

CHAPTER 2

EVALUATION OF COVERAGE AND CONTENT ERRORS

2.0. Introduction

Data evaluation is the assessment of the quality of the data. In evaluating the data, sometimes it is adjusted in order to ensure that it is of an acceptable standard. The adjustment is done on the basis of the responses to the following questions, which were asked during the census:

- Sex of members of household
- Age (in completed years) of members of household
- Residential status of household
- Children still living (with household or elsewhere), and
- Children dead

2.1. Concepts and Definitions

Listed below are the definitions of the major concepts used in this chapter.

Census of Population: Complete enumeration of persons during a specified period in a demarcated geographical area.

Child-Woman Ratio: Number of children age 0-4 years in a population to every 1000 women age 15-49 years in the same population.

Content Error: Mistake made in the recorded information in the census questionnaire either by the respondent or by the interviewer.

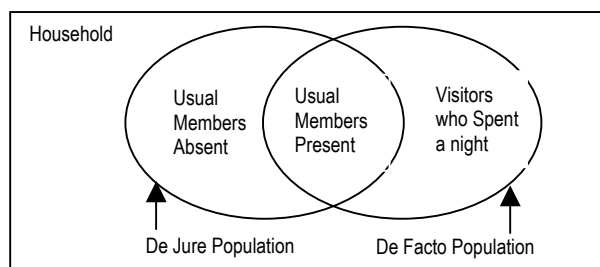
Coverage Error: Under or over-enumeration in a population census due to either omission or duplication.

De facto Population: This refers to the usual household members present and visitors who spent the census night at any given household. This however excludes:

- (a) Foreign diplomatic personnel accredited to Zambia; and
- (b) Zambian nationals accredited to foreign embassies and their family members who live with them abroad and , Zambian migrant workers and students in foreign countries who were not in the country at the time of the census..

De jure Population: This refers to usual household members *present* and usual household members temporarily *absent* at the time of the census. These include institutional populations in places such as hospitals/health centers, prisons and academic institutions (universities, colleges, boarding schools).

Thus, the de facto and the de jure population can be diagrammatically represented as follows:



Dependency Ratio: Ratio of children age 0-14 and persons age 65 years and older, per 100 persons in the age-group 15-64 years old.

Digit Preference: Reporting of age by respondents often ending in certain preferred digits. This results in heaping of population in ages ending with certain digits.

Evaluation of Census Data: Measurement of the quality of Census data.

Sex Ratio: Number of males per 100 females in a population.

2.2. Type of Population used in Evaluating the Coverage and Content Errors

In the analysis of the coverage and content errors, the de facto population has been used. This is so because we would like to analyse the information obtained from the people who gave us their details and not those we did not talk to or collect the information from.

2.3. Methods of Evaluation

During enumeration, checks and controls are instituted to minimise errors in the census. Despite instituting data control measures, there are usually several errors in the census data. For instance, some people may be completely omitted, others may be enumerated more than once, or some characteristics of an individual such as age, sex, fertility and economic activity of the canvassed individual may be incorrectly reported or tabulated. In general, two approaches are used to evaluate the quality of data, direct and indirect methods.

The direct method basically involves the carrying out of what is referred to as a Post Enumeration Survey (PES). In a PES, a sample of households is revisited after the census and data are again collected but on a smaller scale and later compared with that collected during the actual census. The matching process of the two sets of data can then be used to evaluate the quality of the census data. With regard to the 2000 Census of Population and Housing, the PES was carried out between February and March 2001. PES information is, however, only available for use at national level, and therefore, will not be used to evaluate data quality at the provincial level.

Indirect methods usually employ the comparison of data using both internal and external consistency checks. Internal consistency checks compare relationships of data within the same census data, whereas external consistency checks compare census data with data generated from other sources. For instance, one can compare data on education obtained during a census with administrative data maintained by the Ministry of Education.

2.3.1. Coverage Error

The coverage errors are made when omission or duplication resulting in under or over-enumeration occurs. Some factors, which are attributable to this, include errors arising from inaccessibility and co-operation with respondents. Difficulties in communication and lack of proper boundary descriptions also contribute to coverage errors. Coverage errors are usually highlighted by examining certain statistics such as growth rate, age composition, child-woman ratio and dependency ratio.

2.3.1.1. Age Composition

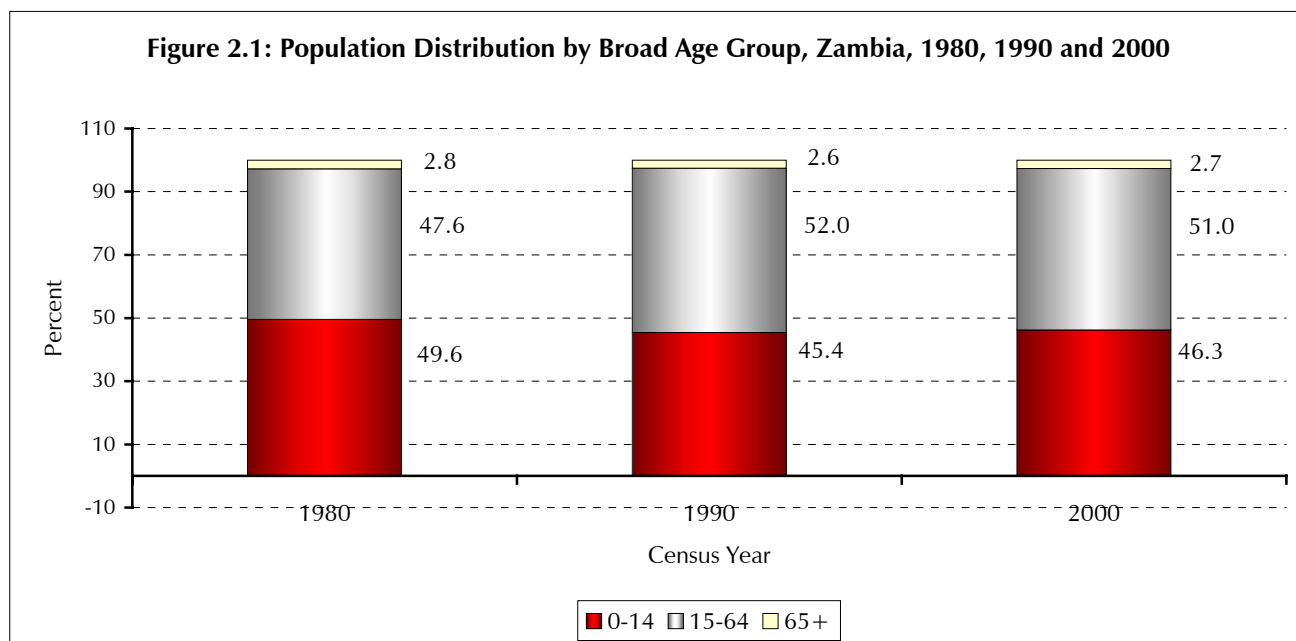
Looking at the age composition overtime can make you assess whether your census data have errors or not. Table 2.1 and Figure 2.1 show Zambia's population composition in broad age groups from 1980 to 2000.

Table 2.1: Population Distribution by Broad Age Groups, Zambia, 1980, 1990, and 2000

Age Group	Population					
	1980	Percent	1990	Percent	2000	Percent
0 - 14	2,809,268	49.6	3,349,677	45.4	4,323,448	46.3
15 - 64	2,693,225	47.6	3,842,799	52.0	4,758,697	51.0
65+	159,307	2.8	190,629	2.6	255,280	2.7
Total	5,661,801	100.0	7,383,105	100.0	9,337,425	100.0

Source: CSO, 1980, 1990 and 2000 Censuses of Population and Housing

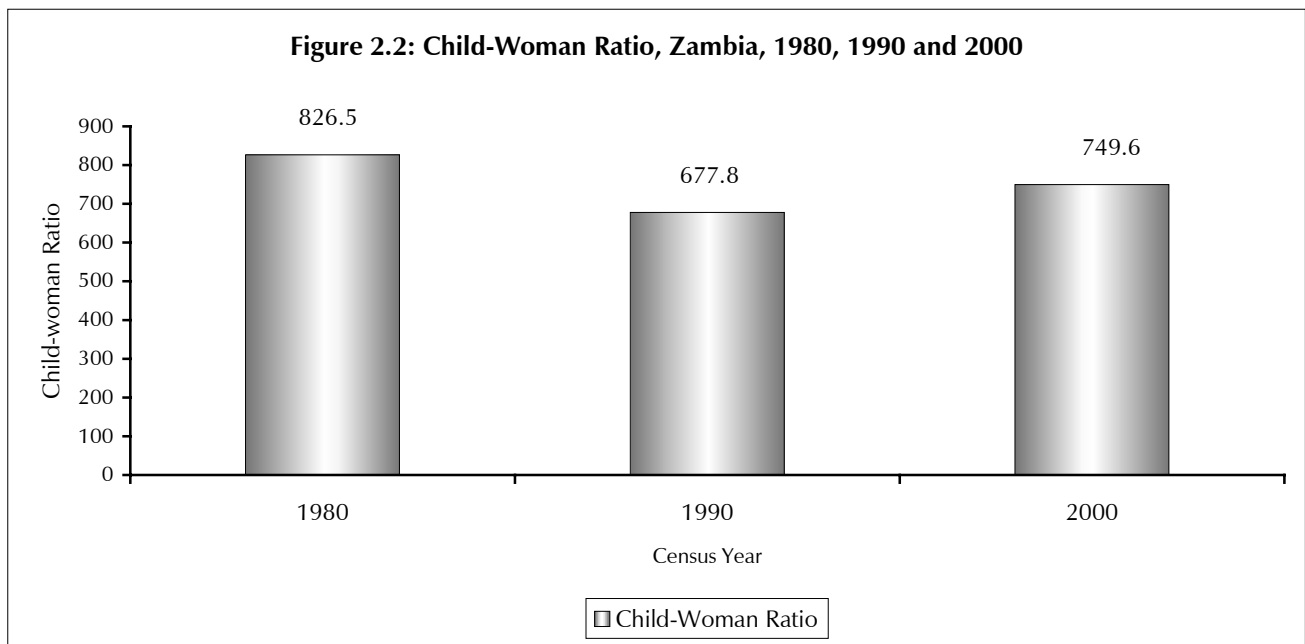
The proportion of children 0-14 years is less in 1990 than in 1980. In addition, the proportion of age 0-14 between 1990 and 2000 just made a marginal increase of one percent. These observations could be attributed to the decline in fertility and the general decrease in child mortality. Generally, the population distribution has remained stable.



