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

Thomas Spoorenberg[1]
Population Division, DC2-Room 1908
Department of Economic and Social Affairs
United Nations
New York, NY 10017, USA

spoorenberg@un.org
thomas.spoorenberg@gmail.com

Abstract:
The Whipple's index is the most widely applied measure of the quality of age reporting. Measuring only preference for age ending by 0 and 5, it is however largely accepted as a fair and reliable measure of the quality of age reporting. This assertion is here tested based on the analysis of 234 census single-years of age returns from 145 countries. The original Whipple's index is compared to a recently proposed synthetic index, the total modified Whipple's index. It is 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 important. This measure does not account well for the quality of age reporting once it is improving and reaches better levels.

[1] The views and opinions expressed in this paper are those of the author and do not necessarily reflect the views of the United Nations. Its contents have not been formally edited and cleared by the United Nations.

1. Introduction

Age misreporting constitutes *"one of demography's most frustrating problems"* (Ewbank 1981: 88). To detect it, demographers have developed different methods to assess the quality of age data. Among these methods, age heaping indices have been proposed to detect the degree of preference or avoidance for specific digits in age reporting.

To evaluate the quality of age reporting, the United Nations selected in the *Demographic Yearbook 1955* the Whipple's index (*W*) for its simplicity and the wide use it has already found in other sources (United Nations 2007: 5). Since then, this index has been widely and thoroughly used to assess the quality of age reporting in census and survey data. While it measures only the attraction for age ending in 0 and 5, it is largely accepted as a fair and reliable measure of the general quality of age reporting. However, to the best of our knowledge, this widely accepted statement has never been tested.

Due to different assumptions at the base of the computation of the other indices measuring the quality of age reporting (Bachi index, Myers' index, Zelnik's index) *W* cannot be compared directly to them. However, the successive modifications made to *W* have resulted in the proposition of a new synthetic index accounting for the attraction/repulsion for the ten digits – the total modified Whipple's index (W_{tot}) (Spoorenberg 2007). Since this new index is based on the same assumption of the original Whipple's index (linearity and rectangularity over a 5-year age group), it allows conducting systematic comparisons of the results obtained by the application of *W*.

The objective of this short research note is to apply this new W_{tot} index to 234 single years of age census data conducted between 1985 and 2003 in 145 countries to measure the quality of age reporting and compare systematically these results to the values given by the original Whipple's index. As it is widely assumed that the digit preferences for 0 and 5 "are usually connected with other sources of inaccuracy in age statements and the [Whipple's] index can be accepted as a fair measure of the general reliability of the age distribution" (United Nations 2007: 6), the paper aims to test this assertion.

2. Methods

While simple, the Whipple's index is highly sensitive to age heaping on ages ending in 0 and 5. This index applies to single years of age returns between ages 23 and 62 inclusive. It is obtained by summing the number of persons in this age range, and calculating the ratio of reported ages ending in 0 or 5 to one-fifth of the total sample. It varies between 1 (indicating no preference for ages ending by 0 and 5) and 5 (indicative of a complete report on ages ending by 0 and 5):

$$W = \frac{[5 \cdot (P_{25} + P_{30} + P_{35} + \cdots + P_{55} + P_{60})]}{P_{23} + P_{24} + P_{25} + \cdots + P_{61} + P_{62}} \qquad (1)$$

where P_x is the population of age x in completed years.

Later modifications of the *W* index gave a more precise and synthetic mean to assess the accuracy of single-year age data. *W* measures only preferences for (or avoidance of) ages ending by 0 and 5 *indistinctively*. However, opposite effects of digit preferences and avoidances can potentially cancel each other out and affect the original Whipple's index. To remove this constraint, two modifications have been made to the original formulation of *W* (Roger et al. 1981; Noumbissi 1992).

Based on the same basic principles of the original Whipple's index (linearity and rectangularity over a 5-year age range), the latter modification allows to measure age heaping for all digits. Through the digit-specific modified Whipple's index (Noumbissi 1992), age heaping could hence being computed *distinctively* for all ending digits (and not the sole 0 and 5 as in the original Whipple's index).

