Demography India
Vol. 21, No. 1(1992), pp. 113-127

S. Mitra and Martin L. Levin

Modeling Survivorship Function Through Principal Component Analysis

The Problem

AN attempt to summarize life tables by factor analytic technique was first made by Ledermann and Breas (1959), using 154 of the 159 abridged life tables compiled originally by the Population Division of the United Nations for the construction of model life tables (1955). Ledermann and Breas factor-analyzed the correlation matrix of the age-specific probabilities of dying finding five factors to be significant, a relatively large number considering the number of original variables in an abridged life table. The reason so many factors were found perhaps lies in the nature of the pattern of variation of the probabilities of dying over the life span. As is well known, the curve generated by these probabilities is U-shaped with a highly variable minimum age of mortality. This feature of the distribution appears to have contributed to what we regard as a large number of factors. One may wonder why other life table functions were not subjected to the same treatment to test their differential performance. For the construction of their model life tables, Coale and Demeny (1966,1983) used the age-specific probabilities of dying plus their logarithmic transformation along with, for reasons they did not specifically detail, the expectation of life at age ten. It seems that the choice of the probabilities of dying were uniformly dictated by the fact that they were immediately and directly derivable from the primary data set in the form of age-specific death rates and it is on these rates that the computation of the life table functions are based. However, from a technical point of view, any life table function can be converted to any other. Accordingly, the selection of one function over another may be guided by theoretical and/or pragmatic considerations. Quite naturally, we decided to experiment with the very first function of age x— the survivorship probability from birth to age x, a monotonically declining function of X in all life tables. It is well known that this function, traditionally denoted by $l(x)$, assumes values at different ages which are not only highly correlated with one another but also with the overall level of mortality as measured by life expectancy, the crude death rate, etc. Accordingly, a data reduction procedure such as principal components would seem to be ideally suited for an in-depth study of the underlying pattern of variation of survivorship function by age.

* Professor S. Mitra and Dr. Martin L. Levin are in the Department of Sociology, Emory University, Atlanta, GA 30322.

Data and Analysis by the Method of Principal Components

A set of 120 abridged life tables each for males and females published by Keyfitz and Fleiger (1971) was selected for this study. These life tables have a standard width of 1 year for the first age interval beginning at age 0,4 years for the second interval, 5 years for all subsequent intervals and end at age 85. The selection of this set of life tables was made arbitrarily from alternatives such as UN model tables (1956), Coale and Demeny's model life tables (1966,83), and life tables published regularly in Demographic Yearbook. In any case, the primary data requirement for such studies is that the life tables cover a wide range of mortality levels and the data are of high quality. The selected data set meets the first requirement reasonably well. However, the assessment of data quality is somewhat more problematic. The mortality statistics and the population composition on which the life tables are based are known to be less accurate, especially in many of the developing countries. Consequently, minor adjustments were made to the values of /(80) in six life tables, four for females and two for males. These adjustments were necessary because the recorded values of age-specific death rates showed an inverse relationship with age around age 80 (see Appendix I).

As mentioned earlier, the *l(x)* values for any two ages say *x* and *X2* are expected to be positively correlated. It is also expected that its value at any age is the strongest determinant of its value at the next a(ge. As a result, the correlation matrix of the *l(x)'s* with one another should[1] contain declining values along any column below the major diagonal. Table 1 confirms this proposition.

The correlations between successive ages (separated by four years for the first interval and by five years thereafter), are not only the largest of their respective columns but their values are also quite large. Of 17 such correlations, the lowest values recorded are between the last pair of ages, 80 and 85, equalling .87 for the male matrix and .89 for the female matrix. Eleven correlations for the male tables and 13 for the female have values in excess

*

of .99. The minimum correlation is between the two extreme ages 1 and 85, equalling .11 and .47 for the male and female tables respectively. In general, the female correlations are larger than the corresponding male correlations, and the difference grows as the gap between the ages widens. It seems that relatively lower mortality rates among females at any age makes their survivorship probabilities more predictable than the male probabilities. In any case, 80 per cent of the correlations for the male data, and 90 per cent of the correlations for the female data exceed .6, indicating a simple underlying structure of the variables. Indeed, the principal components analysis reveals that structure to be comprised of only two factors with eigenvalues greater than 1 and with a joint explanation of 97.1 per cent of the variance for male and 97.8 per cent for female life tables. These values may be seen in Table 2.

