping. Other literature has explored this phenomenon, and this analysis build on previous work, expanding the analysis of the extent of age heaping across multiple countries, and across time. This paper addresses this by making a comparison of the magnitude of Whipple's index of age heaping across all Demographic and Health Surveys from 1986-2015 in Sub-Saharan Africa. We use a random slope multilevel model to evaluate the trend in the proportion of respondents within each survey rounding their age to the nearest age with terminal digit 0 or 5. We find that broadly speaking the trend in the proportion of misreported ages has remained flat, in the region of 5% of respondents misreporting their age. We find that Nigeria and Ghana have demonstrated considerable improvements in age reporting quality, but that a number of countries have considerable increases in the proportion of age misreported, most notably Mali and Ethiopia with demonstrate increases in excess of 10% points.

Introduction and motivation

Much attention has been paid to ensuring that basic data quality within DHS surveys, such as age, is correctly measured.

Most of the measures produced from DHS surveys depend to a critical degree on reports of ages and dates (Pullum 2006). At the most basic level, inclusion of women age 15-49 in the surveys depends on accurate reports of the ages of women near the boundaries of that age interval in the household survey. The inclusion of children under five (or another specified age) for the questions about child health, immunizations, and nutrition also depends on accurate reports of their birth dates. Many measures are age-specific, such as estimates of age-specific fertility rates and infant and child mortality rates. Estimates of levels and trends in such rates may be affected by misreporting of ages and dates of birth for a woman and her children, or dates of death for her children. Misreporting of adult ages often takes the form of a preference for numbers ending in the digits 0 or 5. Since standard age intervals begin with those digits, misreporting can shift women into the next higher age interval. Age displacement of children can seriously distort estimates of current levels and recent trends in fertility and mortality. Additionally, age heaping can have implications for the quality of analyses into other phenomena, such as cause specific death rates (Benjamin et al. 2013). This has lead to a plethora of studies evaluating the quality of basic demographic data in the DHS in a variety of contexts (e.g. Johnson et al.,2009, Robles and Goldman 1999, Channon et al. 2011, Cleland 1996).

As has been noted, there have been considerable efforts to improve the quality of demographic data in other areas such as birth weight (Channon et al. 2011). In this case, increasing utilisation of written demographic information made possible by greater levels of literacy and healthcare systems engagement has led to improvements over time in demographic data quality.

This analysis expands on previous works, by expanding the range of countries evaluated as well as capturing trends across time, to account for potential structural change which may improve the quality of retrospective data (Channon et al. 2011). This builds on the work of Pullum (2006) which was comprehensive but necessarily limited to DHS up to 2003. Data for this analysis are drawn from all available Demographic and Health surveys in the sub saharan Africa region.

As such, this paper addresses two major research aims

1. Capturing the overall trend in the quality of age recall data across multiple waves of DHS surveys
2. Evaluate the extent of cross national variation in the extent of age heaping

Literature Review

"The difficulty of estimating ages and dating events is recognized as one of the major problems with data collection and analysis in Africa" (Barbieri et al, 2005, pp626)

Age heaping is a frequently encountered phenomenon in demography and global health studies and presents significant problems for accurate collection and analysis of data. Age heaping or age *preference* is the tendency for people to incorrectly report their age or date of birth. Studies have shown that individuals' heaping behaviours have a tendency to favour certain ages, most commonly those ending in '0' or '5' as these are *attractive* numbers (A'Hearn et al 2009). Preference for certain years when reporting date of birth has also been observed, notably for those ending in '0' or '5'.

Pullum (2008) found in a ten year period (1993-2003) that heaping in weight and height measurements also favoured these digits. How age heaping behaviours are developing over time is key to comprehending how and why the phenomenon occurs: by contrasting past and more recent literature using Demographic and Health Survey (DHS) data, contributing factors may be identified.