Since the use of the ten digit-specific modified Whipple's indexes (W_i) is not convenient to compare changes through time and across countries, a new synthetic index, the modified total Whipple's index (W_{tot}), has been recently proposed (Spoorenberg 2007). It is computed as the sum of the absolute difference between the digit-specific modified Whipple's index (W_i) and 1:

$$W_{tot} = \sum_{i=0}^{9} |W_i - 1| \tag{2}$$

The complete information on age heaping for all digits is thus fully used. As such it provides an essential complement to the digit-specific W_i indices and gives a more accurate and sensitive measure of overall age reporting quality. Moreover, because W_{tot} is based on the same assumption as *W*, both indices can be compared. Initial application of the W_{tot} index has shown that the improvement of age reporting given by the original Whipple's index is partial since only digit preferences for ages ending in 0 and 5 are account for (Spoorenberg 2007).

For each census year, original Whipple's index (*W*) and W_{tot} index are computed for both sexes. In order to compare both indices and assess their difference, the values of the two indicators are normalized as follow:

$$N_i = \frac{|VO_i - MTV_i|}{MTV_i} \tag{3}$$

where *i* stands for *W* or W_{tot} indices; VO_i the observed value for *W* or W_{tot}; and MTV_i the maximum theoretical value for *W* or W_{tot} (i.e. 16 for W_{tot}; Spoorenberg 2007).

Since the minimum value of the original Whipple's index (*W*) is not 0 but 1, a first transformation is required before proceeding to its normalization. For *W*, the absolute difference from unity is first computed; this allows setting all *W* values to a new standard starting from 0. Hence, the maximum theoretical value for *W* is not anymore 5, but 4. These transformed *W* indices can then be normalized using formula (3) and compared to the normalized W_{tot}.

To assess the difference between the two indices, the normalized *W* (N_W) and W_{tot} (N_{Wtot}) values are plotted against each other (N_{Wtot} on *x*-axis; N_W on *y*-axis). If the normalized values given by both indices are identical (signifying hence that *W* is a fair and reliable measure of the quality of age reporting), countries should align on a diagonal (zero-percent difference); if these values differ, a country would appear above or below the zero-percent diagonal. A country below this diagonal indicates that the normalized value of *W* is lower than the normalized value of W_{tot}, and that *W* is a better measure of the quality of age reporting than W_{tot}. However, since W_{tot} uses the information on attraction/repulsion for the ten digits, it would be in most of cases lower than *W*. Above the zero-percent diagonal, a country presents N_{Wtot} lower than N_W, indicating that W_{tot} better assesses the quality of age reporting than W.

Plotting the N_W against N_{Wtot} gives also the opportunity to consider the quality of age reporting. A country with very accurate age reporting would be close to one (meaning that W or W_{tot} value is close to 0 and thus N_i close to 1). And the further from one, the more inaccurate the quality of age reporting.

The percentage difference between N_W and N_{Wtot} is then put in relation to N_W in order to see whether a relationship emerges. The idea is to consider if for a given quality of age reporting (measured by W), the percentage difference differs.

3. Data

Data used in this analysis are taken from the *Table 1a: Population by single years of age by sex and urban/rural residence: each census, 1983-2003* (United Nations n.d.). This dataset comprised 145 countries for which at least one single-year age returns are available until the age of 62. Single-year age returns are given for male, female and both sexes at three different levels: Total population, urban population and rural population. However, because data for a large number of countries are not provided by residence area, the present analysis focuses only on the total population level. In total, 234 census single years of age data are analyzed.

4. Results

The normalized values of W and W_{tot} for the six world regions according to UN publication's standard (Africa; America, North; America, South; Asia; Europe; Oceania) are presented on figure 1. Each panel of figure 1 presents a different world region (classified by alphabetical order). In each panel, the different diagonal lines indicate the difference in percentage between the normalized values (N_i) of the two indices (respectively 0, 5, and 10 percent) and give the opportunity to consider whether the quality of age reporting is better assessed by the W or the W_{tot} index. The normalized results are presented only for both sexes (figure 1), since the patterns for men and women are very similar.