The statistics for the third factor were included in Table 2 to demonstrate its relative insignificance compared to the eigenvalues for the first two factors. This suggests that, for all practical purposes, two factors provide a reasonable summary of the 18 survivorship functions expressed as

TABLE 1. CORRELATION MATRIX OF THE Sl^RVIVORSHIP PROB
Male

Age	1	5	10	15	20	25	30	35	40	45	50	55	60	65	70	75	80
5	.97860																
10	.96088	.99591															
15	.95498	.99343	.99959														
20	.94641	.98879	.99760	.99874													
25	.93651	.98342	.99451	.99609	.99873												
30	.92550	.97625	.98971	.99191	.99601	.99906											
35	.91686	.97084	.98495	.98720	.99203	.99681	.99899										
40	.90897	.96423	.97977	.98257	.98837	.99385	.99732	.99~903									
45	.89989	.95594	.97195	.97499	.98103	.98786	.99247	.99579	.99823								
50	.88782	.94454	.96055	.96384	.96963	.97754	.98315	.98803	.99219	.99747							
55	.87318	.92958	.94619	.94988	.95547	.96351	.96961	.97507	.98125	.98928	.99634						
60	.83583	.89407	.91326	.91809	.92433	.93287	.94032	.9463*	.95575	.96729	.98006	.99277					
65	.78860	.84677	.86508	.86986	.87471	.88431	.89223	.89991	.91094	.92571	.94525	.96757	.98739				
70	.69975	.75733	.77702	.78273	.79046	.80099	.81100	.81984	.83533	.85456	.88009	.91280	.94952	.97888			
75	.58370	.63661	.65653	.66285	.67250	.68320	.69426	.70368	.72199	.74560	.77537	.81675	.86672	.91256	.97100		
80	.35367	.40408	.4261 1	.43369	.44558	.45705	.46852	.47324	.49782	.52459	.55647	.60413	.66960	.73477	.83797	.93038	
85	.10848	.14465	.16437	.17092	.18263	.19275	.20242	.21195	.22793	.25427	.28287	.32627	.38776	.45753	.57877	.70901	.86858

Table 1 *(contd. on* p. 116)

Table 1 (*contd. from.* 115)

Female

Age	1	5	10	15	20	25	30	35	40	45	50	55	60	65	70	75	80
5	.97442																
10	.96067	.99734															
15	.95382	.99484	.99947														
20	.94696	.99156	.99808	.99938													
25	.93980	.98739	.99565	.99768	.99940												
30	.93351	.98342	.99291	.99545	.99801	.99952											
35	.92905	.98040	.99035	.99305	.99615	.99824	.99945										
40	.92810	.97825	.98806	.99070	.99393	.99637	.99794	.99920									
45	.92383	.97412	.98430	.98728	.99053	.99321	.99501	.99677	.99895								
50	.91565	.96722	.97832	.98188	.98532	.98834	.99052	.99288	.99596	.99858							
55	.90959	.96047	.97159	.97531	.97855	.98177	.98426	.98702	.99098	.99488	.99816						
60	.89147	.94371	.95592	.96038	.96394	.96780	.97092	.97464	.97990	.98575	.99244	.99703					
65	.88145	.92917	.93908	.94264	.94589	.94983	.95342	.95820	.96544	.97293	.98093	.98903	.99517				
70	.85131	.89550	.90466	.90767	.91137	.91594	.92039	.92682	.93627	.94573	.95647	.96670	.97894	.99129			
75	.81954	.85731	.86554	.86785	.87138	.87601	.88076	.88779	.89834	.90879	.92014	.93191	.94746	.96689	.98927		
80	.71525	.75135	.76159	.76479	.76921	.77424	.77953	.78697	.79585	.80624	.81964	.83149	.85241	.87591	.91909	.95776	
85	.47103	.50106	.51378	.51835	.52321	.52822	.53430	.54020	.54615	.55494	.58154	.56936	.60594	.63375	.69449	.76339	.89390

where $sl(x)$ is the standardized form of $l(x)$; the two factor loadings for age x are $f(1,x)$, and $f(2, x)$ and the two factors F(l) and $F(2)$ also appear in their standardized forms. The factor loadings (after varimax rotation) as well as the communalities given by

$$h^2(x)=f^2\ (1, x)+f^2(2,x) \tag{2}$$

may be seen in Table 3.

