On the Extent of Digital Preference in Reporting of Ages in Pakistan

by

FARHAT YUSUF*

INTRODUCTION

In Pakistan, like many other developing countries of the world, age distributions availabe from the decennial population censuses and sample surveys have shown substantial distortions and irregularities [2; 3; 4; 6, pp. 64-75; 9, pp. 638-658; 13; 14, pp. 64-95]. Some of these distortions could be real and may have been the result of events such as the Bengal famine of 1943 and the post-Independence migration between India and Pakistan. Others could be due to the coverage and response problems encountered in the collection of age data. Among the coverage and response problems, two are of most importance: underenumeration of females and erroneous age-reporting. In countries like Pakistan, which have low literacy rates (19.2 per cent literates according to the 1961 Census of Pakistan), most of the people do not know their correct ages. As a result they tend to report their ages either in round numbers or instead ask the enumerators to write down whatever age they think proper. This phenomenon of reporting ages in round numbers is usually called "digital preference". As a result of this the single-year age distributions show distinct peaks and troughs at ages ending with certain digits.

Measures are available with the help of which one can study the extent of digital preference in age reports [4; 5; 12]. In this paper an attempt has been made to use one such measure, Myers' index [7, pp. 395-415], for studying the extent of digital preference in age-reporting in Pakistan. This index measures preferences of respondents in reporting ages ending with each of the ten unit digits (*viz.* 0, 1, 2, . . . , 9). Knowing the pattern of digital preference one can make necessary compensatory adjustments in the age distributions either by smoothing or by using age groupings which minimize the effect of digital preference.

---

* Dr. Farhat Yusuf is a Research Demographer at the Pakistan Institute of Development Economics and is also associated with the Population Growth Estimation (PGE) project as its Deputy Principal Investigator. He wishes to acknowledge his gratitude to the Central Statistical Office of the Government of Pakistan, the Pakistan Institute of Development Economics, the Population Council Inc., New York and the National Centre for Health Statistics of the United States Government, for their support to the PGE Project, which made this research possible. The author also expresses his deep sense of gratitude to Dr. Lee L. Bean and Mr. William Seltzer for their helpful comments and suggestions. The views expressed in this paper are his own and do not necessarily reflect those of the PGE project or its sponsoring organizations.

DATA

The basic data required for this study of digital preference are the single-year age distributions. Unfortunately, these were not tabulated in the 1951 and 1961 Censuses of Pakistan. However, these tabulations were available for the age data collected by the Population Growth Estimation (PGE) project from a probability sample of the national population.

The data utilized in this paper were collected in the cross-sectional (CS) surveys conducted by the Central Statistical Office, on behalf of the PGE, in 10 sample areas each in East and West Pakistan. In these surveys the enumerators were expected to visit each household in the sample area in order to collect information about the household composition and the number of births and deaths occurring during the 12 months prior to the interview.

Between 1962 and 1965, four CS enumeration visits were undertaken each year, followed by a final enumeration visit in January 1966 (*see*, Table I). It can be seen from the table that visits 1.0, 2.1, 3.1 and 4.1 (marked with an asterisk) were the full-enumeration visits, while the remaining 13 visits were the fill-up visits. During the *full-enumeration visits*, the enumerators were asked to collect

TABLE I

VISIT NUMBER AND MONTH ENUMERATION BEGAN FOR 17 CS SURVEY VISITS, PGE: 1962-1966

Visit number	Year	Month
1.0*	1962	January
1.1		April
1.2		July
1.3		October
1.4/2.0	1963	January
2.1*		April
2.2		July
2.3		October
2.4/3.0	1964	January
3.1*		April
3.2		July
3.3		October
3.4/4.0	1965	January
4.1*		April
4.2		July
4.3		November
4.4	1966	January

Notes: 1) Visits marked with an asterisk (*) were the full-enumeration visits, while those not marked were the fill-up visits.

2) Visits 1.4, 2.4, and 3.4 were also called as 2.0, 3.0 and 4.0 respectively.