The 1973 report of Nagi, Stockwell and Snavley examined analysis of digit preference in African censuses and focused on the underlying factors which explained differences within African nations, particularly between Islamic and non-Islamic states, socio- economic indicators, and gender. "in the majority of cases the extent of age heaping is greater for women than for men, with the differences being especially pronounced for the Islamic countries." (Nagi et al, 1973, pp167). Religious factors were considered here as the concept of age can often be formed and defined by the socio- cultural context of upbringing and belief systems. They found that modernization had an impact on the level of errors in age reporting across both Islamic and non- Islamic states. This suggests socio- economic indicators are also a substantial determinant of data accuracy. These authors concluded that any social and economic measures used in their study should be considered tentative at best owing to the potential inaccuracy of data and small sample size. Most intriguingly age heaping was more pronounced in sub- Saharan African countries with "only half of the 26 indexes of preference as high as the Islamic" (Nagi et al, 1973, pp167). Cleland (1996) notes that demographic investigation in less developed countries was rarely undertaken prior to 1946 and recognised a need for change in methods used for data collection and analysis. This was particularly true when contrasted with pronounced improvements in data from the Americas and Asia.

Gage (1995) extended the focus from age appraisal to consider significant life stages including first marriage, first birth, and first intercourse. She found that heaping was a common occurrence in work undertaken across North Africa and Asia especially in terms of preference towards years ending in 0 and 5 when estimating first marriage. While Burkina Faso, Cameroon, Nigeria, and Rwanda had significantly high heaping indices of year of first birth there was little evidence of heaping in Latin America. The exception was Nigeria with high indices of heaping years of first union and of first birth, representing a significant deviation. Gage explains Nigeria's frequent age heapings as a result of the use of older data (from 1990) and poor knowledge of dates. The measurement of women having a first birth before age 20 was the only constant in multiple DHS surveys. Furthermore, "heaping on the year of first birth is most severe in Pakistan" (Gage, 1995, pp24) indicating heaping was not exclusive to sub- Saharan Africa, where the geography of regions did little to predict heaping indices as there were diverse results within continents. Subsequent authors (al- Haddad et al2013) found that Nigeria has followed the trend of other African Nations in its preference for 0 and 5 but there are pronounced differences between genders and a North-South discrepancy, potentially arising from economic and educational disparities. "Age heaping has been established as a proxy for population numeracy and northern Nigeria is both less economically developed and has potentially more heterogeneity in quality of primary and secondary education compared to southern Nigeria." (al- Haddad et al, 2013, pp18).

Becker and Diop- Sidibé (2003) considered the use of new techniques to reduce inaccuracies, with particular focus on postpartum variables rather than just age of respondents in the survey population. More recent versions of the DHS record additional variables, "births; pregnancies; contraceptive use and discontinuation; postpartum amenorrhea; postpartum abstinence; and breastfeeding, for each month from January of the fifth or sixth year before the survey to the month of interview." (Becker and Diop- Sidibé, 2003, pp127). This calendar system is adjusted to fit the life events of women instead of relying on standardized calendars in an attempt to increase accuracy. They found that these calendars did indeed reduce heaping at reported durations of *6, 12, 18,* and

24 months for breastfeeding and amenorrhea. (Becker and Diop- Sidibé, 2003, pp129). However, the calendar method only reduced heaping with regard to abstinence in two out of five countries, which may not be considered a substantial gain. Conceptions and application of chronology are better translated in the use of a calendar that is inclusive of postpartum variables but is not likely to be practical for collecting adult population ages. Haandrikman, Rajeswari, Hutter, and Ramesh (2004) used a similar technique of alternate measures of *timepaths* using 'local calendars' as did Bello (2012) in his work with outpatients in Nigeria. This framework involves referencing local events and festivals which corresponded to the individual's personal life events. This method was relatively successful in that respondents were extremely capable of remembering events as their memory was triggered by the use of a local calendar. From work in Bangladesh, Callahan and Becker (2012) also reported little evidence of heaping when using a calendar to record contraceptive practice. It was a way of taking individuals *back in time* which concurrently resulted in less duration heaping (Haandrikman et al, 2004, pp358).