Source: CSO, 1980, 1990 and 2000 Censuses of Population and Housing

2.3.1.2. Child-Woman Ratio

In 1980, the child-woman ratio was 826.5 per 1000 women age 15-49 years. It declined to 677.8 in 1990 and increased in 2000 to 749.6 per 1000 women age 15-49 years. This is in line with the changes in the proportion of the population in the 0-14 year age group. The decline in the proportion of the population 0-14 years and the decline in child-woman ratio between 1980 and 1990 (see Table 2.2 and Figure 2.2) could have been caused by a decline in fertility. Between the 1990 census and the 2000 census, there is a subsequent increase in the child-woman ratio and the proportion age 0-14.



Source: CSO, 1980, 1990 and 2000 Censuses of Population and Housing

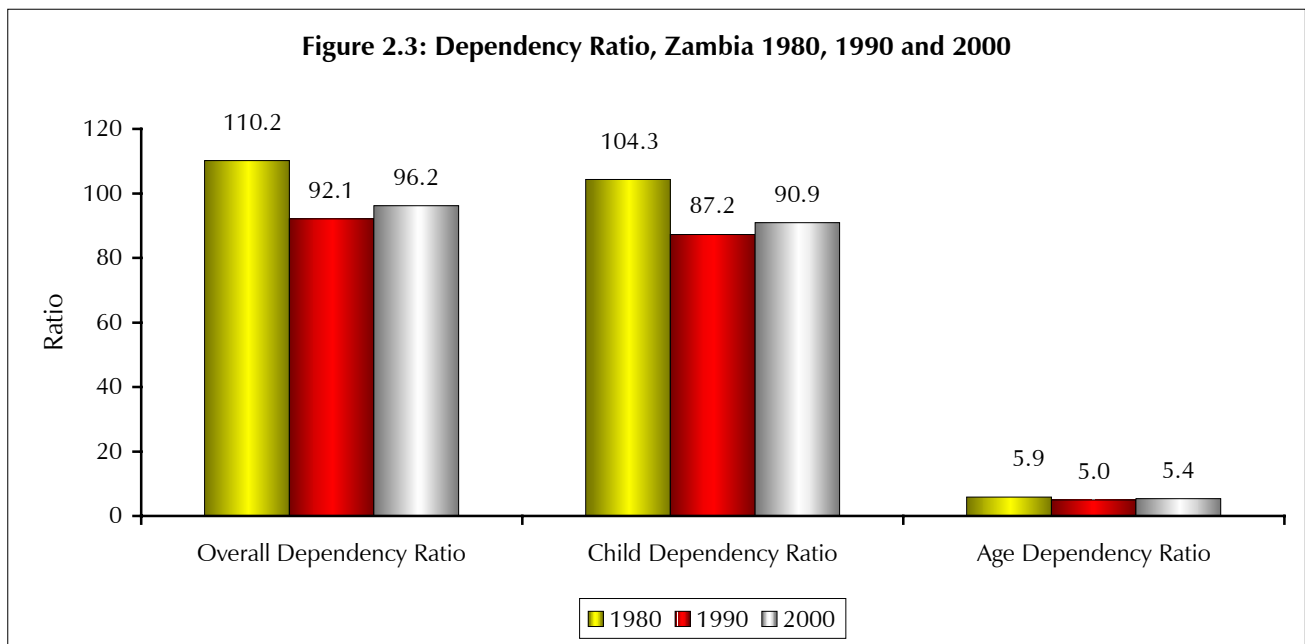
2.3.1.3. Dependency Ratio

The overall dependency ratio for the population of Zambia for 1980, 1990 and 2000 Censuses were 110.2, 92.1 and 96.2 persons, respectively, per 100 persons in age group 15-64 years. This means that for every 100 persons in the age range 15-64 years, there are 96.2 persons in the age groups 0-14 and 65 years or over. Child and Aged dependency ratio have followed a similar pattern. The decline in dependency ratios between 1980 and 1990 could be attributed to an increase in the proportion of population age 15-64 years, while its increase in 2000 could be attributed to a decrease in the proportion age 15-64 (See Table 2.2 and Figure 2.3).

Table 2.2: Dependency Ratio and Child-Woman Ratio, Zambia, 1980, 1990, and 2000

Ratio	1980	1990	2000
Overall Dependency Ratio	110.2	92.1	96.2
Child Dependency Ratio	104.3	87.2	90.9
Aged Dependency Ratio	5.9	5.0	5.4
Child-Woman Ratio	826.5	677.8	749.6

Source: CSO, 1980, 1990 and 2000 Censuses of Population and Housing



Source: CSO, 1980, 1990 and 2000 Censuses of Population and Housing

2.4. Content Errors

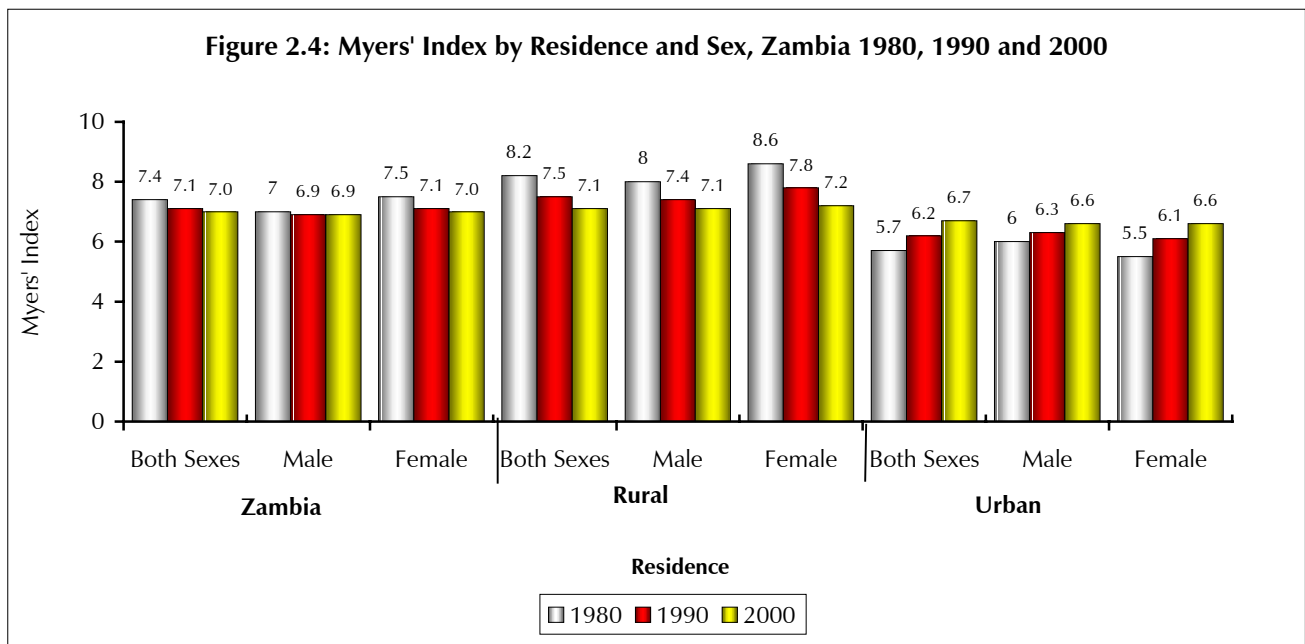
Content errors usually refer to instances where characteristics such as age, sex, marital status, economic activity, etc. of a person enumerated in a census or survey are incorrectly reported or tabulated. Content errors are caused by either a respondent giving a wrong response or by an enumerator recording an incorrect response. For instance, a question about age in a census can be solicited by asking either "date of birth" or "completed number of years". These two questions may yield different ages. During the 2000 Census, age was recorded in completed years. Some content errors can be estimated by the use of the Myers' Index, Sex Ratios, Age Ratios, and Survival Ratios.

2.4.1. Digit Preference

The tendency of respondents to report ages ending with certain digits in preference to other digits is called "digit preference". Digit preference is most pronounced among population subgroups having a low educational status. The causes and patterns of digit preference vary from one culture to another. Age misreporting, net under-enumeration and non-reporting or misclassifications of age contribute to heaping (Shryock, et.al. 1976).

Investigation of age heaping in Zambia is done through the calculation of Myers' Index. This index has been calculated for 1980, 1990 and 2000 Censuses data using PAS (Population Analysis Spreadsheet) a United Nations developed software package in Excel, and is presented in Table 2.3. A high Myers' Index implies poor age reporting whereas a low Myers' Index indicates good age reporting. The maximum value of Myers' Index is 90 and the minimum value is 0. In the case of Zambia, in all the three censuses, the index is on the lower side (less than 10), which implies that the age reporting is good.

Table 2.3 shows an overall marginal decrease in the index. This implies an improvement in the quality of age data. The table further shows a higher index for females than males in all the three censuses, indicating that age was more accurately reported among males than females. This trend is also observed in the rural part of Zambia. In the urban areas, the opposite was observed; there, the index has been rising with succeeding censuses, indicating deterioration in the quality of age data.



Source: CSO, 1980, 1990 and 2000 Censuses of Population and Housing

Table 2.3: Summary of Myers' Index for Digit Preference in Age Data by Residence and Sex, Zambia, 1980, 1990 and 2000

Residence	Sex	1980 Census	1990 Census	2000 Census
Zambia	Both Sexes	7.4	7.1	7.3
	Male	7.0	6.9	7.3
	Female	7.5	7.1	7.3
Rural	Both Sexes	8.2	7.5	7.5
	Male	8.0	7.4	7.4
	Female	8.6	7.8	7.6
Urban	Both Sexes	5.7	6.2	6.9
	Male	6.0	6.3	6.9
	Female	5.5	6.1	6.8

Source: CSO, 1980, 1990 and 2000 Censuses of Population and Housing

The urban areas further show a higher index among males than females, thus showing more accurate age reporting among females than males. Despite the deterioration in the quality of age data in the urban areas, its index is still lower than the rural one in all the three years, implying better age reporting in the former than in the latter.

Age heaping was further explored by assessing the most preferred digits presented in Table 2.4.