For all world regions, N_{Wtot} are always almost equal or lower than $N_{W.}$ As expected no countries are below the zero-percent difference diagonal. The countries appear on the zero-percent difference line – indicating that W gives the same results than W_{tot} –, or above this diagonal – indicating that W_{tot} better assesses the quality of age reporting. Two countries present percentage differences close to 10 percent (Saint Helena: Trista da Cunha (Africa) and Pitcairn (Oceania)). However, it should not be given too much importance to these two exceptions. Both countries count very small population (respectively 296 and 66 inhabitants) which affects the distribution of the population over the ages and increases the impact of a small departure from linearity and rectangularity in the computation of W and W_{tot}. In general, the percentage difference between N_W and N_{Wtot} ranges between 0 and 5 percent and does not seem to be dependent of the quality of age reporting.

[Figure 1 about here]

To secure this conclusion, N_W is plotted against the percentage difference between N_W and N_{Wtot} (figure 2). Based on the results presented on figure 2, a relationship appears. The variation in the percentage difference increases with the quality of age reporting. At low level of age reporting quality (i.e. low N_W value), the percentage difference reduces, while at high quality of age reporting (i.e. high N_W value) the variability increases. The countries characterized by N_W lower than 0.7 exhibit percentage difference lower than 1 percent. With the increase of N_W, the variability of the percentage difference increases; for example, over a N_W of 0.95, the percentage difference ranges between almost 0 percent and close to 5 percent. These patterns are observed for all world regions.

That the variability of the percentage difference between N_W and N_{Wtot} increases when N_W increases means that the attractions for ages ending by 0 or 5 are the main causes to low quality of age reporting. In other words, when the quality of age reporting improved the attraction on 0 and 5 age-digit reduces. This verifies a classic and well-established fact.

At the same time, because the importance of the attractions on 0 and 5 age-digit reduces, the age report on the other age digits gains significance, explaining hence the increase in the variability of the percentage difference between N_W and N_{Wtot} at higher level of age reporting quality.

[Figure 2 about here]

The results plotted on figure 2 point to an important conclusion: the original Whipple's index cannot account well for the quality of age reporting once a certain level of quality of age reporting is reached. This rather simple result has however a direct consequence for the choice of a fair and reliable measure of the quality of age reporting.

5. Conclusion

Based on the analysis of 234 census returns by single-year of age from 145 countries, this research note has shown that the original Whipple's index (*W*) is a fair and reliable measure of the quality of age returns *only when the attractions on ages ending by 0 and 5 are important*. However, this measure cannot reflect completely the quality of age reporting once the attractions on 0 and 5 age-digits reduce. This means that, with the improvement of the age declarations in most of the countries around the world, the changes affecting the quality of age reporting can be only partly accounted for by the application of the original Whipple's index. If one wants to assess with more precision the quality of age reporting and its change through time the original Whipple's index is not a completely fair and reliable measure. In this case, the total modified Whipple's index offers a simple alternative which fully accounts for the changes in the attraction/repulsion on *all* age-digits.

6. References

Ewbank, D.C. 1981. *Age Misreporting and Age-Selective Underenumeration: Sources, Patterns, and Consequences for Demographic Analysis*. Washington: National Academy Press, Committee on Population and Demography, Report No. 4.

Noumbissi A. 1992. L'indice de Whipple modifié : une application aux données du Cameroun, de la Suède et de la Belgique. *Population* 47 (4) : 1038-1041.

