WORKING PAPER

A SIMPLE MODEL FOR THE STATISTICAL ANALYSIS OF LARGE ARRAYS OF MORTALITY DATA: RECTANGULAR vs. DIAGONAL STRUCTURE

John R. Wilmoth
Graziella Caselli

June 1987
WP-87-58

IIASA
International Institute
for Applied Systems Analysis

A Simple Model for the Statistical Analysis of Large Arrays of Mortality Data: Rectangular vs. Diagonal Structure

John R. Wilmoth
Graziella Caselli

June 1987
WP-87-58

INTERNATIONAL INSTITUTE FOR APPLIED SYSTEMS ANALYSIS
A-2361 Laxenburg, Austria

Foreword

The rapidly accumulating data on mortality offers an opportunity for study of the modes and circumstances under which improvement has been occurring in various countries and regions. Do the age-specific rates tend to decline by constant amounts over time, or in constant ratios? Has improvement been continuous and similar from year to year or has it been intermittent? In the latter case has it occurred from a certain point of time onwards, or from a certain cohort onward? IIASA has helped answer such qu estions through various innovations in demographic method, including the use of contour maps introduced by Vaupel, Gambill, Yashin, and Bernstein (1985).

The present paper, tackling some of the issues analytically rather than graphically, fits a flexible model that provides for additive and multiplicative components. Mr. Wilmoth and Professor Caselli have fitted their model to the postwar male mortality series for France.

Among other conclusions, Wilmoth and Caselli isolate certain positive diagonal effects, which is to say that certain cohorts (for example those of 1896-1909 and 1925-35) had unexpectedly high mortality compared with those just older or just younger. Could this have been due to their passing through difficult war years in childhood and early adolescence? It will require other kinds of evidence to say for sure. But the present paper is an innovative way of discerning the facts.

Nathan Keyfitz
Leader, Population Program

Abstract

A simple descriptive model is proposed for the analysis of large, non-additive mortality arrays. Similar in form to additive-plus-multiplicative models discussed by other authors, the model goes one step further by introducing a diagonal term. An exemplary application of the model to French male post-War mortality data demonstrates three important characteristics of the data being analyzed: 1) the structure of the data matrix is largely additive; 2) some rectangular non-additivity exists, implying that mortality has declined with varying speed at different ages or, equivalently, that the shape of the age-curve of mortality has changed over time; and 3) residual non-additive diagonal structure exists, indicating that some "peculiar" cohorts have had mortality experiences which deviate by as much as 2 or 3% from levels which would be expected considering only the age and period of death.

A Simple Model for the Statistical Analysis of Large Arrays of Mortality Data: Rectangular vs. Diagonal Structure

John R. Wilmoth, Graziella Caselli***

I. Introduction

Large arrays of age-specific mortality data are now available for various European countries (for example, Vallin 1973). The size of these matrices can be considerable: over fifty years of data by single-year age groups up to advanced ages. The result is often a matrix of over 5000 data points, which presents problems of statistical analysis which are not easily resolved.

The motivation for their resolution is nevertheless easily demonstrated and, furthermore, well-known. The evolution of mortality in developed countries since the Second World War has been marked by significant declines at nearly all ages in the probability of death from one birthday to the next. The speed of the decline has varied across the age range, and certain age groups have seen periods of decline, followed by stability or even slight increase (Caselli and Egidi, 1981).

Of particular interest have been theoretical analyses of the causal factors which have produced the observed evolution and hence the structure of the data matrix. Traditionally, one has attempted to classify these factors into three dimensions: events related to the age at death, the period of death, or the cohort of birth (see Hobcraft et. al., 1982). An inherent difficulty with this approach derives from the fact that one is trying to analyze two-dimensional data in terms of a three-dimensional factor space.

Our purpose in this work is not to resolve, from a theoretical standpoint, the deeper issues of causal statistical analysis of such a data matrix; nor do we intend to undertake an extensive discussion of the demographic and biological theory of mortality change. Rather, we propose a simple model for describing the structure of a large array of mortality data. The model provides concise pictures of the evolution of mortality at the various

*Office of Population Research, Princeton University, 21 Prospect Avenue, Princeton, New Jersey 08544, USA
**Dipartimento di Scienze Demografiche, Universita Degli Studi di Roma "La Sapienza", Via Nomentana 41, 00161 Rome, Italy

ages and of the age-pattern of mortality during different periods, and yields interesting results on some peculiarities of the mortality experience of certain cohorts.