TABLE 2: EIGENVALUES AND PERCENTAGE OF VARIANCE EXPLAINED BY THE FIRST THREE FACTORS BY SEX

Factor	Male		Female	
Eigenvalue	Percent Variance		Eigenvalue	Percent Variance
1	15.22	84.6	16.48	91.6
2	2.25	12.5	1.13	6.3
3	0.29	1.6	0.22	1.2

TABLE 3: ROTATED FACTOR MATRIX AND COMMONALITIES BY SEX

Age	Male			Female		
	Factor 1	Factor 2	Commonality	Factor 1	Factor 2	Commonality
1	.94176	.12076	.90150	.90283	.29695	.90329
5	.97189	.17140	.97395	.93386	.32795	.97965
10	.97435	.19658	.98799	.93414	.34261	.98999
15	.97317	.20530	.98921	.93302	.34843	.99192
20	.97112	.21892	.99099	.93107	.35557	.99332
25	.96880	.23395	.99331	.92783	.36427	.99356
30	.96448	.24935	.99239	.92355	.37395	.99278
35	.95993	.26338	.99084	.91879	.38539	.99270
40	.95259	.28718	.98990	.91930	.39886	.99429
45	.94081	.32093	.98811	.90605	.41600	.99399
50	.92403	.36228	.98509	.89338	.44002	.99175
55	.89876	.42023	.98436	.88007	.46265	.98857
60	.85056	.50162	.97508	.85144	.50391	.97888
65	.78325	.58846	.95976	.81843	.54943	.97170
70	.66520	.72040	.96146	.75089	.63425	.96611
75	.51050	.84256	.97052	.67413	.71785	.96976
80	.23750	.95592	.97019	.49906	.85769	.98469
85	-.04544	.92951	.86605	.18065	.94930	.93380

A factor loading, say, $f(1, x)$ is also the correlation between the corresponding factor and the variable i.e., $l(x)$ and F(l) in this example. For the first factor, we see from Table 3 that the first loading is rather large with a value of .94 at age 1. The second loading (age 5) is slightly larger, and the largest loading is .97 at age 10. The magnitude of the loadings declines slowly from that point until age 45 and then at an accelerated rate thereafter to assume a small negative value of-.05 at the last age. In contrast, the second factor loadings start with a small value of. 12 at age 1, increase thereafter to a maximum value of .96 at age 80, ending at age 85 with a slightly smaller value of .93. These opposing patterns of the two sets of factor loadings can be seen clearly in Figure 1.