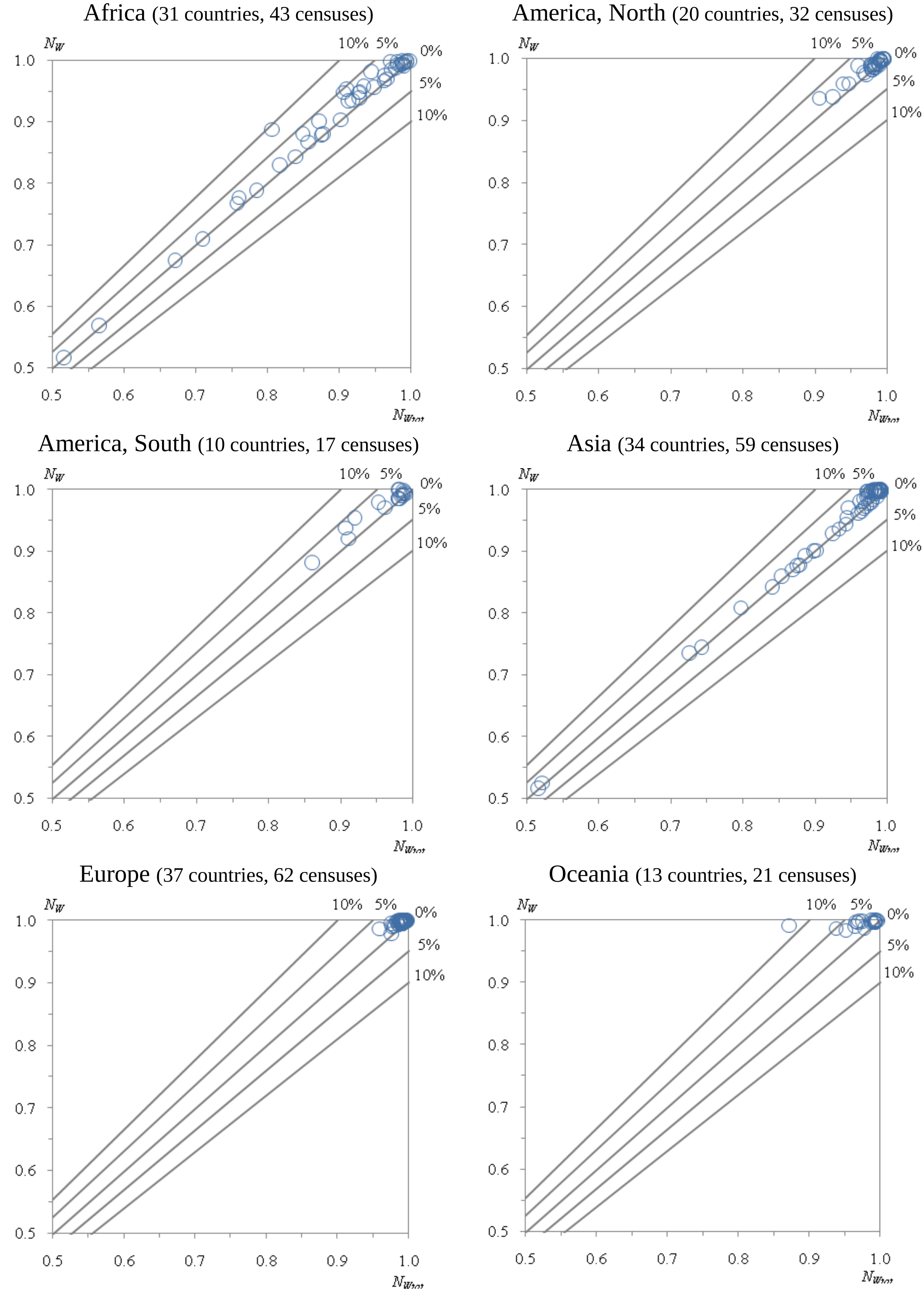
Roger G., Waltisperger D., Corbille-Guitton C. 1981. *Les structures par sexe et âge en Afrique*. Paris : Groupe de Démographie Africaine, IDP-INED-INSEE-MINCOOP-ORSTOM.

Spoorenberg, T. 2007. Quality of age reporting: Extension and application of the modified Whipple's index. *Population-E* 62 (4): 729-742.

United Nations. 2007. *2004 Demographic Yearbook*, *Fifty-sixth issue*. New York: Department of Economic and Social Affairs, ST/ESA/STAT/SER.R/35.

United Nations. n.d. *Demographic Yearbook* Special Census Topics, Volume 1 – Basic Population Characteristics, Table 1a: Population by single years of age, sex, urban/rural residence: each census, 1985-2003. <http://unstats.un.org/unsd/demographic/products/dyb/dybcens.htm>. New York. Accessed: 25 March 2008.