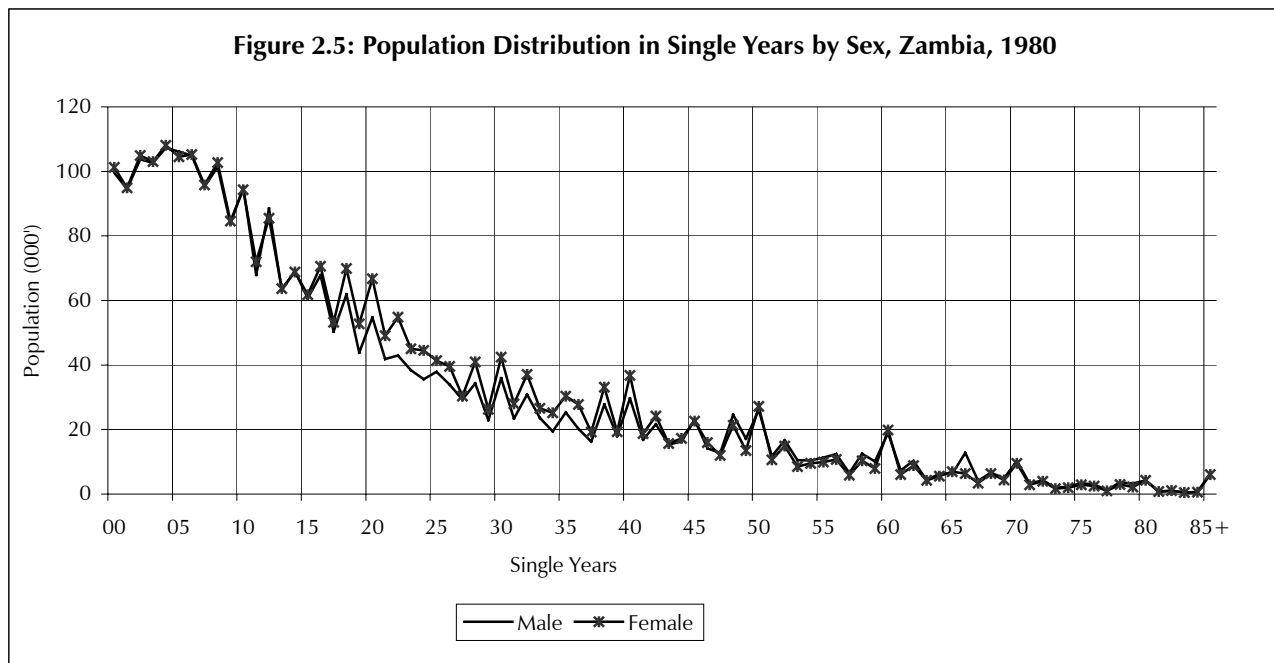
Table 2.4: Most Preferred Digits by Residence and Sex, Zambia, 1980, 1990, and 2000

Residence	Most Preferred Digits and Census Year			
	Sex	1980	1990	2000
Zambia	Both Sexes	0, 8, 2	0, 8, 2	0, 5, 8
	Male	0, 8, 2	0, 8, 2	0, 5, 8
	Female	0, 8, 2	0, 8, 2	0, 8
Rural	Both Sexes	0, 8, 2	0, 8, 2	0, 5, 8
	Male	0, 8, 2	0, 8, 2	0, 5, 8
	Female	0, 8, 2	0, 8, 2	0, 8
Urban	Both Sexes	0, 8, 2	0, 8, 2	0, 8
	Male	0, 8, 2	0, 8, 2	0, 5, 8
	Female	0, 8, 2	0, 8, 2	0, 2, 8

Source: CSO, 1980, 1990 and 2000 Censuses of Population and Housing

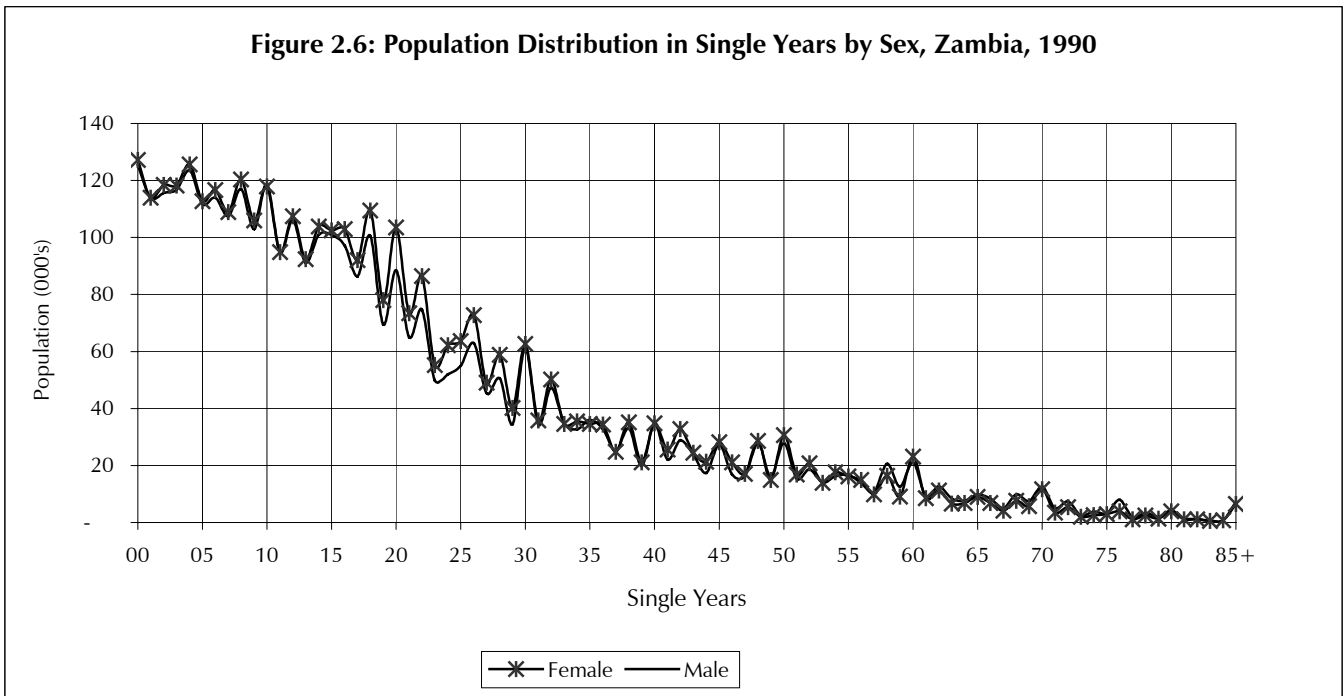
Table 2.4 shows that there was age heaping in all the three censuses. The table shows the most preferred digits in decreasing order of preference for the three censuses. Preference for digits 0, 8, and 2 among males and females is observed in both 1980 and 1990. In the 2000 Census, 0, 5 and 8 were the most preferred digits by males at national level and by residence as well. Females in rural areas preferred 0 and 8 while those in urban areas preferred 0, 2, and 8. The preference for these digits among males may be attributed to the greater tendency to over estimate the age, whilst for females; it may be due to underestimation of their age.

Age misreporting errors are also presented in Figures 2.5 to 2.10. The peaks on the curves indicate the most preferred ages in reporting while the troughs indicate the under reported ages. A comparison of Figures 2.5 to 2.7 shows that the peaks and troughs are higher for ages reported below 60 years in all the three censuses. There is no noticeable difference in the height of the peaks and troughs for ages reported after 60 years. This may suggest that both males and females tend to misreport their ages before age 60.



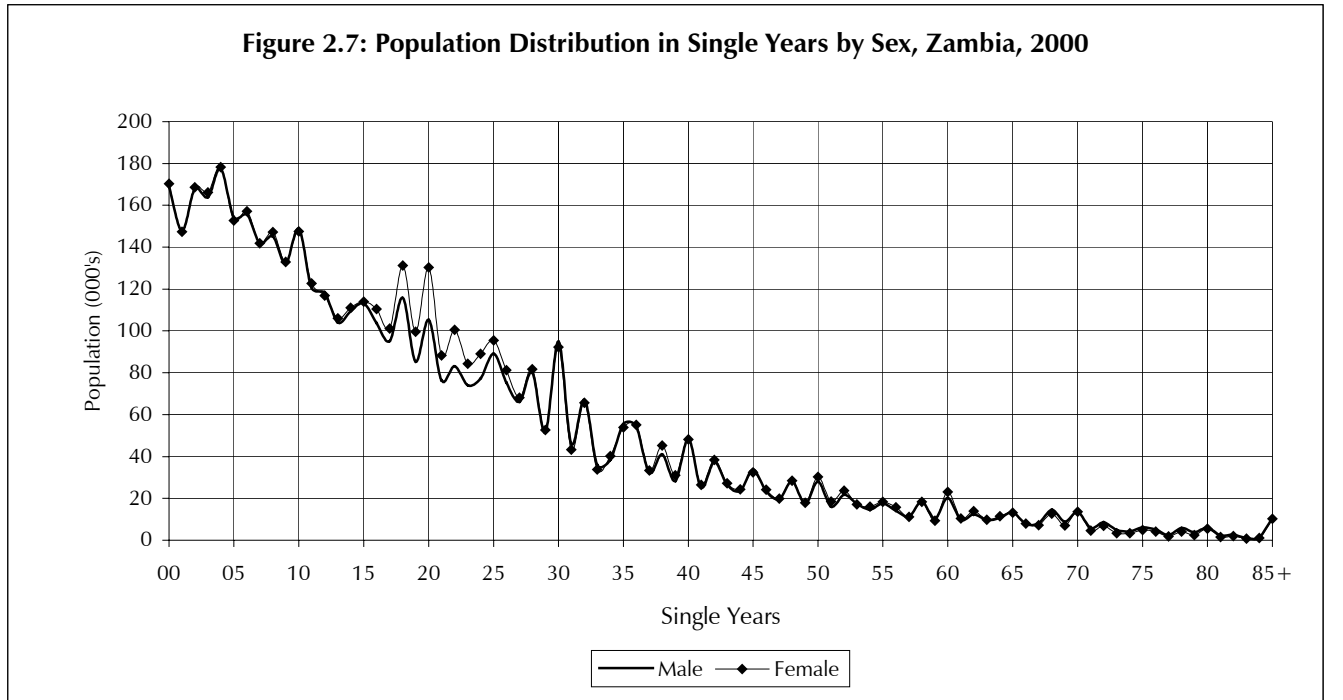
Source: CSO, 1980 Census of Population and Housing