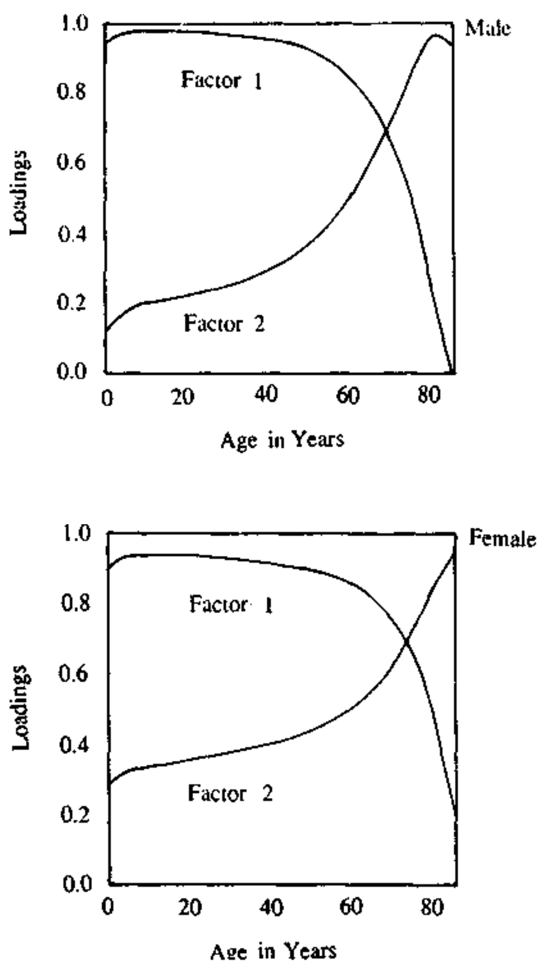


Fig.1. Factor Loadings by Age and Sex

Since, the loadings are quite large at earlier ages and relatively small at older ages for the first factor while the pattern is the reverse for the second, it may be said that the younger ages are heavily influenced by the first factor while the older ages are impacted by the second set. The first factor may therefore be labelled as the youth factor and the second as the elderly factor. Such distributions of the factor loadings, as well as the fact that the two factors are orthogonal to one another, make them complement one another in jointly accounting for the variation of the associated variable. Thus, a low loading on one factor for a particular age is associated with a high loading on the other, producing relatively large and stable values of the communalities. Consequently, compared to the factor loadings, the communalities have a small range of variation, .87 to .99 for the male and .90 to .99 for female tables. The communalities based on these two factors can be interpreted as the squared multiple correlations of *sl(x)* or of *l(x)* on F(l) and F(2). Observe from: Table 3 that 14 of the 18 communalities exceed .97. Consequently, these two factors should be able to reproduce the survivorship function with a high degree of accuracy.

TABLE 4: FACTOR SCORE COEFFICIENT MATRIX BY SEX

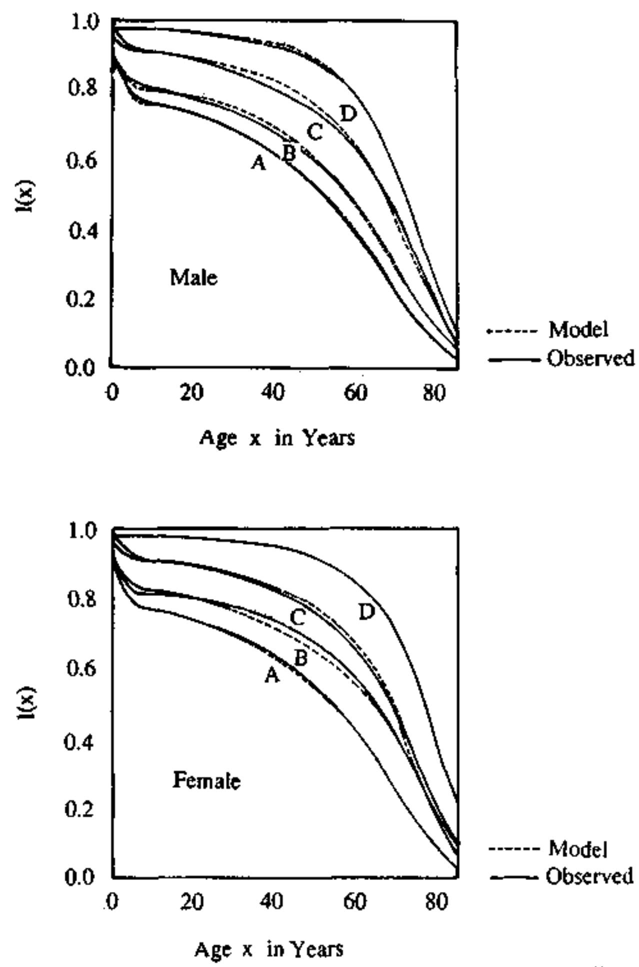
Age	*Male*		*Female*	
	Factor 1	Factor 2	Factor 1	Factor 2
1	.10538	-.08814	.12964	-.11476
5	.10209	-.07318	.12675	-.10442
10	.09883	-.06397	.12165	-.09445
15	.09743	-.06048	.11930	-.09006
20	.09523	-.05502	.11627	- .08446
25	.09279	- .04898	.11236	- .07735
30	.09003	-.04251	.10783	-.06919
35	.08744	- .03653	.10256	- .05965
40	.08310	-.02644	.09653	- .04867
45	.07676	-.01194	.08845	-.03413
50	.06870	.00616	.07668	-.01316
55	.05716	.03178	.06523	.00708
60	.03930	.06957	.04323	.04552
65	.01818	.11215	.01859	.08843
70	-.01596	.17909	- .02889	.17053
75	-.05347	.24753	-.07834	.25504
80	-.10514	.32946	-.17316	.41288
85	-.13820	.35969	-.28819	.58819