Johnson, Grant, Khan, Moore, Armstrong and Sa (2009) considered fieldwork- related factors in DHS surveying. Factors such as language, interviewer behaviour, sex, number of visits to the household, and time spent in the fieldwork group were examined. The impact of the interviewer in the field was integral to accuracy of data: "where male interviewers are assigned to conduct the household interview as well as male interviews, male interviewers may have a tendency to "rush" household interviews, causing increases in incomplete and inconsistent responses" (Johnson et al, 2009, pp8). Gendered issues in the field were prominent, as sex- specific age displacement at the household level was identifiable. Both the sex of interviewers and the sex of respondents had an impact on accuracy to some extent as did behaviour and attitudes towards fieldwork. Considering the impact of these factors, there are potential hazards to the accuracy of data arising from human error. Errors in reporting are not always due to respondents' social, cultural, economic and or educational circumstances: a fundamental factor, that of language, affected all areas of the research process as translation to local languages and communication between interviewer and respondent was a foundation. "The most striking finding is that, in most of the countries that have these data, missing information on age or date of birth is highest when the interviewer and respondent use the same language, but it is a language that is not included on the list of main languages spoken in the country" (Johnson et al, 2009, pp45). Language was something of a barrier but these authors suggest screening interviewers based on their local language skills with careful observation through the fieldwork period. This would combat both language and interviewer indifference as confounding factors, while care with translations of questionnaires and consistency in household visits has potential to decrease inaccuracies.

Barbieri et al (2005) looked at age differences between spouses and contraceptive practice in sub-Saharan African countries as sex and age were anticipated to be principal cultural identifiers. Sex and age were found to be substantially linked to power relations. This caused men to frequently overestimate their ages and underestimate that of their spouses as advanced age was a way to perpetuate gender based power. Features of male power were contradicted by the introduction of formal education that allowed women to negotiate or impose innovative behaviour in respect to contraceptive practice. "we postulated that the age difference between the spouses was likely to influence individual and couple decision-making power with respect to contraception." (Barbieri et al, 2005, pp643). Evidently age was not a simple passive or static concept as it perpetuated strong cultural behaviours. Siegel and Swanson (2004) looked at the methods and components of demography itself. They discovered an age preference for ages ending in 0 and 5 but noted the causes of these patterns may vary from one culture to another. "In some cultures, certain numbers may be specifically avoided (e.g., 13 in the West and 4 in East Asia)." (Siegal and Swanson, 2004,

pp136). Number avoidance and number preference is a phenomenon in censuses even in more developed countries: “such heaping occurred in the 1970 and 1980 censuses of the United States, and both heaping on ages and on years of birth ending in 0 and 5 were evident in the 1990 census of the United States.” (Siegal and Swanson, 2004, pp140). This was considered a result of asking only year of birth, and accuracy was evidently more attainable when the question of current age was also included. This suggests issues with the method of reporting, in particular the wording of questions or having different cultural definitions of age in childhood (no 0 age or not regarded as a member of the household). They also found a correlation between populations with low educational status and pronounced age heaping.

Pullum (2006) assessed DHS data from 1985 to 2003 to identify evidence of misreporting, particularly on the effect of education and schooling levels. “we find consistently negative and statistically significant effects of mean education on misreporting. The strongest effect, by far, is on the measures of incompleteness. For example, it is far more common to give an age but not a birth date, or to give ages and birth dates that are inconsistent in contexts where the mean level of education is low.” (Pullum, 2006, pp54). Incompleteness, to answer only some questions, was a serious issue and was often combined with upward and downward transfers outside the age range eligibility (15-49) of the surveys. Pullum (2006) further found that age heaping in the household surveys was seriously affected by respondent’s schooling, suggesting a *poor knowledge of ages and dates of birth*. sub- Saharan Africa was above average especially in “age transfers—downward for women, upward for women, and upward for children—and incompleteness of all kinds of dates.” (Pullum, 2006, pp54), but promisingly, over time, appears to be less of a problem. Pullum (2006) comments on how age heaping in itself may not produce much distortion unless it has a *systematic component of bias.* This suggests that distortion could be avoidable, for example in the case of transfers if changes were made in analytical strategies to cope with it.