Figure 2.6: Population Distribution in Single Years by Sex, Zambia, 1990



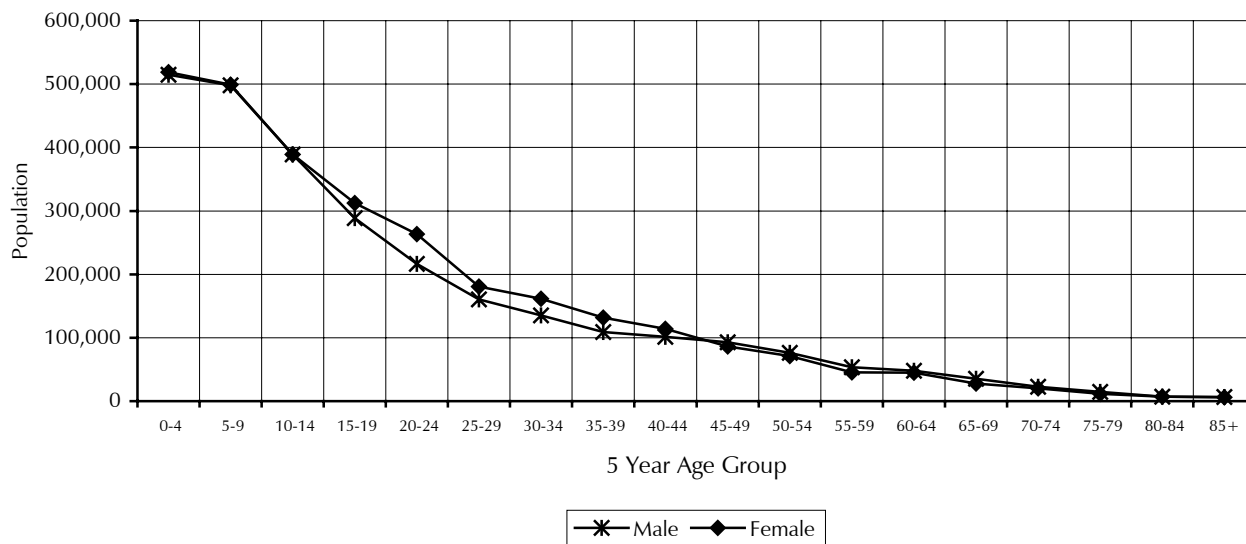
Source: CSO, 1990 Census of Population and Housing

Figure 2.7: Population Distribution in Single Years by Sex, Zambia, 2000



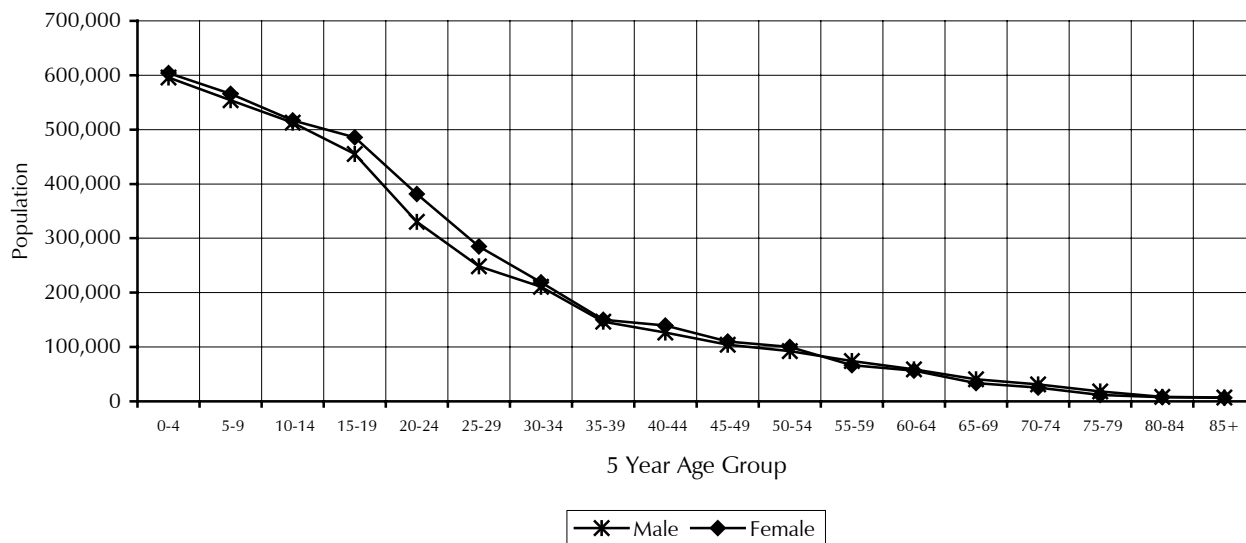
The smoothness of the curves in Figures 2.8 to 2.10 show that grouping single year age data into five-year age groups improves irregularities in age data arising from age misreporting.

Figure 2.8: Population Distribution by 5 Year Age Group and Sex, Zambia, 1980



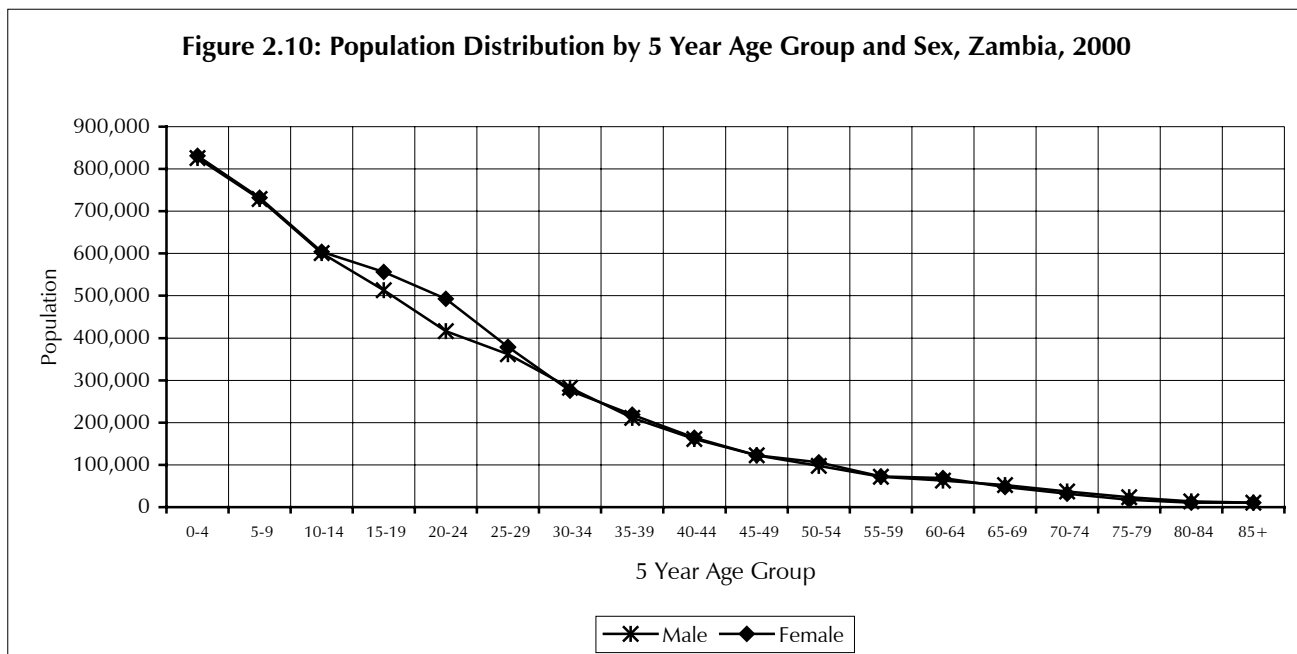
Source: CSO, 1980 Census of Population and Housing

Figure 2.9: Population Distribution by 5 Year Age Group and Sex, Zambia, 1990



Source: CSO, 1990 Census of Population and Housing

Figure 2.10: Population Distribution by 5 Year Age Group and Sex, Zambia, 2000

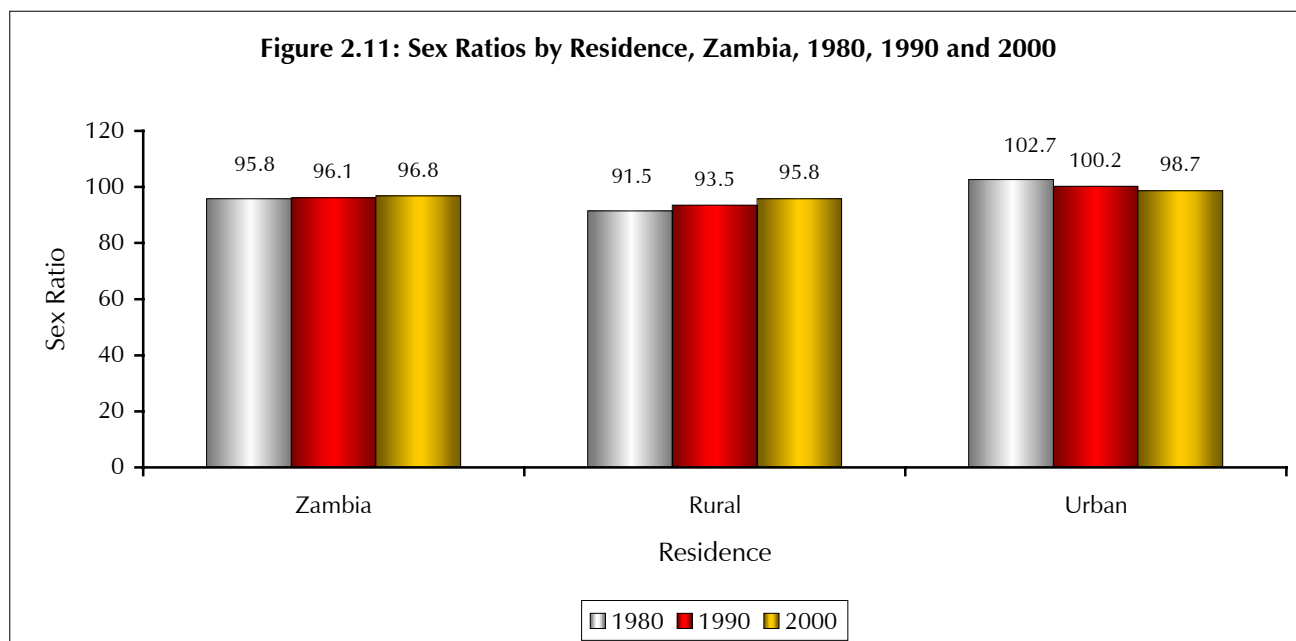


2.4.2. Sex Ratios

The presence of errors of omission, age misreporting, and out-migration may be detected by looking at the pattern of sex ratios. A sex ratio is the number of males per 100 females. A sex ratio of more than 100 shows an excess of males over females while a sex ratio of less than 100 shows an excess of females over males. A sex ratio of 100 indicates an equal number of males and females.

