

ART. VI.—*Note on the Veracity of the Returns of Age in the Census of 1901.*

By H. W. SEGAR, M.A., Professor of Mathematics, University College, Auckland.

[*Read before the Auckland Institute, 6th July, 1903.*]

Plates IV. and V.

THE statistics of the ages of the people afforded by census returns are marred by an element of uncertainty due to the ignorance, carelessness, and wilful misrepresentation of individuals. A person ignorant of his own age, ignorant of and careless in ascertaining the ages of members of his household, or wilfully misrepresenting the age of himself or the ages of others, is very apt to choose numbers ending in 0 or 5. Many, for instance, knowing an age to be about 40, will put it down in a round number as 40 without further bother. The result is that in census returns generally we find a greater number of people returned as being of such ages than is consistent with the numbers returned for neighbouring ages, as is shown in the following table:—

TABLE I.—NUMBER OF PEOPLE RETURNED AT CERTAIN AGES IN NEW ZEALAND CENSUS, 1896.

Ages.	Number of People.	Ages.	Number of People.
39-40	7,224	49-50	4,926
40-41	10,186	50-51	7,909
41-42	5,556	51-52	4,092
44-45	5,751	54-55	5,285
45-46	7,676	55-56	5,699
46-47	5,741	56-57	5,562

Here the numbers recorded as being of ages 40, 45, 50, and 55 are evidently in excess of what should be.

The numbers ending in 0 are more patronised than those ending in 5. Also, numbers ending in 2 and in 8 are favourites in the same way, but to a much less extent. The result is that at these ages, and especially at those which are multiples of 10, we have clusters, as it were, of population in the census returns of ages, and corresponding deficiencies in the other ages. Now, as a matter of fact, in all populations that have not been subject to some very remarkable and unusual event that would tend to increase or diminish

the number of people at particular ages above or below the normal, the number of people at the several ages should change gradually and continuously from age to age. If the population be grouped into quinquennial age-periods, we find that the numbers in the successive age-periods do change in this manner.

Figs. 1, 2, illustrate these remarks by the graphical method. Fig. 1 gives the line which represents the numbers returned at the several ages from 20 to 80 years in the census of 1896, and also a similar line constructed by taking for the central year of each quinquennial age-period the average population of each year of age in that period. The ordinates for these averages are measured along those corresponding to the central years—namely, those ending in 2 or 7. Fig. 2 illustrates similarly the results of the census of 1901.

The curve given by taking the population in quinquennial age-periods offers a great contrast to the other. The absence of any pronounced angular projections indicates the absence of any serious tendency for the ages of people to be entered in special quinquennial periods. The population, then, as arranged in quinquennial age-periods may be taken as being practically correct, and the corresponding curve gives us a fair average curve with which to compare the other. The excesses entered in special years are consequently seen in each case to be drawn in the main from the neighbouring years.

A comparison of the lines for the two census years shows at once a remarkable resemblance in the forms of the lines representing the population recorded according to each year of age, indicating a remarkable general persistence in the tendencies we are considering. The variations from the average are very small until the age of 20 is reached, and this is why the figures are taken only as beginning at that equation. The variations are not considerable until the age of 30 years, and are greatest, relatively to the population of the age concerned, at 50 and 60 years.

A careful comparison of the two diagrams will make it obvious that in 1901 the variations due to erroneous returns of age were considerably less than in the previous census. We say a *careful* comparison because the mere magnitudes of the variations do not, as a rule, differ greatly in the two cases; it is the variations relatively to the population at each age that differ considerably. If the number of erroneous returns made in 1901 were proportionate to those of 1896 they ought, absolutely considered, to be much greater, because of the much greater population between the ages 20 and 80 years in the later of the two years. This point is also illustrated in the following table, which takes the populations returned at

each year of age ending in 5 or 0 and compares them with the means of the numbers returned for the two neighbouring ages in each case. The percentages of excess are then given.

TABLE II.—VERACITY OF CENSUS RETURNS OF AGE

Age.	1896.			1901.		
	Mean of Numbers of Adjacent Ages.	Number of Age specified.	Per-centage of Excess.	Mean of Numbers of Adjacent Ages.	Number of Age specified.	Per-centage of Excess
20	14,961	14,850	-0·7	16,898	17,142	1·4
25	12,919	13,266	2·7	15,158	15,266	0·7
30	9,100	12,079	32·7	11,337	14,682	29·5
35	8,373	9,226	10·1	9,840	10,287	4·5
40	6,390	10,186	59·4	7,816	10,795	51·7
45	5,746	7,676	33·6	6,617	8,044	21·6
50	4,509	7,909	75·4	5,133	8,134	58·4
55	5,424	5,699	5·1	4,784	4,949	3·4
60	3,048	5,445	78·6	3,883	5,965	53·1
65	2,419	3,088	25·6	3,580	4,095	14·4
70	1,172	1,763	50·4	1,896	2,662	40·4
75	873	927	6·2	1,100	1,157	5·2
80	290	414	42·8	466	550	18·0

It will be noticed in this table that in every case, except for the age of 20, there is a diminution in the percentage of excess, that in the majority of cases this diminution is considerable, and that in the one case in which there is an increase the excess in both censuses is only very small. Such small excesses, moreover, should be ignored, as they are no greater than the variations in the annual number of births.