3) Visit 4.3 was delayed because of the Indo-Pakistan War. It was conducted in November instead of October 1965.

information about age, sex, marital status, occupation, *etc.*, of all persons reported as usual members of the household (whether present or absent) along with visitors, guests, servants, *etc.*, who were present at the time of enumeration [8]. This information was recorded in duplicate, the original being kept by the enumerator and the carbon copy was sent to the PGE headquarters. In the full-enumeration visit the enumerator had no record of the responses in the previous fill-up visit. During the *fill-up visits*, the enumerators were told to varify the information about household composition collected in the preceding full-enumeration visit and to record any changes or discrepancies therein. In both the full-enumeration and fill-up visits, the information about presence and absence of each enumerated person was recorded.

Information about age was obtained anew in each full-enumeration visit for persons enumerated at that visit. Although there were specific instructions to ask about age at each fill-up visit, however, it is doubtful if this instruction was followed. Moreover, the enumerators were also asked not to alter the age reports in the full-enumeration visits if the discrepancy in the fill-up visits was of less than three years. For persons not enumerated in the full-enumeration visit, for example, migrants into the area or persons missed in the full-enumeration visit, age was obtained as of the visit in which they were first enumerated.

Appendix Tables A-1 and A-2 show the basic data used in this paper[1]. These single-year age-sex distributions are for all persons reported as present in the four mid-year enumeration visits (*viz.*, visits 1.2, 2.2, 3.2 and 4.2) in 10 sample areas each in East and West Pakistan. It may be noted that for the computation of Myers' indexes only the age range 10-79 years was used (*see*, Appendix B).

QUALITY OF AGE-REPORTING IN PAKISTAN

In populations with no misstatement of age, because of the effect of mortality alone, it is expected that the number of persons reported at any age x will always be more than persons aged $x+1$. To overcome this problem, Myers' index is derived by forming equally weighted population totals for each of the ten terminal digits. Such weighted population totals are known as *blended populations*[2]. The index itself is derived by summing the absolute differences of per cent blended populations for each terminal digit from 10 per cent. Theoretically, the index can vary between 0 and 180. In developed countries, the value of this index is quite low. For example, in United States the value of Myers'

[1] This information is published in full since to-date no similar data have been made available for the population of Pakistan.

[2] For details, *see* Appendix B.

index was only 4.4 in 1950[11]. On the other hand, in countries with defective age-reporting the index may be quite high. In India, for example, the Myers' indexes computed from the 1961 Census age distributions were 70.7 for males and 75.1 for females[4].

For Pakistan, Myers' indexes were computed from the PGE age-sex distributions given in Tables A-1 and A-2. The values of these indexes are presented in Table II. It appears from the table that the quality of age-reporting for persons enumerated in the PGE sample was substantially better in East than in West Pakistan. Similar differences were observed in the 1961 Census of India where it was noted that the states bordering East Pakistan had lower values of the Myers' index compared with those states which had common borders with West Pakistan[4]. To some extent this inter-province differential in the quality of age-reporting in Pakistan may be due to the fact that the literacy rate in East Pakistan is higher than in West Pakistan (21.5 per cent literates in East Pakistan compared with 16.3 per cent in West Pakistan, according to the 1961 Census). Some other recent studies of age distributions of Pakistan also give support to this hypothesis of positive relationship between levels of literacy and the quality of age-reporting[1, pp. 85-90].

TABLE II

MYERS' INDEX OF DIGITAL PREFERENCE FOR THE AGE RANGE 10-79 YEARS, PGE: 1962-1965

Sex/province	1962	1963	1964	1965
Males				
East Pakistan	60.4	54.1	54.0	59.4
West Pakistan	81.4	71.4	77.3	65.9
Females				
East Pakistan	60.8	57.1	57.8	59.0
West Pakistan	81.9	72.2	79.2	65.0

Another interesting fact which is revealed by Table II is that over the four years of PGE operations, the quality of age-reporting in West Pakistan sample areas improved substantially while in East Pakistan it remained more or less constant. As pointed out above, during the four years under study, each household in the sample area was enumerated 17 times. Thus, it is possible that these 17 quarterly enumeration visits helped to bring up the more illiterate West Pakistani respondents to the level of age-reporting comparable to that in East Pakistan. However, it is clear from our experience in both East and West Pakistan that even after 17 enumeration visits in four years the quality of age-reporting remained quite poor.