Overall, the sex ratios for Zambia are given in Figures 2.11 and 2.12 and Tables 2.5, 2.6, 2.7, and 2.8. Sex ratio in Zambia has increased slightly, from 95.8 in 1980 to 96.1 in 1990 and to 96.8 males per 100 females in 2000. In all the three censuses, the sex ratio is below 100 showing slightly more females than males in Zambia. In terms of residence, sex ratio is higher in urban areas than in rural areas for all the three censuses. This indicates a preponderance of males over females in the urban areas and the opposite in the rural areas. Further analysis also shows that the sex ratio for rural areas has been increasing whilst that for urban areas has been declining. The pattern of sex ratios cannot only be attributed to errors in the data but are also influenced by sex selective migration too. In this case, selective migration from rural areas to urban areas could have been in favour of males. This could have reduced the overall sex ratios in the rural areas and increased sex ratios in the urban areas. However, the urban-rural drift being observed in Zambia in the recent past could have reduced the overall sex ratios in the urban areas and increased sex ratios in the rural areas (Refer to the Migration and Urbanization Report, available at CSO).

Figure 2.11: Sex Ratios by Residence, Zambia, 1980, 1990 and 2000

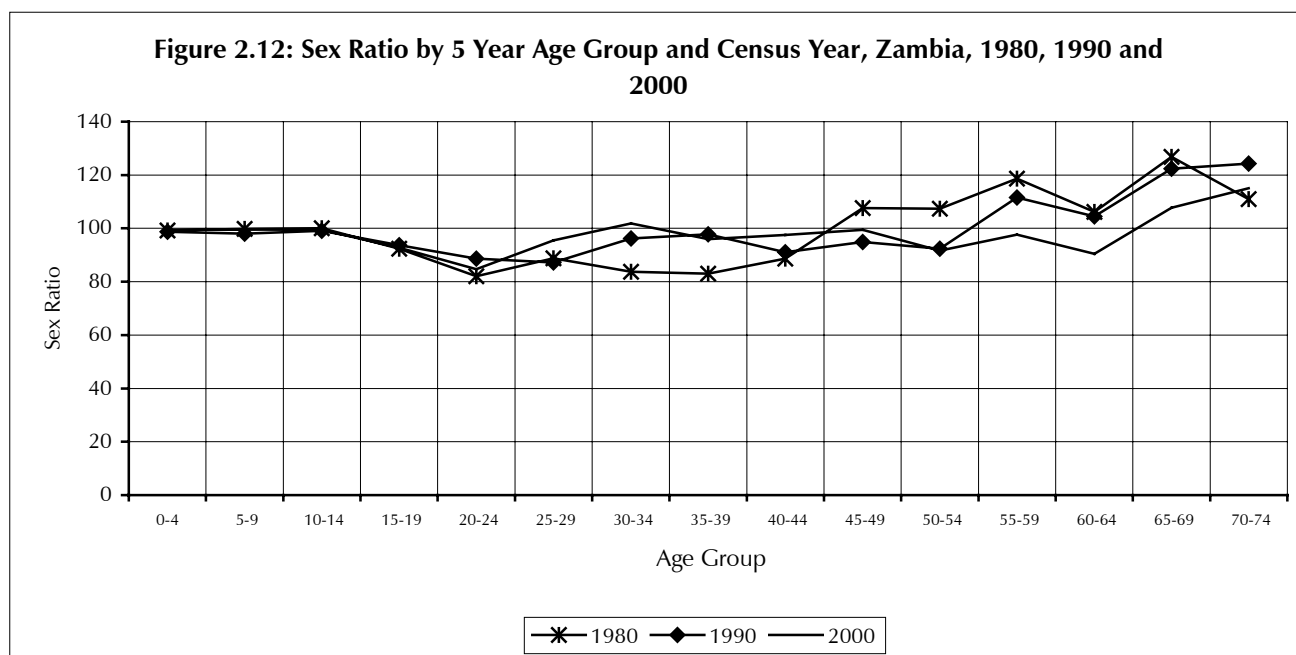


Source: CSO, 1980, 1990 and 2000 Censuses of Population and Housing

In the absence of significant fluctuations in births, deaths and migration, the sex ratios are expected to be high at infant ages because the sex ratio at birth is favourable to males. After early childhood, the ratios are expected to decline continuously to reach very low levels at the highest ages when female mortality is much lower than the male mortality.

An analysis of age-specific sex ratios for 1980 reveals a deficit of males in age groups 15-44 years whereas that for 1990 reveals a deficit of males in age groups 0-54 years. Ratios for 2000 show a deficit of males in the ages 0-4, 10-29, 35-44 and 50-64. (See Figure 2.12).

Figure 2.12: Sex Ratio by 5 Year Age Group and Census Year, Zambia, 1980, 1990 and 2000



Source: CSO, 1980, 1990 and 2000 Censuses of Population and Housing

There are many possible factors that may explain this, including high male mortality. The tendency also, by men to over estimate their age could have shifted men into older ages while the tendency by women to under-state their age could have shifted them into younger ages, hence, causing errors in age and sex data. A lower sex ratio in the age group 0-4 and 5-9 may suggest under enumeration of children, since sex ratio is supposed to be high at such age groups.

Table 2.5: Sex Ratio by Age and Residence, Zambia, 1980, 1990, and 2000

Age Group	1980			1990			2000		
	Zambia	Rural	Urban	Zambia	Rural	Urban	Zambia	Rural	Urban
0 - 4	99	99	100	99	99	99	99	99	99
5 - 9	100	101	98	98	99	96	100	101	97
10-14	100	104	94	99	104	93	99	103	93
15-19	92	93	92	94	96	90	92	94	89
20-24	82	80	85	87	84	90	84	83	86
25-29	89	78	102	87	84	92	95	92	101
30-34	84	65	112	96	90	105	103	96	113
35-39	83	65	115	98	82	119	97	90	108
40-44	89	69	132	91	68	138	98	91	110
45-49	108	84	177	95	73	153	100	89	121
50-54	107	86	189	92	75	153	93	78	134
55-59	119	103	185	112	99	160	99	87	138
60-64	106	100	140	105	98	139	91	87	108
65-69	127	122	154	122	118	143	109	109	110
70-74	111	111	111	124	123	133	116	118	108
75-79	126	127	119	155	155	152	134	140	111
80-84	100	100	101	110	113	94	122	127	100
85+	106	105	117	111	114	93	104	109	78

Source: CSO, 1980, 1990 and 2000 Censuses of Population and Housing

Sex ratios higher than 100 were observed in age groups above 45 years in 1980, age groups above 55 years in 1990 and age group above 65 years in 2000. The 2000 pattern of sex ratio seems to be consistent with the 1980 and 1990 pattern because, the cohorts, which had lower sex ratios 10 years before, still have lower sex ratio 10 years later. However, the pattern of sex ratio of 1980, 1990 and 2000 suggest under enumeration of children since sex ratio is supposed to be high at age groups 0-4 and 5-9.

2.4.3. Age Ratio

The quality of age data can also be evaluated by examining age ratios. An age ratio may be defined as the ratio of the population in a given age group to one-third of the sum of the populations in the age group itself, the preceding and the following age groups, times 100 (Shryock et al, 1976). When there are no major changes in fertility, mortality or migration, the age ratios do not deviate much from 100, hence, any substantial deviation is explained in terms of age misreporting. Age ratios disaggregated by sex are given in Tables 2.6, 2.7, and 2.8 and Figure 2.13.

Figure 2.13: Population Distribution of Zambia by Age Ratio and Sex, 2000

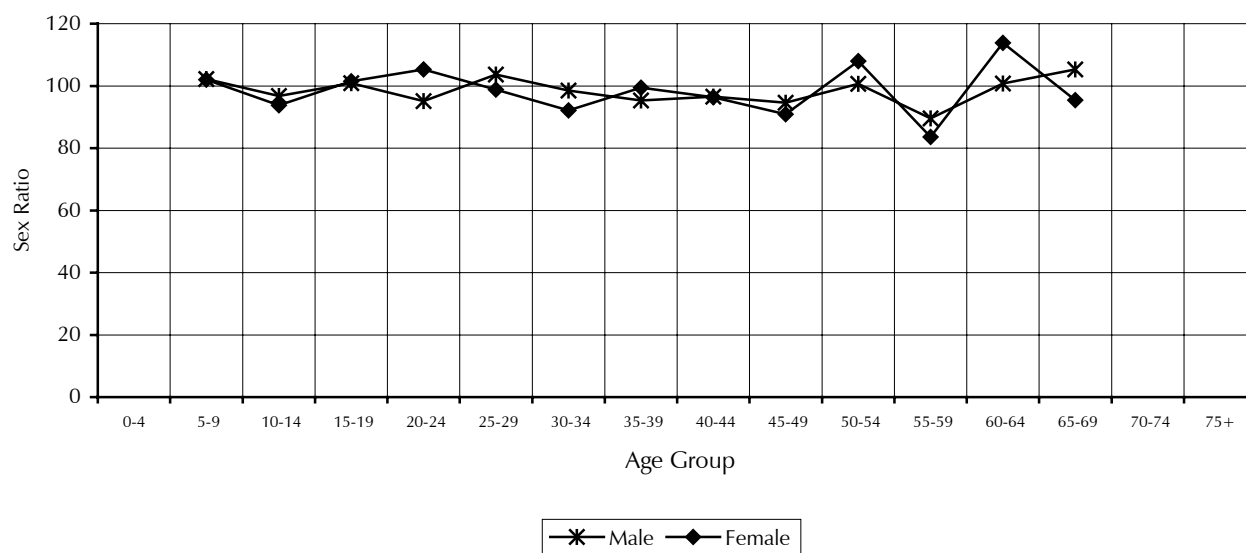


Table 2.6: Population by Five Year Age Group, Sex, Age Ratio and the Age-Sex Accuracy Index, Zambia, 1980