A’Hearn, Baten, and Crayen (2009) analysed age heaping by considering human capital through history across Europe and taking into account the circumstances for comprehension of age. This study is in contrast to those with focus on contemporary developing countries as it considers the social capabilities of centuries-old European groups. A study of grave monument inscriptions in the Roman Empire by Richard Duncan- Jones found “age heaping on multiples of five at rates not dissimilar to those for medieval Tuscany or developing countries of the 1950s and 1960s, and higher for women than men.” (A’Hearn, Baten, and Crayen, 2009, pp786). Age preference and gender disparities are a trend regardless of location or period, appearing as a common human idiosyncrasy. A’Hearn and colleagues (2009) conclude though that *individual human capital, collective number discipline, and administrative capacity* (which reflects low age heaping) is not dependant on mass schooling or industrialization. This is a somewhat puzzling finding as modern societal structures, such as formal education and economic development, have been suggested in previous studies to be fundamental in determining age heaping patterns. They explored the question of whether the age heaping phenomenon reflects more *the characteristics of individuals or of the society they inhabit* (A’Hearn, Baten, and Crayen, 2009, pp789). Crayen and Baten (2010) examined global trends in numeracy from 1820-1949 and found in developing countries a link between low numeracy and vague ideas about date of birth which were potentially down to degree of schooling. This is in contrast to the work of A’Hearn and colleagues (2009) who suggest that numeracy is not dependent on formal structures of mass education, but sociocultural practice and that age heaping can in fact index basic numeracy. Crayen and Baten (2010) do however put forward other factors which influence numeracy (which in turn determines age heaping) “besides the state bureaucracy, market demand or schooling success is the frequent use of calendars or astrological elements as in the Chinese culture, for example.” (Crayen and Baten, 2010, pp89). This is interesting as it is in line with

the idea that numeracy can in fact be developed through other cultural channels. Pardeshi (2010) concurs with this view as he suggests that the use of calendars and astrology determine age heaping, and with the concept advanced by A'Hearn and colleagues of the use of human capital as a source element. Crayen and Baten (2010) also imply that infant protein malnutrition syndrome was (historically) and is (in today's poorest economies) a limiting factor in an adult's cognitive abilities (which can cause misreports in age) indicating poverty and underdevelopment are still prominent controversies in global health studies.

Considering the numerous factors that interact with age heaping and heaping at large, globally and historically, it appears that they cannot simply be reconciled. DHS data overall have provided detailed insights into developing countries but over time the methodologies and areas of focus have evolved. Research models can better cope with attitudes and behaviours in the field and the process, in recent years, allows for improved cultural translations. This has indeed reduced heaping in some areas and analyses from sub- Saharan Africa show some improvement in data accuracy, along with increased levels of development (spread of formal education, industrialization, and improved gender equality). By using a cross-cultural and historical evaluation we see how heaping is frequently associated with particular features such as numeracy, literacy, levels of schooling, attractiveness of certain numbers, gender, and development. There appears to have been a broad understanding and therefore adjustment for temporality being socially, culturally and economically defined: "A society in which individuals know their age only approximately is a society in which life is not governed by the calendar and the clock but by the seasonal cycle; in which birth dates are not recorded by families or authorities; in which few individuals must document their age in connection with privileges (voting, office-holding, marriage, holy orders) or obligations (military service, taxation); in which individuals who do know their birth year struggle to accurately calculate their age from the current year." (A'Hearn, Baten, and Crayen, 2009, pp 785). This improved awareness of the complexity of these interacting factors indicates that age heaping remains an important but mutable phenomenon.