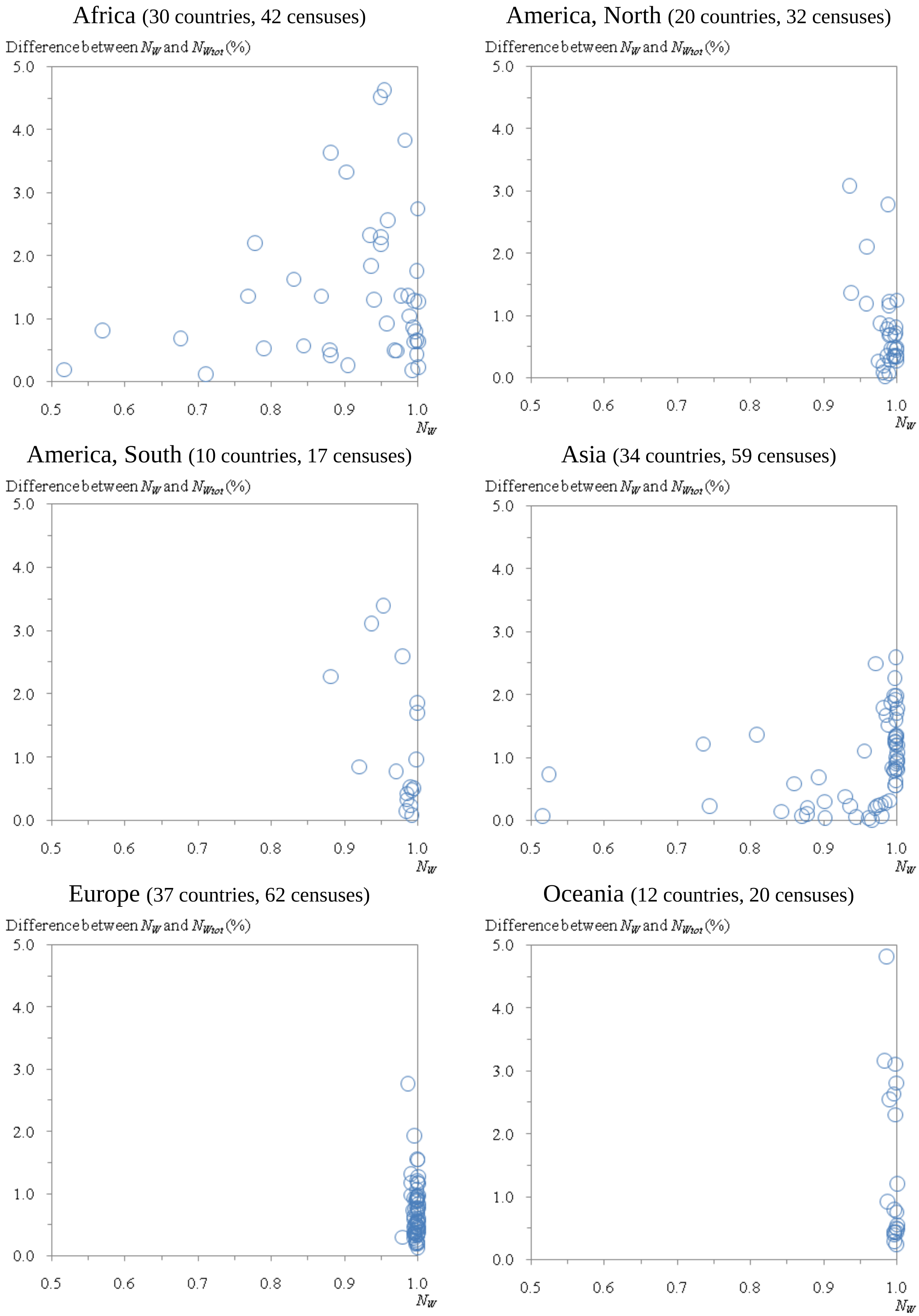
Figure 1: Normalized values of the total modified Whipple's index (N_{Wtot}) and the original Whipple's index (N_W) by world regions, 145 countries, 234 censuses, 1985-2003.



Source: United Nations n.d.

Note: the lines (and their labels) indicate the percentage difference between the two indices.

Figure 2: Normalized values of the original Whipple's index (N_W) and percentage difference between N_W and N_{Wtot} by world regions, 143 countries, 232 censuses, 1985-2003.



Source: United Nations n.d. *Note*: Countries with high percentage difference (Saint Helena:Tristan da Cunha and Pitcairn, respectively 9.233 and 11.997 percent) are not shown on figure 1.