There was no substantial sex differential in the quality of age-reporting in either East or West Pakistan. This is perhaps due to the fact that respondents in most cases were males and they apparently did not have different digital preferences for reporting ages of females.

Table III presents the per cent blended populations (age range 10-79 years) for each of the ten terminal digits. These were derived from the PGE age distributions presented in Tables A-1 and A-2, by using the method described in Appendix B. Here it may be mentioned that in populations with no misstatement of age, it is expected that the blended populations for all digits will be equal (*viz.*, 10 per cent of the total blended population). Thus, if for a given digit the blended population is different than 10 per cent, this difference is taken as a measure of preference for that digit.

TABLE III

BLENDED POPULATIONS FOR THE AGE RANGE 10-79 YEARS, PGE: 1962-1965

Terminal digit	1962		1963		1964		1965	
	Male	Female	Male	Female	Male	Female	Male	Female
East Pakistan								
0	27.45	29.38	25.24	28.96	25.99	30.94	24.65	27.36
1	3.58	3.61	4.87	4.89	4.76	5.06	2.75	2.88
2	10.92	9.61	10.55	9.57	10.62	9.59	12.43	10.95
3	4.12	3.46	5.28	4.68	4.62	3.96	4.53	4.02
4	5.58	5.55	5.24	4.94	6.08	4.97	5.87	5.96
5	21.37	20.48	20.83	18.68	19.07	17.56	21.82	19.59
6	8.89	8.63	9.46	8.78	10.58	10.26	9.55	10.07
7	4.17	3.77	4.66	3.78	4.29	3.68	4.90	4.40
8	10.47	10.52	10.65	11.08	10.71	10.12	10.95	11.53
9	3.45	4.99	3.42	4.64	3.28	3.86	2.55	3.24
Total	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
West Pakistan								
0	33.22	35.40	30.67	32.21	32.29	33.09	28.72	29.02
1	1.94	1.59	3.54	3.59	2.06	1.71	2.71	2.39
2	9.50	8.90	9.42	10.23	9.46	9.58	9.86	10.52
3	3.40	3.26	3.81	3.19	3.13	3.10	4.03	4.01
4	4.04	4.12	3.78	4.31	4.10	4.62	5.00	5.94
5	27.46	25.53	25.02	23.64	26.33	26.35	24.02	22.79
6	6.40	6.22	8.79	7.89	7.31	6.86	7.79	7.73
7	3.32	3.60	4.11	4.19	3.09	2.75	4.82	4.83
8	8.60	9.05	8.02	8.20	10.01	10.17	10.24	10.18
9	2.12	2.33	2.84	2.55	2.22	1.77	2.81	2.59
Total	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00

It appears from Table III that digit 0 was most popular in both East and West Pakistan. The blended population for this digit varied between 25 to 31 per cent in East Pakistan and between 29 to 35 per cent in West Pakistan. The next popular digit was 5 which had blended populations ranging between 17 to 21 per cent in East and 23 to 27 per cent in West Pakistan. In order of preference these two digits (*i.e.* 0 and 5) were followed by digits 2, 8, 6, 4 and then the remaining odd digits. Moreover, it can also be noted from Table III that digits 0 and 5 gained popularity mainly at the expense of digits 1, 3, 4, 7 and 9 which had blended populations varying between 2 to 6 per cent. However, digits 2, 8 and 6 got about their right share in the total blended populations, *viz* 10 per cent approximately.

SUMMARY AND CONCLUSIONS

Three main conclusions can be drawn from the analysis of PGE age data presented in this paper. First, the general quality of age-reporting in Pakistan is so poor that it is not advisable to place much reliance on the age distributions prior to making compensatory adjustments. Second, the quality of age-reporting in East Pakistan is comparatively better than in West Pakistan. The third conclusion refers to the general pattern of digital preference: digits 0 and 5 were most popular with digits 2, 4, 6 and 8 next in popularity and followed by the remaining odd digits. Similar patterns of digital preference have been found in other developing countries of the world [4;12].