Data

DHS are nationally representative, cross-sectional household surveys with multi-stage cluster sampling designs, and respondents are women of reproductive age (15-49 years). Exact details of the sampling designs are available on a country by country basis. For this analysis all countries within the sub Saharan Africa region are extracted. The countries included, the years of survey and the proportion of 0 and 5 terminal digits are presented in Table 1.

Method

Whipple's index of age heaping

This analysis uses Whipple's index of age heaping to measure data quality, defined in equation 1 (Henry et al 1976), where the number of respondents within the data with their age ending in digit *x* is defined as N_x. In all cases we use the age reported at interview accurate to the nearest year. This equation effectively denotes the proportion of ages ending with terminal digits commonly heaped, in this case 0 and 5.

$$Whipple = \frac{N_0 + N_5}{\sum_{x=0}^{9} N_x}$$

Eq 1.

Where no ages are heaped, we expect this index to take the value 0.2. Deviation from this number indicates some degree of terminal digit preference, for example 0.25 indicating that 5% of ages have been heaped at either a zero or five terminal digit.

Regression model

We specify the dependent variable in our model as the excess proportion of ages ending in 0 or 5 from (Whipple's index of heaping-0.20), denoted as y_{tj} where y is the proportion of respondents with heaped ages, indexed by year of survey t and country j. Survey years are hierarchically nested within countries. We therefore specify a multilevel model in the form of equation 2, where the logit of the index of heaping is a function of the year of the survey with intercountry variation captured a random effect parameter at the country level, v_j which is assumed to be Normally distributed with variance σ^2.

$$logit(y_{tj}) = \beta_0 + \beta_1 t + v_{0j}$$
$$v_{0j} \sim N(0, \sigma^2)$$

Eq 2.

To overcome the truncation of the proportion of age heaped at zero, we use a logit link to allow the specification of the model in the linear form of Equation 1. Since the proportion of ages heaped in some countries are low for some years, we performed sensitivity checks to ensure that our results are robust to the choice of link function using the arcsin link as an alternative. In all models tested, we extracted the same $\hat{\beta}$ regardless of link. Similarly, we explored different specifications of the year of survey parameter by introducing square and cubic terms for the effect of year to account for non-linearity. Neither of these specifications improved model fit on -2LogLikelihood significance tests. We therefore retain a linear specification as the most parsimonious model.

We performed tests for differences in the trend in the proportion of ages heaped over time by introducing a random slope parameter at the country level. This model is described in equation 3

$$logit(y_{tj}) = \beta_0 + \beta_1 t + v_{0j} + v_{1j} t$$
$$v_{0j} \sim N(0, \sigma^2), v_{1j} \sim N(0, \sigma^2)$$

Eq 3

In equation 3, the random effect parameter v_{1j} allows deviation from the overall trend in Whipple's index of heaping over time according to indexation by country j. This parameter is allowed to correlate with v_{0j}.

Model estimation is conducted by taking the logit of Whipples index of heaping, and using this as the response variable in a linear multilevel analysis. Models are estimated using MLwiN 2.36, with Restricted Iterative Generalised Least Square (2nd order Penalised Quasi Likelihood) estimation being used to account for the low number of observations per country.

Table 1: Countries for analysis and years of survey with Whipple's indicies

Country	Year of survey																													
	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
Benin											0.23					0.32					0.36						0.34				
Burkina Faso								0.25						0.27				0.27							0.26					
Burundi		0.23																							0.26					
Cameroon						0.26							0.26						0.26							0.26				
Chad												0.36							0.34											0.41
Comoros											0.26																0.28			
Congo																				0.23							0.21			
Cote dIvoire									0.26					0.25													0.23			
DRC																						0.23							0.23	
Ethiopia															0.30					0.37						0.35				
Gabon															0.23												0.20			
Gambia																												0.28		
Ghana			0.31					0.28					0.28					0.25					0.27						0.25	
Guinea														0.32						0.36							0.33			
Kenya				0.27				0.24					0.24					0.23					0.25						0.24	
Lesotho																			0.21					0.22					0.21	
Liberia	0.27																					0.24						0.24		
Madagascar							0.24					0.24							0.23					0.25						
Malawi							0.23								0.25				0.24						0.23					
Mali		0.23									0.31					0.31					0.34							0.34		
Mozambique												0.24						0.23								0.23				
Namibia							0.22								0.22							0.20						0.21		
Niger							0.36						0.33								0.37						0.37			
Nigeria					0.44													0.35					0.36					0.35		
Rwanda							0.23								0.24					0.22			0.23		0.22					0.22