The above is but a rough method of dealing with the question, but it is really sufficient for the purpose in hand, as the reductions are so considerable. As, however, the excess is measured for any year from the mean of the recorded populations of the two neighbouring years, which are themselves less than the true populations of those years, the excess obtained is not the excess of the recorded population over the true population for the year of age under consideration, but is considerably greater than it. The diminution of the one excess involves, however, a somewhat proportional diminution in the other, and so the above argument is not vitiated, although we must not take the percentages of Table II. as measuring the proportion of erroneous returns for the several years of age.

There is another way of approaching the problem. The census tables, giving the numbers recorded at each age, may be adjusted by distributing the total numbers in certain groups of ages over the single years according to actuarial laws and methods. This has been done by the Actuary of the Government Life Insurance Department for the last two censuses, and the results are recorded, not in the ordinary census volumes, but in the reports of the Registrar-General on the same. Table III., following, takes for both censuses these adjusted numbers, compares with them the recorded numbers, and presents the excess as a percentage, the ages considered being from 20 to 75, at which latter age the tables of the Government Actuary cease.

TABLE III.—VERACITY OF CENSUS RETURNS OF AGE.

Age.	1896.			1901.		
	Adjusted Population at Ages specified.	Recorded Population at Ages specified.	Percentage of Excess.	Adjusted Population at Ages specified.	Recorded Population at Ages specified.	Percentage of Excess.
20	14,755	14,850	0·7	16,725	17,142	2·5
25	12,620	13,266	5·1	15,254	15,266	0·1
30	9,975	12,079	21·1	12,199	14,682	20·4
35	8,359	9,226	10·4	9,963	10,287	3·3
40	7,545	10,186	35·0	8,245	10,735	30·2
45	6,150	7,676	24·8	7,161	8,044	12·3
50	5,708	7,909	38·6	5,847	8,134	39·1
55	5,085	5,699	12·1	4,968	4,949	-0·3
60	3,845	5,445	41·6	4,451	5,965	34·0
65	2,565	3,038	18·4	3,728	4,095	9·8
70	1,390	1,763	26·9	2,248	2,662	18·4
75	750	927	23·6	1,083	1,157	6·8

As was to be expected, the percentages of excess in this table are smaller than in Table II. They give a reasonably accurate estimate of the proportions of the returns, at the several ages considered, that are erroneous. Comparing the results for the two censuses, we get the same general result as before; in every case but two there is a reduction, and generally this reduction is considerable. The two exceptions consist of the age of 20—and in this case the percentages of excess are small in both censuses—and the age of 50—and in this case, although the percentage of excess is large in both census years, the increase in the later year is very small.

Thus the fact is conclusively established that in the recent census there was a much greater approach to truth in the

returns of age than in the preceding census. It would have been of interest to have considered some earlier censuses and found what the tendency was in this matter in earlier times, but the census returns prior to 1896 do not give the population for each year of age, and it is impossible to pursue the question.

As to what the reasons are—and such must exist—for this very substantial and satisfactory progress we cannot suggest anything with confidence. Countries differ very much in the matter we have been discussing, and on the whole the better the education of the people the more truthful are their returns of age. In India, on the one hand, the returns reach the height of the ridiculous; in 1890, for instance, out of 100,000 persons of all ages, the number returned as of 39 years was 322, and of 41 years was 216, while the number returned as of 40 years was no less than 5,240! In Germany, on the other hand, the results are far better than our own. One can hardly claim, however, that the education of the people of New Zealand has improved in five years to such an extent as would be necessary to explain the phenomenon. If the improvement I have pointed out were confined to or more conspicuous in the older ages, the influence of the old-age pension scheme, that has made many an old person better acquainted with his age, and would be likely to make him more precise in recording it, might be put forward as a main cause, but the improvement is about equally conspicuous on the whole throughout the period from 30 to 80 years of age. Of course, a higher moral sense in the people would suffice to explain everything; let us hope this is the true cause.

I have calculated tables for the sexes taken separately similar to those discussed above, but which I do not propose to publish. They reveal the fact that there is, on the whole, nothing much to choose between the sexes in the matter of inaccurate returns of age, in spite of the fact that the chief blame for this offence is popularly given to females. If females are more strongly tempted from motives of vanity to make inaccurate returns, it may be that a finer moral sense prevents them yielding to the temptation to any greater extent than the males. But, however the causes of such inaccuracies may differ in kind or in degree in the two sexes, the results in the main are the same. A greater preference in the males than in the females for the ages 35 and 45, and a greater preference in the females than in the males for the ages of 50, 60, and 70, are apparently the only decided differences that the tables reveal.