Reproducibility of Life Tables by Two Factors

To reproduce a life table at this stage of the principal components analysis, it is necessary to compute the two factor scores corresponding to that life table. The procedure is quite routine with the factor scores derived from the equation

$$F(i) = \sum_x w(i, x)\, sl\,(x) \qquad (3)$$

where $w(i, x)$ is the factor weight for the ith factor corresponding to age x. The factor weights generally resemble factor loadings, and the result here is no exception as shown in Table 4.



	Life Expectancy	
	Male	Female
A———Djakarta 1961	44.54	46.13
B———South Africa (Colored) 1961	49.85	54.56
C———Honduras 1966	59.24	60.70
D———Japan 1964	67.72	72.94

Fig. 2. Observed vs. Model Value of the $l(x)$ Function: Selected Countries

Substituting (3) in (1) gives the standardized values of *l(x)* which can then be converted to produce the model estimates *Ml(x)* of *l(x)*. In view of the large communalities found for each *x*, these *Ml(x)* values should be highly correlated with the */(x)* values for any life table. However, by itself, a high correlation is not a sufficient indicator of a good match between the observed and the model values. The regression *of Ml(x)* on *l(x)* should also produce an intercept close to 0 and a slope close to 1. The results of this study show this to be the case. For male tables the lowest correlation is .9926 and for female tables is .9877. A perfect

TABEL 5 : PREDICTED VALUES OF LIFE EXPECTANCY «(O) FOR SELECTED FACTOR SCORE COMBINATIONS BY SEX

Male — Second Factor Score

First Factor Score	-2.25	-1.75	-1.25	-.075	-0.25	0.25	0.75	1.25	1.75	2.25
-2.25	45.40	46.95	48.49	50.03	51.58	53.12	54.66	56.21	57.75	59.29
-1.75	48.08	49.62	51.16	52.71	54.25	55.79	57.34	58.88	60.42	61.97
-1.25	50.75	52.30	53.84	55.38	56.93	58.47	60.01	61.56	63.10	64.64
-0.75	53.43	54.97	56.51	58.06	59.60	61.14	62.. 69	64.23	65.77	67.32
-0.25	56.10	57.65	59.19	60.73	62.28	63.82	65.36	66.90	68.45	69.99
0.25	58.78	60.32	61.86	63.41	64.95	66.49	68.04	69.58	71.12	72.67
0.75	61.45	63.00	64.54	66.08	67.63	69.17	70.71	72.25	73.80	75.34
1.25	64.13	65.67	67.21	68.76	70.30	71.84	73.39	74.93	76.47	78.02

Female — Second Factor Score

First Factor Score	-2.25	-1.75	-1.25	-0.75	-0.25	0.25	0.75	1.25	1.75	2.25
-2.25	47.16	49.21	51.25	53.30	55.34	57.39	59.43	61.48	63.52	65.57
-1.75	50.07	52.11	54.16	56.20	58.25	60.29	62.34	64.38	66.43	68.47
-1.25	52.97	55.01	57.06	59.10	61.15	63.19	65.24	67.29	69.33	71.38
-0.75	55.87	57.92	59.96	62.01	64.05	66.10	68.14	70.19	72.23	74.28
-0.25	58.78	60.82	62.87	64.91	66.96	69.00	71.05	73.09	75.14	77.18
0.25	61.68	63.72	65.77	67.81	69.86	71.90	73.95	75.99	78.04	80.08
0.75	64.58	66.63	68.67	70.72	72.76	74.81	76.85	78.90	80.94	82.99
1.25	67.48	69.53	71.58	73.62	75.67	77.71	79.76	81.80	83.85	85.89

Prediction Equations:

Male: $y = 64.3844 + 5.3496 * F1 + 3.0862 * F2$

Female: $y = 69.4301 + 5.8061 * F1 + 4.0902 * F2$.

correlation (value of 1) is found for 5 male and 4 female tables. The largest absolute value of the intercept is .02686 for the male data and .04040 for the female data and ninety per cent of the tables produced intercepts less than .01 in absolute value. One hundred and nine male and 101 female tables show slopes in the interval .99 to 1.01. Overall goodness of fit measured by these coefficients is excellent. Accordingly, we may expect model estimates of the survivorship function *l(x)* to match the actual values quite well. A visual demonstratiaon of this fit is provided in Figure 2 for four countries.