It is unfortunate that the single-year age distributions were not made available in the 1951 and 1961 Censuses of Pakistan. Thus, it is not possible to compare the quality of PGE single-year age distributions with those of the Census. In addition, it may be pointed out that Myers' indexes computed from the census data for other countries are not strictly comparable to the indexes derived from the uninflated PGE sample figures. It is hoped that the single-year age distributions will be made available for the 1971 Census of Pakistan so as to enable an extensive study of distortions and irregularities in Pakistani age data. Until the single-year age distributions are available from the Census of Pakistan, the information presented in this paper, although based on a sample survey, is quite relevant to demographic analysis in Pakistan.

The age distribution is one of the most important input variables for some of the recent and more sophisticated methods of demographic analysis, for example, the stable and quasi-stable population techniques. These demographic techniques, under certain assumptions, have been utilized for estimating vital rates and other demographic parameters for Pakistan [9,pp. 638-658;14,pp.64-95]. It is obvious that if the age distribution of any population is distorted the estimates

of demographic parameters derived from such an age distribution will be adversely affected. To avoid such situations, the distortions in age distributions are analyzed by methods which enable us to determine the nature and extent of digital preference in age reports. Knowing the pattern and extent of digital preferences, compensatory adjustments can be made by using age groupings which minimize the effect of digital preference and by smoothing the age distributions. The necessity of doing both can be illustrated from Table III in which the cumulative proportions for digits 0 through 4 are slightly greater than the proportions for digits 5 through 9. Thus, even grouping the age distributions in 0-4, 5-9, 10-14,..... age groups will still produce distorted age distributions and for such analysis as life table construction smoothing will have to be done.

REFERENCES

1. Ahmed, Khwaja Mansoor, "Literacy and its Relation with Error in Age Distribution—A Study Based on 1961 Population of West Pakistan", *Proceedings of the Pakistan Statistical Association*, Vol. 13, 1965-66.

2. Hashmi, Sultan S., *Main Features of the Demographic Conditions in Pakistan.* Country background paper presented to the Asian Population Conference, New Delhi. (Karachi: Central Statistical Office, 1963).

3. Hashmi, Sultan S., *The People of Karachi: Demographic Characteristics.* (Karachi: Pakistan Institute of Development Economics, 1964).

4. India, *Census of 1961, Age Tables, Paper No. 2 of 1963.* (New Delhi: Manager of Publications, 1963).

5. Jaffe, A.J., *Handbook of Statistical Methods for Demographers.* (Washington: United States Department of Commerce, Bureau of the Census, 1960).

6. Krotki, Karol J., "A First Glance at the Pakistan's Age Distribution", *Pakistan Development Review*, Vol. 1, Summer 1961.

7. Myers, R.J., "Errors and Bias in Reporting of Ages in Census Data", *Transactions of the Actuarial Society of America*, Vol. 41, 1940.

8. Pakistan, Population Growth Estimation, *Instructions for Cross-Sectional Surveys.* Special Studies Series PGE 3. (Karachi: Central Statistical Office, Government of Pakistan, February 1964).

9. Robinson, Warren C., William Seltzer and Sultan S. Hashmi, "Quasi-Stable Estimates of the Vital Rates of Pakistan", *Pakistan Development Review*, Vol. 5, Winter 1965.

10. Rizvi, Hafsa and Shuja-ur Rehman, "Three Years' Experience in the Collection of Age Data, PGE: 1962-64". Paper presented to the Fourth Pakistan Sociological Conference, Abbottabad, August 1967.

11. Spiegleman, Moritimer, *Introduction to Demography.* (Chicago: The Society of Actuaries, 1955).

12. United Nations, *Methods of Appraisal of Quality of Basic Data for Population Estimates, Manual II.* (New York: United Nations, 1963).