Age Group	Population		Age Ratio		Deviation from 100		Sex Ratio	Difference
	Male	Female	Male	Female	Male	Female		
0 - 4	514,698	518,793	-	-	-	-	99.2	-
5 - 9	498,074	499,191	110.2	109.9	10.2	9.9	99.8	0.6
10-14	389,247	389,288	99.0	96.0	-1.0	-4.0	100.0	0.2
15-19	288,545	312,240	95.3	95.7	-4.7	-4.3	92.4	-7.6
20-24	216,454	263,573	96.4	106.9	-3.6	6.9	82.1	-10.3
25-29	160,543	180,829	91.4	85.1	-8.6	-14.9	88.8	6.7
30-34	134,993	161,250	100.1	103.3	0.1	3.3	83.7	-5.1
35-39	109,077	131,496	92.4	95.5	-7.6	-4.5	83.0	-0.8
40-44	101,217	114,120	100.2	104.7	0.2	4.7	88.7	5.7
45-49	92,993	86,438	104.7	93.3	4.7	-6.7	107.6	18.9
50-54	76,335	71,099	104.1	108.0	4.1	8.0	107.4	-0.2
55-59	53,678	45,246	86.4	77.9	-13.6	-22.1	118.6	11.3
60-64	47,935	45,134	108.0	123.7	8.0	23.7	106.2	-12.4
65-69	35,126	27,713	99.8	84.8	-0.2	-15.2	126.8	20.5
70-74	22,461	20,235	-	-	-	-	111.0	-15.8
75+	28,619	25,161	-	-	-	-	113.7	-
Total Popn	2,769,995	2,891,806	-	-	-	-	-	-
Mean	-	-	-	-	5.1	9.9	-	8.3

Source: CSO, 1980 Census of Population and Housing

Age-Sex Accuracy Index = 3 times mean difference in sex ratios plus mean deviations of male and female age ratios.
 $= 3 \times 8.3 + 5.1 + 9.9$
 $= 39.9$

Table 2.7: Population by Five Year Age Group, Sex, Age Ratio and the Age-Sex Accuracy Index, Zambia, 1990

Age Group	Population		Age Ratio		Deviation from 100		Sex Ratio	Difference
	Male	Female	Male	Female	Male	Female		
0 - 4	596,079	604,265	-	-	-	-	98.65	-
5 - 9	554,045	565,479	99.94	100.85	-0.06	0.85	97.98	-0.67
10-14	512,633	517,171	101.60	98.39	1.60	-1.61	99.12	1.14
15-19	455,045	485,773	107.95	108.12	7.95	8.12	93.67	-5.45
20-24	330,433	381,404	93.92	98.96	-6.08	-1.04	86.64	-7.04
25-29	248,568	285,021	91.85	94.91	-8.15	-5.09	87.21	0.57
30-34	210,810	219,189	106.62	100.72	6.62	0.72	96.18	8.97
35-39	146,862	150,238	87.03	83.83	-12.97	-16.17	97.75	1.58
40-44	126,705	139,258	100.83	107.01	0.83	7.01	90.99	-6.77
45-49	104,452	110,032	95.38	91.99	-4.62	-8.01	94.93	3.94
50-54	92,312	99,962	103.23	113.18	3.23	13.18	92.35	-2.58
55-59	74,390	66,612	98.23	85.10	-1.77	-14.90	111.68	19.33
60-64	59,145	56,582	102.75	113.29	2.75	13.29	104.53	-7.15
65-69	40,737	33,274	90.02	81.34	-9.98	-18.66	122.43	17.90
70-74	31,361	25,234	-	-	-	-	124.28	1.85
75+	34,000	26,028	-	-	-	-	130.63	-
Total Popn	3,617,577	3,765,520	-	-	-	-	-	-
Mean	-	-	-	-	5.1	8.4	-	6.07

Source: CSO, 1990 Census of Population and Housing

Age-Sex Accuracy Index = 3 times mean difference in sex ratios plus mean deviations of male and female age ratios.
 $= 3 \times 6.07 + 5.1 + 8.4$
 $= 31.7$

Table 2.8: Population by Five Year Age Group, Sex, Age Ratio and the Age-Sex Accuracy Index, Zambia, 2000

Age Group	Population		Age Ratio		Deviation from 100		Sex Ratio	Difference
	Male	Female	Male	Female	Male	Female		
0 - 4	825789	830931	-	-	-	-	99.4	-
5 - 9	729181	731901	102.2	102.0	2.2	2.0	99.6	0.2
10-14	601279	604367	96.8	93.8	-3.2	-6.2	99.5	-0.1
15-19	513320	556676	100.9	101.5	0.9	1.5	92.2	-7.3
20-24	416083	492589	95.1	105.3	-4.9	5.3	84.5	-7.7
25-29	361901	379247	103.6	98.8	3.6	-1.2	95.4	11.0
30-34	282439	275434	98.5	92.1	-1.5	-7.9	102.5	7.1
35-39	211356	218631	95.3	99.4	-4.7	-0.6	96.7	-5.9
40-44	161179	164597	96.6	96.4	-3.4	-3.6	97.9	1.3
45-49	122486	122834	94.6	90.9	-5.4	-9.1	99.7	1.8
50-54	97850	105762	100.7	108.0	0.7	8.0	92.5	-7.2
55-59	71905	72933	89.6	83.6	-10.4	-16.4	98.6	6.1
60-64	62678	68797	100.8	113.8	0.8	13.8	91.1	-7.5
65-69	52499	47994	105.3	95.4	5.3	-4.6	109.4	18.3
70-74	37066	31869	-	-	-	-	116.3	6.9
75+	47279	38573	-	-	-	-	122.6	-
Total	4,591,731	4,745,678	-	-	-	-	-	-
Mean	-	-	-	-	3.62	6.18	-	6.31

$$\begin{aligned}
 \text{Age-Sex Accuracy Index} &= 3 \text{ times mean difference in sex ratios plus mean deviations of male} \\
 &\quad \text{and female age ratios.} \\
 &= 3 \times 6.31 + 3.62 + 6.18 \\
 &= 28.7
 \end{aligned}$$

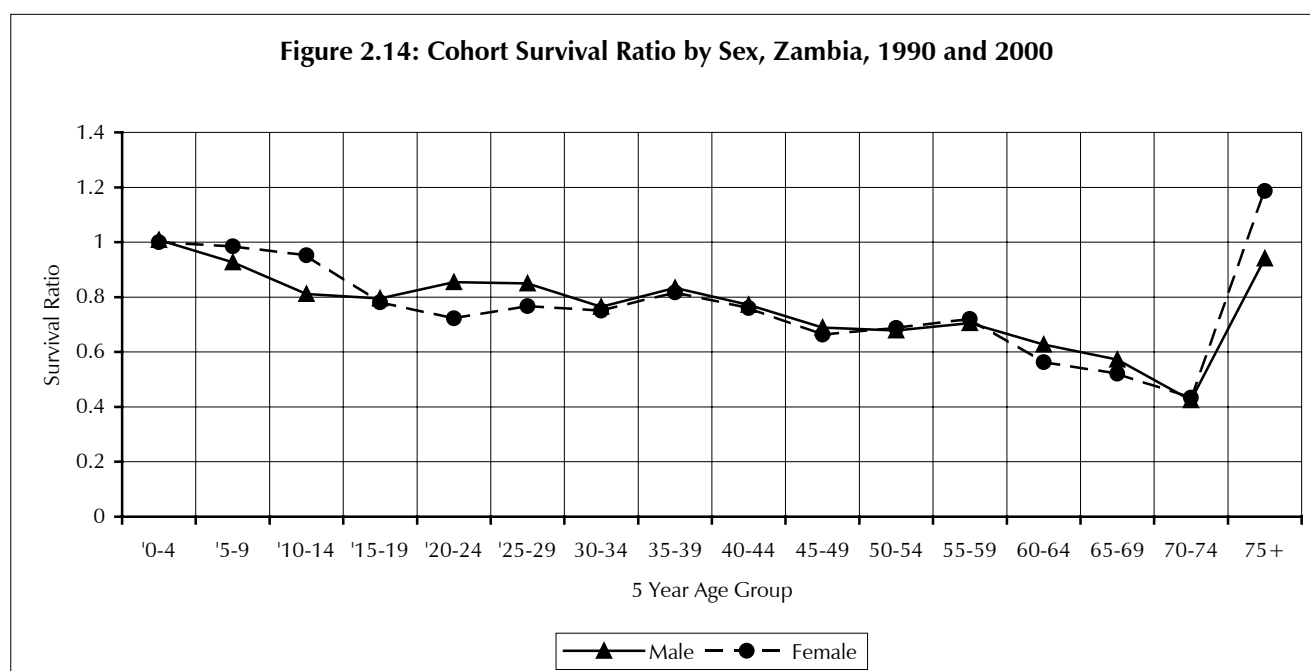
The irregular patterns of the age ratios show that data could be affected by errors from age misreporting, digit preference, omission, migration or fluctuations in births and deaths. In all the three Censuses females have higher Average Age Ratio Deviation than males. Average Age Ratio Deviation was 9.9 for females and 5.1 for males in 1980, 8.4 for females and 5.1 for males in 1990 and 6.2 for females and 3.6 for males in 2000. This shows that age reporting is better for males than for females.

It is worthy to note that over time, there has been an improvement in the quality of data depicted by the declining Age Accuracy Index. The Age Accuracy Index declined from 39.9 in 1980 to 31.7 in 1990 and to 28.7 in 2000. The United Nations define age data as "accurate, inaccurate and highly inaccurate" if the age accuracy index lies below 20, between 20-40, and 40 and above, respectively. In as far as the United Nations Age-Sex Accuracy Index is concerned, the 1980, 1990 and 2000 age data were inaccurate. However, the 2000 age data show some improvement over the 1980 and 1990 age data (Refer to Tables 2.6, 2.7, 2.8 and Figure 2.13 for details).

2.4.4. Survival Ratios

Survival ratios represent the probability that individuals of the same birth cohort or group of cohorts will still be living 10 years later. Survival ratios can also be used to evaluate the quality of data. Evaluating the quality of age and sex data from two censuses is done under the assumptions that the population should be closed to migration and influence of abnormal mortality such as through wars, disasters, and diseases over a 10-year period. Cohort survival ratio refers to the survival ratio of the population in a given age group to the next age whereas overall survival ratio refers to the ratio of the population of age say 10 years and above, who will survive to 15 years and above, and so on. With respect to Zambia, one needs not to rule out the effect of high mortality caused by the Acquired Immune Deficiency Syndrome (AIDS) on the pattern of survival ratios.

Cohort survival ratios are expected to be highest at age group 10-14 where mortality is assumed to be lowest and then to decline continuously thereafter. Table 2.9 shows fluctuations rather than the expected pattern i.e. there is no systematic decline in the cohort survival ratios with the increase in age. These distortions in data could either be due to age misreporting, under enumeration or over enumeration at some age groups.