Sao Tome and Principe																								*0.20*						
Senegal	*0.23*						*0.28*					*0.23*								*0.27*						*0.29*	*0.27*		*0.26*	
Sierra Leone																							*0.37*					*0.34*		
Swaziland																					*0.21*									
Tanzania							*0.25*				*0.24*			*0.26*						*0.25*					*0.25*					
Togo			*0.27*										*0.28*																*0.29*	
Uganda				*0.25*						*0.26*						*0.26*					*0.24*					*0.25*				
Zambia							*0.20*				*0.20*					*0.21*						*0.22*							*0.22*	
Zimbabwe			*0.23*						*0.23*					*0.22*							*0.21*					*0.22*				

Notes:

Figure within table indicate proportion of responses ending in 0 or 5.

Values within table above 0.20 indicate digit preference for 0 or 5

Empty spaces within table indicate survey not carried out in that year

Results

Results from the modelling procedure are presented in Table 2. We find no evidence of a secular trend toward an improvement in the proportion of ages heaped, with the coefficient from both Model I and Model II being both statistically non-significant and substantively small.

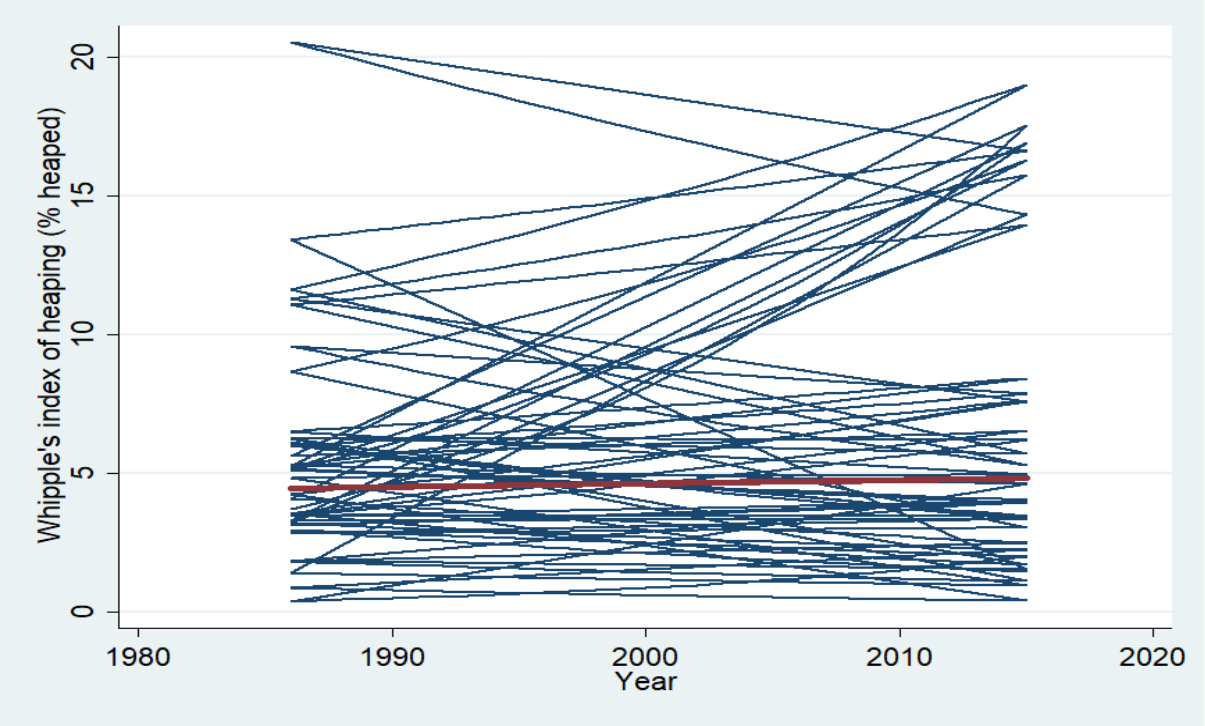
Table 2: Estimated multilevel model for proportion of ages heaped