13. Yusuf, Farhat, *Differential Fertility in Lahore, Pakistan.* Unpublished Ph.D. thesis presented at the Australian National University. (Canberra: April 1966).

14. Zelnik, Melvin and Masihur Rahman Khan, "An Estimate of the Birth Rate in East and West Pakistan", *Pakistan Development Review,* Vol. 5. Spring 1965.

Appendix A

TABLE A-1

SINGLE-YEAR AGE-SEX DISTRIBUTIONS FOR DEFACTO MID-YEAR POPULATIONS OF 10 CS SAMPLE AREAS IN EAST PAKISTAN PGE: 1962-1965

Age	1962		1963		1964		1965	
	Male	Female	Male	Female	Male	Female	Male	Female
<1	875	890	969	1,007	949	921	989	1,037
01	568	582	657	621	581	602	753	737
02	1,044	1,159	901	948	1,021	993	966	990
03	1,004	996	968	1,019	930	946	1,106	1,024
04	953	944	1,065	1,030	1,048	1,104	1,035	1,042
05	1,010	917	953	924	979	944	1,102	1,107
06	971	994	1,123	1,069	1,161	1,167	1,147	1,162
07	861	834	850	910	863	912	905	898
08	1,059	983	1,135	1,084	1,231	1,217	1,254	1,184
09	617	606	732	720	770	739	719	736
10	1,052	911	1,073	862	1,045	888	1,094	890
11	425	402	500	419	510	422	484	448
12	951	588	997	650	1,059	721	1,045	710
13	423	336	487	346	462	371	478	372
14	525	442	550	418	616	459	635	500
15	489	355	536	400	556	389	492	435
16	522	456	541	477	578	510	566	626
17	219	256	262	217	286	235	328	303
18	599	581	585	581	546	514	594	667
19	203	336	198	272	233	254	214	223
20	632	828	615	945	626	885	587	811
21	143	190	171	206	181	272	160	175
22	429	499	422	444	422	451	546	586
23	197	190	252	250	204	213	214	245
24	271	303	226	262	292	271	262	334
25	773	864	856	870	707	855	797	858
26	392	406	420	394	440	578	387	471
27	237	193	260	190	209	174	245	221
28	487	452	501	489	499	446	492	486
29	119	157	139	146	123	134	83	128
30	1,018	986	1,026	1,041	1,051	1,149	994	1,015
31	115	97	160	132	172	146	81	83
32	419	294	411	314	403	317	503	353
33	112	72	181	125	139	96	159	97
34	127	93	143	97	165	90	156	133
35	828	683	851	609	766	612	961	704
36	316	234	335	233	459	297	388	261
37	110	62	126	85	126	75	153	106
38	239	194	283	230	324	243	350	266
39	87	75	79	117	63	72	56	59
40	871	775	845	790	850	887	877	749

(Contd.)

TABLE A-1 (*Contd.*)

Age	1962		1963		1964		1965	
	Male	Female	Male	Female	Male	Female	Male	Female
41	98	58	121	116	137	89	39	39
42	245	170	254	199	247	193	299	246
43	60	44	92	63	101	54	69	37
44	61	67	78	51	80	41	91	66
45	628	508	605	457	609	433	750	506
46	119	99	156	117	187	121	159	119
47	64	37	68	42	53	53	60	34
48	131	148	222	183	223	148	215	192
49	57	63	73	75	59	62	31	34
50	701	628	671	647	716	732	649	714
51	49	33	112	75	92	69	31	14
52	138	91	108	117	159	108	171	92
53	39	14	61	50	52	34	44	33
54	61	36	58	42	62	49	64	51
55	359	254	341	222	354	203	376	281
56	73	80	104	79	92	67	128	117
57	15	3	35	27	24	18	34	20
58	96	58	62	77	79	89	80	68
59	24	22	26	21	16	18	10	15
60	493	455	447	375	497	486	485	417
61	28	20	67	52	44	45	19	6
62	63	41	77	42	76	56	94	52
63	16	5	25	16	30	7	31	10
64	28	16	21	21	30	17	29	12
65	203	169	192	165	156	117	218	158
66	27	9	52	49	45	18	38	26
67	5	2	14	3	22	9	14	4
68	32	25	38	28	38	34	45	29
69	12	11	9	10	14	4	6	2
70	198	164	175	130	191	139	186	113
71	4	2	20	13	10	8	6	4
72	15	12	21	16	23	10	34	25
73	12	1	6	6	6	7	5	3
74	5	7	5	2	6	4	5	5
75	72	87	65	43	54	41	71	54
76	7	6	15	22	16	15	10	8
77	3	2	5	5	4	6	5	2
78	2	10	2	6	10	6	15	13
79	1	2	2	2	2	0	4	0
80+	176	152	136	122	147	108	151	126
Unknown	3	4	2	2	0	0	3	2
Total	26,715	24,830	28,027	25,733	28,408	26,289	29,131	26,981