Source: CSO, 1990 and 2000 Censuses of Population and Housing

Female ratios should be higher than the male ratios because females normally have lower mortality compared to males. However, Figure 2.14 and Table 2.9 indicate more male survivals in the 15-34 and 60-69 age groups.

Table 2.9: Cohort Survival Ratios By Sex, Zambia, 1990 and 2000

Age Group	Survival Ratios	
	Male	Female
0 - 4	1.0087	1.0002
5 - 9	0.9265	0.9844
10-14	0.8117	0.9525
15-19	0.7953	0.7807
20-24	0.8548	0.7222
25-29	0.8503	0.7671
30-34	0.7646	0.7509
35-39	0.8340	0.8176
40-44	0.7723	0.7595
45-49	0.6884	0.6628
50-54	0.6790	0.6882
55-59	0.7057	0.7205
60-64	0.6267	0.5632
65-69	0.5720	0.5214
70-74	0.4244	0.4332
75+	0.9412	1.1865

Source: CSO, 1990 and 2000 Censuses of Population and Housing

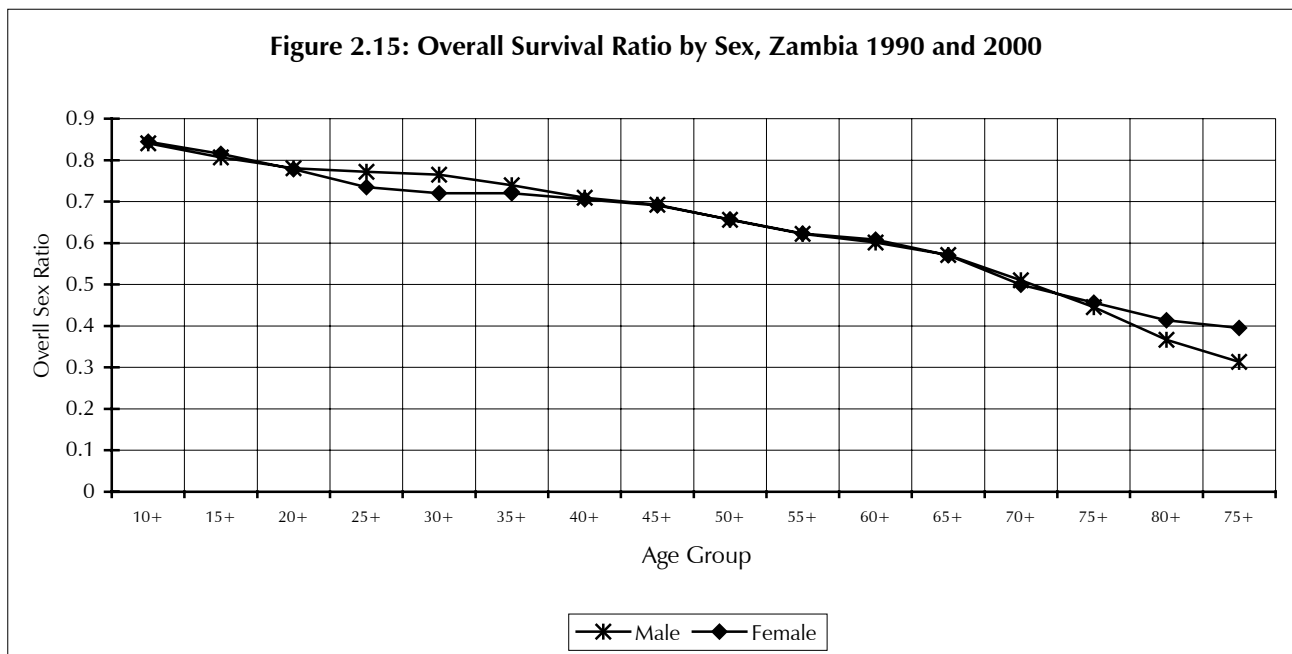
Overall survival ratios show the expected pattern, where the ratio declines continuously with increasing age, without any fluctuations. This suggests that the quality of age data is good. However, the pattern of having higher ratios for females than males is not true at 10+, 15+, 50+, 55+, 60+, 75+, 80+ and 85+ (Table 2.10 and Figure 2.15). This could be caused by age misreporting, under or over coverage in these age groups or high mortality.

Table 2.10: Overall Survival Ratios by Sex, Zambia, 1990 and 2000

Age	Survival Ratios	
	Male	Female
10+	0.8402	0.8446
15+	0.8069	0.8148
20+	0.7800	0.7779
25+	0.7718	0.7345
30+	0.7646	0.7204
35+	0.7391	0.7198
40+	0.7091	0.7053
45+	0.6927	0.6911
50+	0.6558	0.6567
55+	0.6220	0.6228
60+	0.6011	0.6085
65+	0.5711	0.5701
70+	0.5104	0.4992
75+	0.4456	0.4563
80+	0.3669	0.4141
85+	0.3137	0.3955

Source: CSO, 1990 and 2000 Censuses of Population and Housing

Figure 2.15: Overall Survival Ratio by Sex, Zambia 1990 and 2000

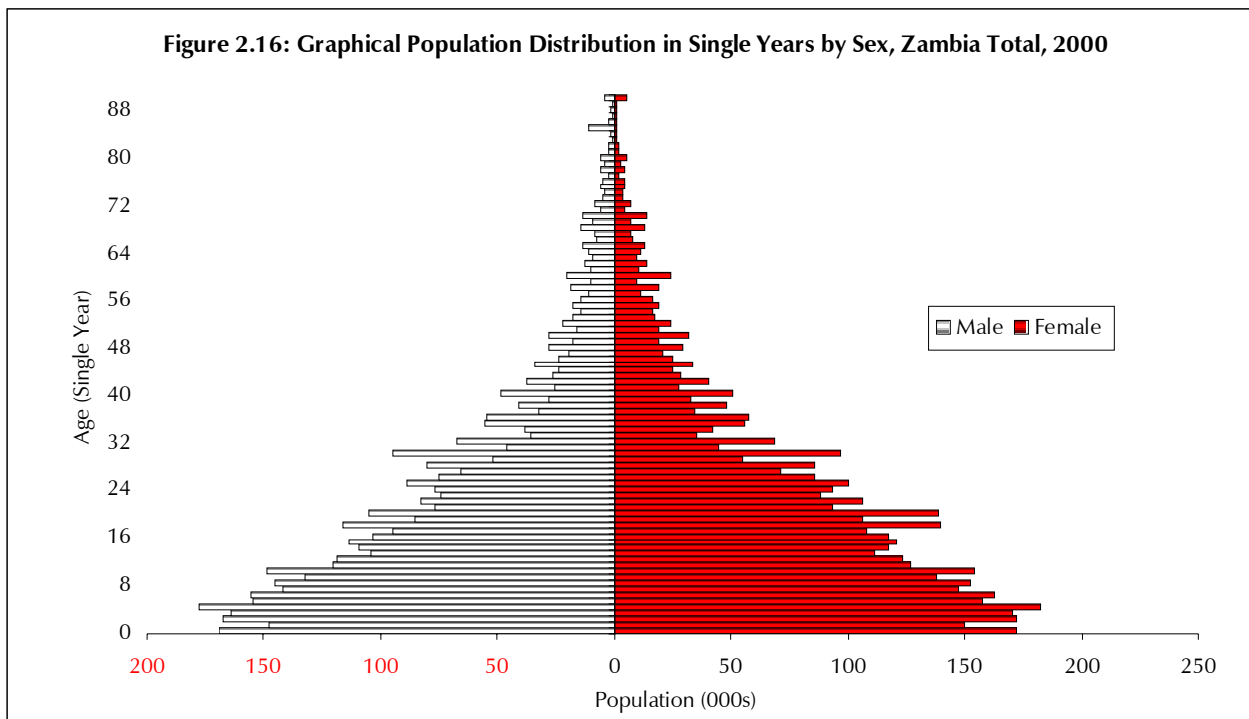


Source: CSO, 1990 and 2000 Censuses of Population and Housing

2.4.5. Population Pyramids

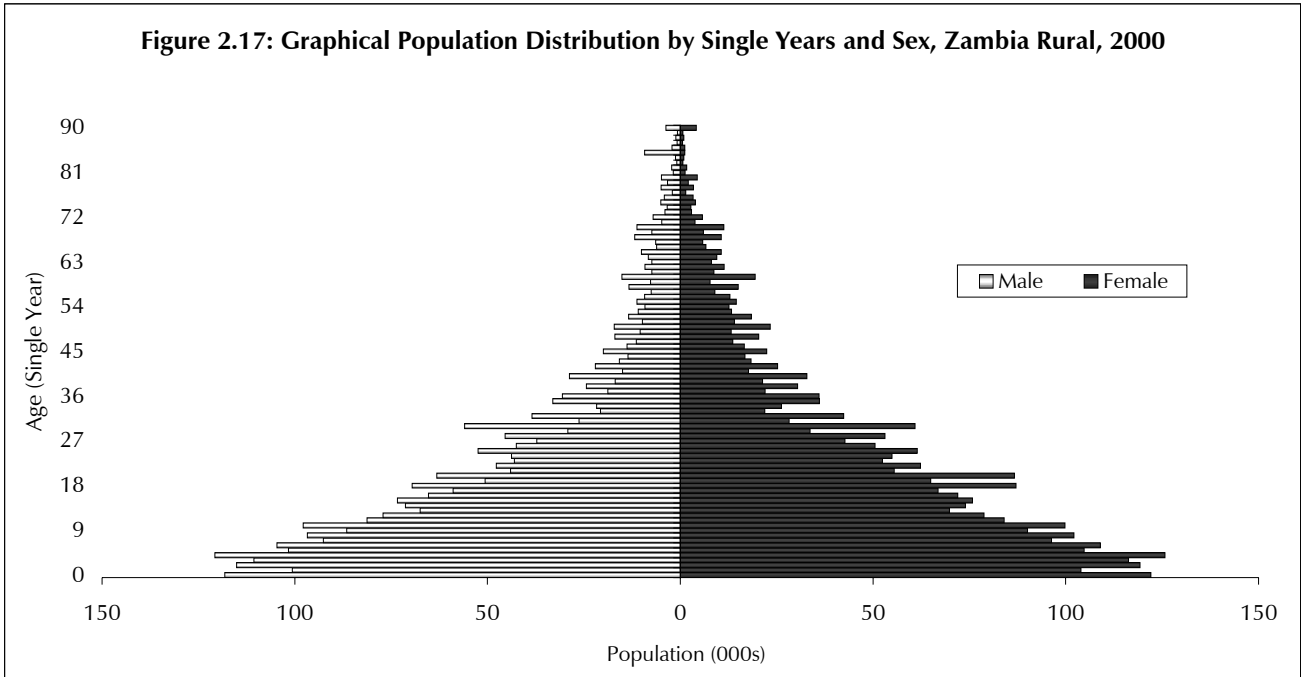
Another way of detecting irregularities in the reported age data of a survey or census is by looking at a Population Pyramid by single years of age. As already observed, when census age data is distributed in single years, one can easily spot out inaccuracies than when it is distributed in five-year age groups. Looking at the population pyramids for the 2000 Census data from Figures 2.16 to 2.18, it can be seen that age misreporting was not severe.