	Model I: Random intercept model		Model II: Random slope model	
	Parameter estimate	95% confidence interval	Parameter estimate	95% confidence interval
Fixed effect parameters				
Survey year (centred)	0.004	(-0.004, 0.012)	0.003	(-0.010, 0.015)
Intercept	-3.032		-3.025	
Random effect parameters				
Random intercept v_{0j}	0.836	(0.416,1.256)	0.786	(0.390,1.182)
Random slope v_{1j}	-	-	0.001	(0.000, 0.001)
Intercept-slope covariance	-	-	0.004	(-0.007, 0.015)

Notes: Model based on 2nd order PQL RIGLS

The introduction of the random slope parameter proved to significantly improve model fit based on both likelihood test. The predicted values by country from Model 2 are presented in Figure 1. The overall trend in the proportion of age heaped in denoted by the red line within individual country trajectories denoted for each blue line. In general, there is a reasonable degree of clustering around the population line: the majority of countries have a portion of age heaped which is consistent over time, and in the range of between 2 and 6%.

Figure 1: Estimated median predicted proportion of ages heaped by country across survey year



Based on the predicted values of Whipple's index of heaping, we identify countries with substantial differences between survey years 1987 and 2015 based on the EB residuals from model 2. We identify 2 countries with large predicted decreases in the proportion of countries which we define as an absolute decrease of 4% points of more. Nigeria exhibits the largest decrease in the proportion of respondents reporting a heaped age, with a decline in the predicted value of Whipple's index of 6.22% points, with the only other country exhibiting a large substantive decrease in the proportion of respondents with a heaped age being observed in Ghana with a fall of 4.28% points.

A number of countries exhibit substantive increases in the proportion of respondents reporting a heaped age, again defined as an increase of 4% points or more between the predicted values of Whipple's index between 1987 and 2015. Sierra Leone, Chad and Ethopia demonstrate increases of 4.46% points, 7.38% points and 7.58% points respectively. We also note exceptionally large increases in the proportion of respondents with a heaped ages in excess of 10% points between 1987 and 2015: Mali exhibits and increase of 11.78 % points and Benin increases by 13.87% points

Conclusions and discussion

Data quality from retrospective sample surveys continues to be of major importance in social science, and basic demographic data is no exception. This paper therefore provides an assessment of the quality of age reported data within the DHS. We use all available DHS for the sub saharan Africa region to assess trends over time in the proportion of age reported which are heaped on terminal digits 0 and 5.

Out initial research hypothesis was that there may be a secular trends toward lower proportions of age heaped. This has been noted for other basic demographic information (Channon et al. 2011) where improvements in data collection procedures and provision of written information to increasingly literate populations has been a means of improving the accuracy of recalled data. However, in our analysis, we find no evidence of a significant decline in the proportion of ages heaped. That said the predicted probabilities are at a relatively low level for most countries, and are

not a substantial concern. We do however identify some major outliers: Nigeria and Ghana have considerable falls in the proportion of ages heaped, while there have been dramatic increases in Sierra Leone, Ethiopia and Chad. Tentatively, these changes can be explained by economic performance: relatively high growth rates in Nigeria and Ghana compared to moribund economic growth in Ethiopia and Chad exacerbated by internal conflict and violence which may have disrupted vital registration procedures. In any case, this study highlights the need to take into account country context when analysing data quality, even for standardised datasets such as the DHS.