An alternative measure of the reproducibility of a life table can be obtained by comparing observed and model life expectancies. Rather than computing model life expectancies from the *Ml(x)* values and comparing them with the observed values of e(0), we chose to regress the observed *e(0)* values on the two factor scores instead. It is gratifying to note that the squared multiple correlations of these regressions were .998 for male and .994 for female tables. Such large correlations imply that the two factor scores are excellent predictors of life expectancy. In Table 5, we show the distribution of *e(0)* for selected combinations of the two factor scores. As may be seen from Table 5, as well as from the regression equations, identical life expectancy can result from many combinations of the two factor scores. In other words, different distributions of life table functions showing varying patterns of mortality may have the same expectation of life at birth.

Note the difference between the ranges of the two factor scores in Table 5. The reasons why the first factor score must have a small positive value for its upper limit is discussed in the next section.

The Test of Consistency

Since the survivorship function *l(x)* is a nonincreasing function of age *x*, the corresponding *Ml(x)* values generated by the principal components model should exhibit the same pattern. Remembering that the principal components analysis, unless otherwise directed, converts variables to their standardized forms, the fundamental equation relating the variables with the factors can be written as

$$\frac{l(i) - m(i)}{s(i)} = f(1, i)\, F(1) + f(2, i)\, F(2) \tag{4}$$

In (4), $l(i)$ corresponds to the age of the ith category; $m(i)$ and $s(i)$ are the mean and standard deviation of $l(i)$; the loadings of the two factors are $f(1, i)$ and $f(2, i)$; and the factor scores are $F(1)$ and $F(2)$. The consistency condition

$$l(i) - l(i + 1) \geq 0 \tag{5}$$

translates into the inequality

$$\begin{aligned} m(i) - m(i+1) \quad &+ F(1)\,[s(i) f(1, i) - s(i+1) f(1, i+1)] \\ &+ F(2)\,[s(i) f(2, i) - s(i+1) f(2, i+1)] \geq 0 \end{aligned} \tag{6}$$

Writing

$$g(1, i) = [s(i) f(1, i) - s(i+1) f(1, i + 1)] \tag{7}$$

and

$$g(2, i) = [s(i)f(2, i) - s(i+1)f(2, i+1)] \tag{8}$$

the lower and upper limit of $F(1)$ for given values of $F(2)$ and i can be obtained from (6) as

$$F(1) \ > \text{or} < \ c(i) + d(i)\, F(2) \tag{9}$$

according as the coefficient of $F(1)$, i.e., $g(1, i)$ is $>$ or < 0, where

$$c(i) = -\frac{m(i) - m(i+1)}{g(1, i)} \tag{10}$$

and

$$d(i) = -\frac{g(2, i)}{g(1, i)} \tag{11}$$

Since (5) must hold for all i, the lower and upper bound of $F(1)$ are finally determined from

$$F(1) \ > \ \mathit{Maximum}\ [c(i)F(2) + di\ F(2)] \tag{12}$$

and

$$F(1) < \mathit{Minimum}\ [c(i) + d(i)F(2)] \tag{13}$$

of the subsets of $F(1)$ generated by those values of i for which $g(1, i)\ >$ and < 0 respectively. The values of $m(i)$, $s(i)$, $f(1, i)$, $f(2, i)$, $c(i)$ and $d(i)$ are presented in Table 6. Equations (12) and (13) must also hold for age 0, where the mean is equal to 1, the standard deviation is equal to zero, and