Figure 2.16: Graphical Population Distribution in Single Years by Sex, Zambia Total, 2000



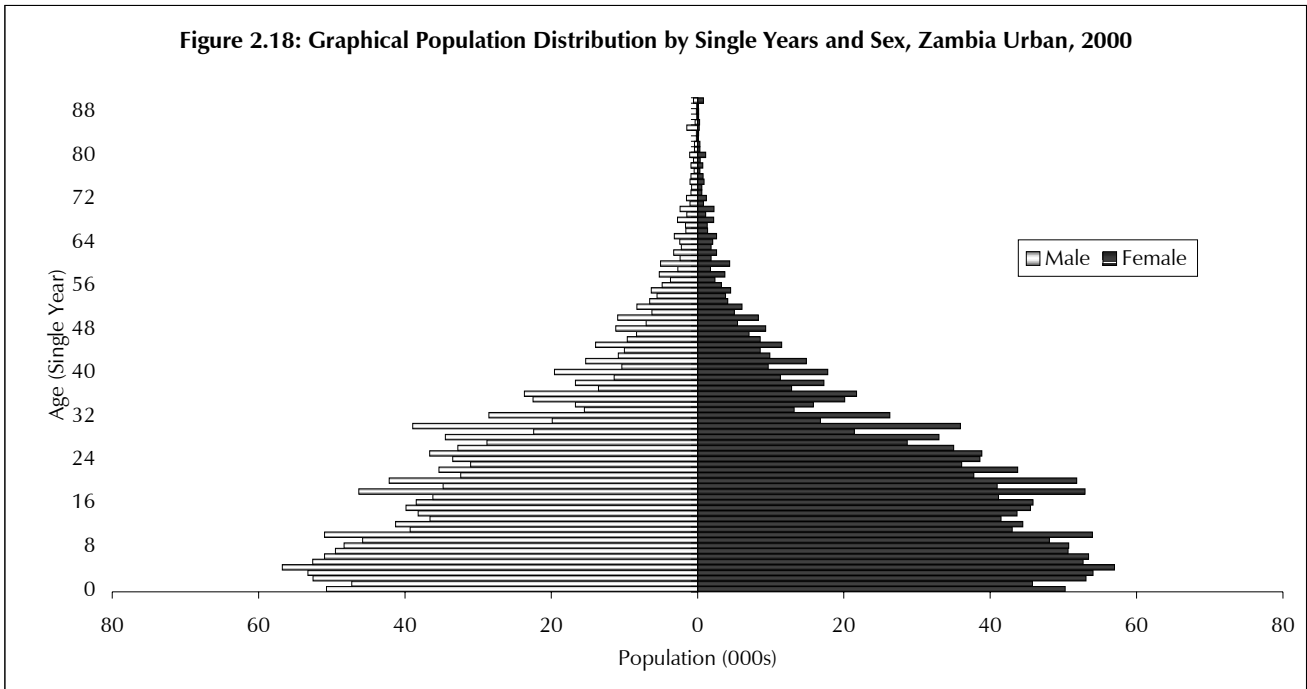
Source: CSO, 2000 Census of Population and Housing

Figure 2.17: Graphical Population Distribution by Single Years and Sex, Zambia Rural, 2000



Source: CSO, 2000 Census of Population and Housing

Figure 2.18: Graphical Population Distribution by Single Years and Sex, Zambia Urban, 2000



Source: CSO, 2000 Census of Population and Housing

Smoothing the age data using selected techniques for light smoothing of the population (Edwardo E. Arriaga: November 1994, pages 11-42) shows that the irregularities in the structure are not severe to warrant the smoothing of data, see Figures 2.19 and 2.20.

The smoothing of data has been done using AGESMTH software program one of the Population Analysis Spreadsheet (PAS) programmes developed by the United Nations. Selected techniques for light smoothing of the population include Carrier Farrag, Karup-King Newton, Arriaga and United Nations. The strong smoothing technique has also been incorporated.

Figure 2.19: Reported and Smoothed Population by 5 Year Age Group and Smoothing Technique, Zambia Males, 2000

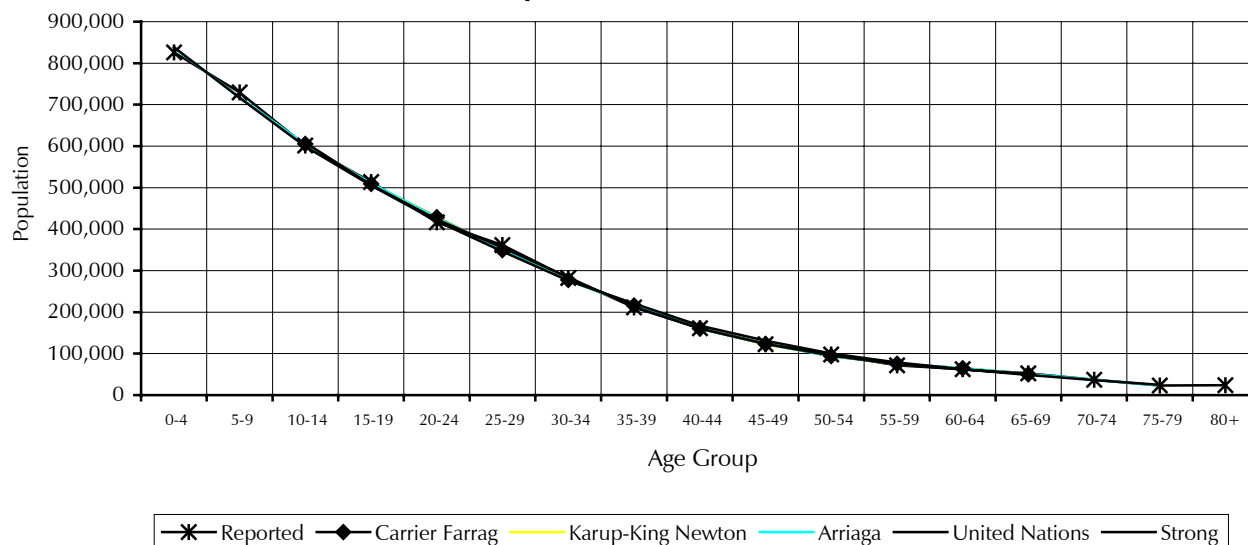
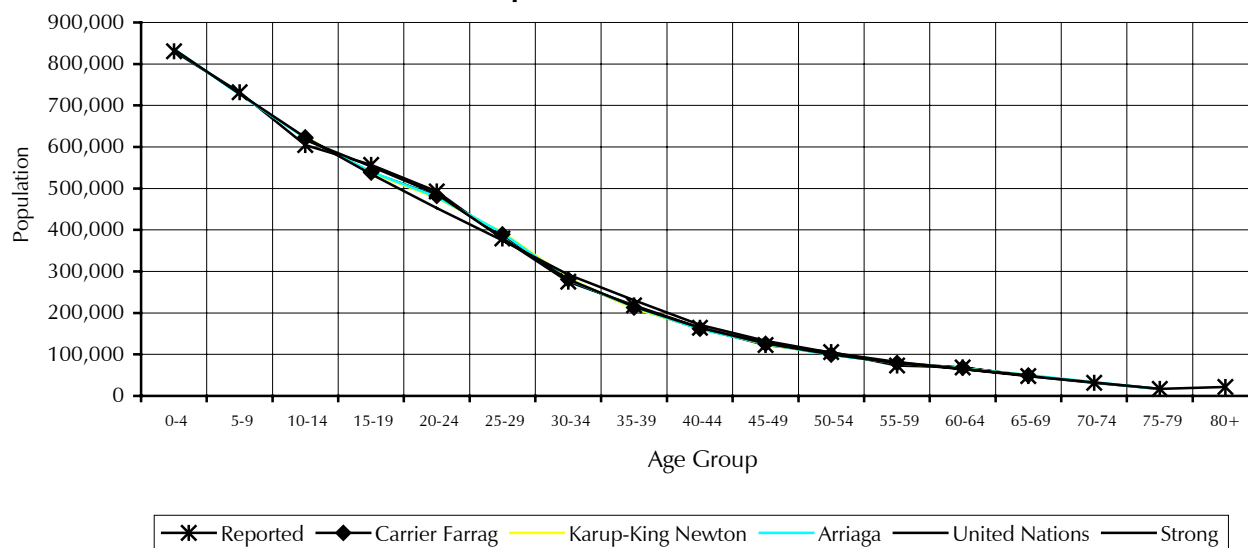


Figure 2.20: Reported and Smoothed Population by 5 Year Age Group and Smoothing Technique, Zambia Females, 2000



Given that the irregularities in the reported proportions are small as can be seen by the smoothness of the curves in Figures 2.19 and 2.20, it is not recommended to smooth the 2000 Census of Population and Housing data because genuine irregularities in the reported pattern might be smoothed out.

2.5. Summary

In the evaluation of content and coverage errors, the notable observations made were that the pattern of Age Composition, Child Woman Ratio and Dependency Ratio in 2000 is in line with the observed fertility and mortality declines. There was also digit preference during age reporting, the preferred digits being 0, 5, 8 and 2. Age heaping in the 2000 Census was also observed, just like in the other two previous censuses. However, the 2000 age-sex data shows an improvement over the 1990 and 1980 age-sex data as evidenced by the decline in the Age-Sex Accuracy Index from 39.9 in 1980, 31.7 in 1990 to 28.7 in 2000. In the analysis of Age-Specific Sex Ratios, under enumeration and over enumeration was identified at ages 0-4 and 5-9 year age groups. Generally, the census taking in Zambia is yielding improved data overtime.