TABLE A-2

SINGLE-YEAR AGE-SEX DISTRIBUTIONS FOR DEFACTO MID-YEAR POPULATIONS OF 10 CS SAMPLE AREAS IN WEST PAKISTAN PGE: 1962-1965

Age	1962		1963		1964		1965	
	Male	Female	Male	Female	Male	Female	Male	Female
<1	830	757	845	784	893	820	1,106	1,071
01	773	667	770	696	768	796	712	644
02	951	926	914	825	951	889	914	846
03	1,035	1,029	989	991	1,015	952	1,155	1,087
04	923	918	971	905	1,059	927	1,051	918
05	951	835	929	857	963	969	1,138	980
06	942	831	971	864	1,034	965	1,049	914
07	774	764	795	769	874	710	958	885
08	1,007	851	912	827	993	1,019	1,013	929
09	509	482	617	512	629	528	744	612
10	970	768	886	738	993	781	965	786
11	398	319	429	371	407	356	457	386
12	838	639	899	615	999	710	959	774
13	418	351	462	376	420	384	601	436
14	538	417	506	405	542	419	567	448
15	613	375	568	372	478	373	540	418
16	603	473	650	465	627	463	512	452
17	290	239	339	271	318	215	375	240
18	733	693	699	595	772	577	735	522
19	231	231	291	239	266	194	333	239
20	858	956	726	800	781	772	810	815
21	139	103	199	196	180	120	213	170
22	549	561	565	602	622	657	565	595
23	191	183	226	191	210	186	262	262
24	236	240	233	305	291	361	288	382
25	1,173	1,256	1,030	1,031	1,159	1,172	1,065	1,025
26	305	292	406	385	424	382	407	371
27	177	190	209	213	179	136	315	323
28	394	373	361	333	541	544	499	505
29	64	53	96	78	57	41	82	92
30	1,203	1,267	1,058	1,175	1,291	1,317	1,185	1,188
31	41	25	113	128	32	28	56	43
32	360	251	332	327	369	282	415	395
33	81	61	107	63	82	63	105	99
34	61	65	64	45	55	63	93	139
35	1,097	871	993	806	1,140	957	1,088	800
36	154	151	278	223	185	179	252	263
37	59	49	84	68	43	46	101	107
38	142	117	120	137	202	213	271	276
39	20	22	33	16	26	12	36	42
40	1,094	971	956	916	1,135	1,059	1,011	888
41	20	12	91	57	27	17	40	27
42	166	122	129	152	160	133	185	212
43	34	33	44	16	26	30	38	42

(*Contd.*)

TABLE A-2 (*Contd.*)