$$g(1, 0) = -s(1)f(1, 1) \tag{14}$$

and

$$g(2, 0) = -s(1)f(2, 1) \tag{15}$$

TABLE 6: MEANS AND STANDARD DEVIATIONS OF l (x), FACTOR LODINGS, *c(x), d(x)* BY SEX

			Lodings			
Age x	Mean	Standard Deviation	Factor 1	Factor 2	c(x)	d(x)
			Male — l(x)			
0	1.0	0.0			1.54993	-.12823
1	.95513	.03074	. 94176	.12076	.80612	-.24072
5	.93768	.05206	. 97189	.17140	.94142	-.41934
10	.93212	.05799	. 97435	.19658	1.34927	-.39276
15	.92826	.06100	. 97317	.20530	2.02818	-.48781
20	.92162	.06450	. 97112	.21892	2.23594	-.49186
25	.91264	.06880	. 96880	.23395	1.88258	-.47705
30	.90285	.07450	. 96448	.24935	2.27262	-.49951
35	.89136	.08012	. 95993	.26338	2.78432	-.69198
40	.87650	.08634	. 95259	.28718	4.87876	-1.11155
45	.85585	.09192	. 94081	.32093	10.10952	-1.88165
50	.82595	.09679	. 92403	.36228	26.55592	-4..50998

Table 6 *(contd. on p.* 124)

Table 6 (contd. from p. 123)

Ladings Age x	Mean	Standard	Factor 1	Factor	c (x)	d (x)
55	.78159	.10137	.89876	.42023	-	2.79696
60	.71628	.10325	.85056	.50162	-0.03553	.66020
65	.62546	.09929	.78325	.58846	-6.95662	.43525
70	.50727	.09137	.66250	.72040	-6.52884	-.00892
75	.37009	.07790	.51050	.84256	-5.44887	-.23525
80	.23432	.06253	.23750	.95592	-7.01047	-1.02345
85	.11566	.04567	-.04544	.92951		
			Female-1(x)			
0	1.0	0.0			.52785	-.32891
1	.96186	.02765	.90283	.29695	.80351	-.37730
5	.94476	.04952	.93386	.32795	.86952	-.50156
10	.94011	.05523	.93414	.34261	.00779	-.48729
15	.93706	.05854	.93302	.34843	.20204	-.51244
20	.93281	.06246	.93107	.35557	.22714	-.52215
25	.92691	.06786	.92783	.36427	.36668	-.55504
30	.91986	.07376	.92355	.37395	.67824	-.60550
35	.91092	.07994	.91879	.38539	2.15260	-.66613
40	.89923	.08631	.91390	.39886	2.94396	-.80207
45	.88386	.09282	.90605	.41600	4.31545	-1.07980
50	.86322	.09949	.89338	.44002	6.71213	-1.19887
55	.83383	.10597	.88007	.46265	15.09233	-2.86504
60	.79288	.11272	.85144	.50391	76.79210	-10.14789
65	.73106	.11825	.81843	.54943	-	1.51547
70	.64299	.11942	.75089	.63425	-8.07646	.25164
75	.52186	.11077	.67413	.71785	-5.07913	-.03664
80	.37444	.09147	.49906	.85769	-4.87633	-.37621
85	.21309	.06953	.18065	.94930		

It is important to mention that, although *the factors produced by the principal com- - ponents are technically uncorr elated, the consistency criterion of the survivorship function as shown in* (12) *and* (13) *introduces a constraint into that relationship. From a practical point of view, the contribution of such a constraint is relatively minor in nature.*

Most of the second factor scores have absolute values under 2.25. For male and female data, only 5 second factor scores lie outside the interval (-2.25, 2.25). The picture is the same for the first factor score with respect only to the lower limit of the interval but not with the upper. Thus, while only 4 male and 6 female first factor scores were less than -2.25, none was higher than 1.12. In fact, the first factor score can not, for all practical purposes, assume values much greater than 1.2 without violating the inequalities (12) and (13). A demonstration of this feature of the joint variation of the two factor scores has been provided in Table 7. There the permissible ranges of the first factor scores corresponding to selected of the subsets of F(2) generated by those values of *i* for which *g(2, i)* > and < 0 respectively.