References

A'Hearn, B., Baten, J., & Crayen, D. (2009). Quantifying quantitative literacy: age heaping and the history of human capital. *The Journal of Economic History,* 69(3), 783-808.

al-Haddad, B. J., Jedy-Agba, E., Oga, E., & Adebamowo, C. (2013). Age Heaping and Cancer Rate Estimation in Nigeria.

Barbieri, M., Hertrich, V., & Madeleine Grieve. (2005). Age Difference between Spouses and Contraceptive Practice in Sub-Saharan Africa. *Population (English Edition, 2002-), 60*(5/6), 617-654. doi:10.2307/4148187

Becker, S., & Diop-Sidibé, N. (2003). Does Use of the Calendar in Surveys Reduce Heaping? *Studies in Family Planning, 34*(2), 127-132.

Bello, Y. (2012). Error Detection in Outpatients' Age Data Using Demographic Techniques. *Int J Pure Appl Sci Technol*, *10*, 27-36.

Benjamin, J S, Elima Jedy-agba, Emmanuel Oga, Clement Adebamowo, and J S Benjamin. 2013. "Age Heaping and Cancer Rate Estimation in Nigeria Age Heaping and Cancer Rate Estimation in Nigeria."

Callahan, R., & Becker, S. (2012). The Reliability of Calendar Data for Reporting Contraceptive Use: Evidence from Rural Bangladesh. *Studies in Family Planning, 43*(3), 213-222

Channon, Andrew A. R., Sabu S. Padmadas, and John W. McDonald. 2011. "Measuring Birth Weight in Developing Countries: Does the Method of Reporting in Retrospective Surveys Matter?" Maternal and Child Health Journal 15 (1): 12–18. doi:10.1007/s10995-009-0553-3.

Cleland, J. (1996). Demographic Data Collection in Less Developed Countries 1946-1996. *Population Studies, 50*(3), 433-450.

Crayen, D., & Baten, J. (2010). Global trends in numeracy 1820–1949 and its implications for long-term growth. *Explorations in Economic History*, *47*(1), 82-99.

Gage, A. J. (1995). An assessment of the quality of data on age at first union first birth and first sexual intercourse for *phase II of the Demographic and Health Surveys program.*

Haandrikman, K., Rajeswari, N. V., Hutter, I., & Ramesh, B. M. (2004). Coping with time: using a local time-path calendar to reduce heaping in durations. *Time & Society*, *13*(2-3), 339-362.

Henry S. Shryock and Jacob S. Siegel, Methods and Materials of Demography (New York: Academic Press, 1976)

Johnson, K., Grant, M., Khan, S., Moore, Z., Armstrong, A., Sa, Z. (2009). Fieldwork- Related Factors and Data Quality in the Demographic and Health Surveys Program.

Leckie, G. and Charlton, C. (2013). runmlwin - A Program to Run the MLwiN Multilevel Modelling Software from within Stata. Journal of Statistical Software, 52 (11),1-40

Nagi, M., Stockwell, E., & Snavley, L. (1973). Digit Preference and Avoidance in the Age Statistics of Some Recent African Censuses: Some Patterns and Correlates. *International Statistical Review / Revue Internationale De Statistique, 41*(2), 165-174. doi:10.2307/1402833

Pardeshi, G. S. (2010). Age Heaping and Accuracy of Age Data Collected During a Community Survey in the Yavatmal District, Maharashtra. *Indian Journal of Community Medicine : Official Publication of*

Indian Association of Preventive & Social Medicine, *35*(3), 391–395. http://doi.org/10.4103/0970-0218.69256

Pullum, T. W. (2006). An assessment of age and date reporting in the DHS Surveys 1985-2003.

Pullum, T. W. (2008). An Assessment of the Quality of Data on Health and Nutrition in the DHS Surveys, 1993- 2003.

Siegel, J. S., & Swanson, D. A. (ed) (2004). The methods and materials of demography.