Age	1962		1963		1964		1965	
	Male	Female	Male	Female	Male	Female	Male	Female
44	32	24	16	17	33	13	52	55
45	851	593	818	615	915	769	842	693
64	59	46	132	84	73	56	167	108
47	31	39	43	39	24	27	63	64
48	92	76	97	79	151	132	179	193
49	8	8	13	13	13	6	16	12
50	912	784	918	730	1,020	794	870	674
51	16	7	37	20	12	6	36	23
52	68	52	93	82	87	71	128	88
53	37	21	17	12	21	11	28	16
54	15	16	12	16	15	17	94	75
55	471	303	411	313	485	415	385	325
56	40	22	65	47	68	41	101	78
57	10	14	23	19	9	12	27	30
58	36	29	24	23	42	34	78	59
59	5	5	5	2	4	2	7	5
60	718	549	719	500	739	571	659	541
61	3	6	20	12	8	4	18	15
62	70	33	62	45	25	23	69	46
63	18	8	13	4	11	6	9	8
64	16	8	11	5	14	6	36	17
65	281	217	267	216	341	247	332	291
66	16	7	31	13	23	10	37	30
67	8	4	9	17	4	6	13	10
68	18	20	18	16	30	32	50	28
69	3	1	4	1	3	1	3	0
70	336	271	350	241	329	230	278	202
71	5	2	7	6	1	3	7	6
72	26	11	25	12	12	14	36	19
73	4	1	2	0	2	1	5	2
74	2	7	1	0	4	2	7	4
75	109	77	100	84	111	95	117	112
76	10	8	11	10	6	9	7	12
77	5	1	5	2	1	1	1	2
78	5	5	3	1	9	3	10	6
79	0	0	1	0	1	0	1	0
80+	316	258	269	225	262	228	197	204
Unknown	39	3	6	3	0	0	0	0
To tal	27,636	24,325	27,553	24,264	29,126	25,715	30,231	26,701

Appendix B

Computational Details of Myers' Index

The Myers' index shows the relative preferences of respondents for each of the ten unit digits *viz*, 0,1,2,...,9. The method consists of computing a *blended population* in which the expectation is to have equal sums for each digit provided there is no digital preference. For computing the blended populations a decision will have to be made about the age range on which to base the computations. Usually the lower limit is not less than 10 years and the upper limit is not more than 80 years. This is because the age-reporting at less than 10 and more than 80 years of age is largely affected by causes other than digital preference. In the present study the age range was taken as 10-79 years.

The Myers' index involves the computation of two series of population totals with a time lag of 10 years. In the present study one series has a range of 10-69 years while the range for the second series is 20-79 years. Denoting P(x) as the population at age x, the ten population totals in the first series will be:

$$\text{total for digit } 0 = P(10) + P(20) + \ldots + P(60)$$
$$\text{total for digit } 1 = P(11) + P(21) + \ldots + P(61)$$
$$\vdots \qquad\qquad \vdots$$
$$\text{total for digit } 9 = P(19) + P(29) + \ldots + P(69)$$

Similarly, the ten population totals in the second series will be:

$$\text{total for digit } 0 = P(20) + P(30) + \ldots + P(70)$$
$$\text{total for digit } 1 = P(21) + P(31) + \ldots + P(71)$$
$$\vdots \qquad\qquad \vdots$$
$$\text{total for digit } 9 = P(29) + P(39) + \ldots + P(79)$$

The ten population totals (one for each digit) of the first series are then multiplied by weights, 1, 2, 3, ..., 10 while the ten population totals of the second

series are multiplied by 9, 8, 7 ..., 2, 1, 0 respectively. The two sets of products are then summed for each terminal digit to arrive at the blended population for that digit. The blended populations for the ten digits are then converted into per cent of the total blended population. The Myers' index is derived by summing the absolute differences of the per cent blended populations for each terminal digit from 10 per cent.

The need for blending arises because even in populations with no misstatement of age it is expected that the number of persons living at successive ages will form a decreasing series due to the effects of mortality alone. For example, in a life table population which is subject solely to the effects of fertility and mortality, the two series of population totals referred to in the preceding paragraph would yield decreasing populations at successive terminal digits. Any index constructed from such a decreasing series would invariably exaggerate digits 0 through 4 at the expense of digits 5 through 9. To overcome this effect of mortality, Myers suggested the method of blending two population totals as already described in the previous paragraph.

By using this method we effectively give equal weight to each terminal digit. It can be shown algebraically that the sum of blended populations (age range 10-79 years) corresponding to the ten terminal digits is equal to the sum of populations in the age ranges 10-69, 11-70, 12-71,..., 19-78 and 20-79 years.