values of the second in the interval (-2.25 to 2.25), derived from the inequalities (12) and (13), have been tabulated. Similar ranges can be obtained for the second factor scores for selected values of the first from

$$F(2) > Maximum \left[\frac{c(i)}{d(i)} + \frac{F(1)}{d(i)} \right] \tag{16}$$

and

$$F(2) < Minimum \left[\frac{c(i)}{d(i)} + \frac{F(1)}{d(i)} \right] \tag{17}$$

TABLE 7: BOUNDARIES OF FIRST SCORE FOR SELECTED SECOND FACTOR SCORES BY SEX

Factor 1 Boundaries

Factor 2	*Male*		*Female*	
	Lower Boundary	*Upper Boundary*	*Lower Boundary*	*Upper Boundary*
-2.25	-4.70772	1.34775	-4.02986	1.52785
-1.75	-5.03719	1.22739	-4.21797	1.46378
-1.2S	-5.15481	1.10703	-4.40607	1.27513
-.75	-5.27244	.98666	-4.59417	1.08649
-.25	-5.39006	.86630	-4.78228	.89784
.25	-5.50769	.74594	-4.9703*	.70919
.75	-5.62531	.62558	-5.10666	.49334
1.25	-5.74294	.41725	-5.12498	.24256
1.75	-5.86056	.20758	-5.14329	-.00822
2.25	-5.97730	-.00209	-5.16161	-.25900

Observe that both the lower and the upper boundaries of the first factor score decline as the second factor score increases shifting, but not significantly constricting, the range of variation of the first factor score in the process. The figures also show, as we have noted earlier, that while both factor scores can simultaneously be small, they cannot both be large at the same time, a fact also confirmed by the data. For the entire range of variation of the second factor scores, however, the observed first factor scores fall outside the prescribed ranges for four male and six female tables. The discrepancies, relatively minor in nature, can be circumvented m such cases by adjusting the factor scores from equations (12) through (17) to derive the model values of life table functions.

Concluding Remarks

Having established the adequacy of two factors to reproduce the life table survivorship function, we now offer some suggestions for their usefulness. Model life tables can be generated (see equation 1) by simply choosing appropriate values of the two factor scores within their permissible ranges of variation (equations 12 and 13). These tables may provide an alternative to the published life tables. Incidentally, we may add that the adequacy of the published tables was never tested through comparisons with the life tables selected to generate them, and we doubt such tests are possible. As we have seen earlier, the method of principal component analysis provides a way of determining the extent to which an observed life table can be reproduced.

Although model life tables can be constructed by selecting appropriate combinations of the two factor scores, the application of principal component analysis is not limited to the construction of life tables alone as the method can be extended to the derivation of factor scores for any life table (see equation 3). Since the factor scores contain most of the information present in a life table, it follows that the factor scores for a country at various points in time can be interpreted as indicators of the manner in which the two dimensions of mortality have changed during that period. Extrapolations of the trends of these values can be made for the construction of projected life tables.

As noted earlier, an infinite number of life tables can be constructed by combining the two factor scores in all possible ways, providing the boundary conditions are met. In a sense, these two factors, youth and elderly, can be viewed as two independent dimensions of mortality paralleling the dimensions of region and life expectancy of Coale and Demeny's regional model tables. However, region permits only a finite number of alternatives, thereby limiting the number of possible choices, whereas the number is virtually limitless with the factors. Further, the regional identity of a life table cannot be deduced from the principles on which the derivations of the model tables are based. In contrast, principal component analysis net only has the ability to capture the finer details of a life table, it also has a built-in method of comparing the model estimates with the actual values.

References

Brass, W., *el al.,* 1968, *The Demography of Tropical Africa.* Princeton University Press, Princeton.

Coale, A. and P. Demeny, 1966, *Regional Model Life Tables and Stable Populations.* Princeton University Press, Princeton. ____ and ____, 1983, *Regional Model Life Tables and Stable Populations.* Second edition. Academic Press, New York.

Keyfitz, N., 1971, *Population.* W. H. Freeman and Company, San Francisco. Leadermann, S'. and J. Breas, 1959, Les dimensions de la mortalite. *Population* 14(4): 637-682. United Nations, 1956, *Manual III: Methods for Population Projections by Sex and Age,* Sales No. 56, XIQ, 3.

APPENDIX-I

Modified Values of 1(80) by Country and Sex

Country	*Sex*	*Values of (1)80*	
		Recorded	*Modified**
Tunisia 1960	female	38,208	38,855
Greenland 1960	female	26,365	30,319
Kuwait 1966	male	17,568	23,805
	female	24,472	31,898
Portugal 1966-68	male	18,165	24,611
	female	29,043	37,049

* The figures are simple averages of 1(75) and 1(85).