We identify two types of structure within the data matrix: rectangular and diagonal. Rectangular structure is that which can be explained in terms of row (age) or column (period) membership, while diagonal structure consists of those residual elements which seem consistent over diagonals (cohorts) but not over rows or columns. Hence, we give preferential treatment to the rectangular structure and treat diagonal structure as a residual element. In short, we attempt to describe the structure of the data matrix using simple additive and multiplicative terms over the rows and columns of the array (rectangular structure), and then attempt further to decompose the residuals by introducing a diagonal effect (diagonal structure).

In general, we find that a single rectangular element, consisting of a simple additive term, is consistently capable of describing an extremely large portion of the total variance in the matrix. Nevertheless, it appears that other rectangular elements (such as multiplicative terms) and the diagonal effects provide a description of the non-additive nature of the data matrix which seems interesting for a theoretical study of mortality.

We provide an exemplary application of the model using French male mortality data during the post-War period. The model, in addition to describing the average pattern of mortality across age and its general evolution across time, highlights ages at which mortality change has been atypical, periods in which the age-curve of mortality differs from the average pattern, and cohorts whose mortality experience has been peculiar.

II. Model Development

Anyone familiar with series of mortality data would not be surprised by the fact that often a simple additive fit proves to be sufficient for describing the structure of the data matrix, especially if our only criterion for a "good fit" is in terms of the percent of total variance explained. Our experience indicates that, after a simple transformation of the yearly death probabilities, a traditional combination of additive row and column effects typically accounts for over 99% of the total variance.

In spite of this near-additivity (seemingly "near-perfect additivity"), we are motivated for theoretical reasons to describe the matrix further in terms of certain non-additive elements. We identify two types of residual non-additivity (rectangular and diagonal) which are characterized by the addition of multiplicative and diagonal terms. The former serve as correction terms for ages or years which deviate from the predominant additive structure, while the latter describe the experience of cohorts whose mortality levels may be considered relatively higher or lower than expected.

We define a family of models and suggest in the next section how to choose the appropriate model within the family based on the data being analyzed. It has proven necessary to transform the observed death probabilities before application of the model. A transformation which approximates the logarithm of the mid-year force of mortality can be easily calculated. Thus, defining

$$f_{ij} = \log\left(\frac{q_{ij}}{1 - \frac{1}{2}\, q_{ij}}\right) \quad , \tag{1}$$

we may specify the model in terms of f_{ij}, where i and j are the row (age) and column (period) indices of the matrix. Other transformations, such as simple logs or logits, are also possible. These alternatives have been tried as well, but with only minor changes in the results.

The family of models which interests us has the following form:

$$f_{ij} = a_i + b_j + c_i d_j + g_i h_j + \cdots + o_k + e_{ij} \quad , \tag{2}$$

where $k = j - i$. The first two terms are the additive part of the model, which consists of row effects, a's, and column effects, b's. The next two terms are multiplicative adjustments for non-additivity and are optional additions to the model (although at least one multiplicative term has appeared desirable in practice almost always). The diagonal effects, o's, are calculated from the residuals of the additive and multiplicative terms, leaving only random errors, e's. The necessary constraints for the statistical identification of the model are given in the Appendix.

III. Application to French Male Mortality Data

To demonstrate an application of the model, we chose to consider post-War French male mortality for the period 1946-1981 and single-year ages 0-89. The observed probabilities of death are age- and cohort-specific and correspond to the "Tables de génération" in Vallin (1973, 1984). That is, q_{ij} gives the observed probability of death between ages i and $i+1$ for the cohort $j-i$, whose members attain age i sometime during the calendar year j. In our case, we have $i = 0,...,89$ and $j = 1946,...,1981$. The result of using age-cohort data in the analysis is that we know with precision the age and cohort of the probability, which is in fact measured over two adjacent period: when we refer to year j, the death may have occurred in year j or $j+1$.

The first step in the application of the model is to identify the proper model within the family of models for the dataset at hand. We propose two criteria, both somewhat arbitrary, which may be used to determine the correct number of multiplicative terms. Fortunately, both criteria seem to lead us consistently to the same choice.

The first criterion (and the inferior one, in our opinion) is the "variance explained" criterion. Assuming that an additive term has already been fit to the data, we may successively add multiplicative terms in accordance with the model given above. Each of these is guaranteed to explain some portion of the remaining variance. Our experience indicates that, when there remains significant non-additive structure in the residuals after fitting an additive or a multiplicative term, the addition of a (perhaps further) multiplicative term tends to account for over 50% of the variance remaining in the residuals. This situation may be observed for the first or second multiplicative term, but rarely have we seen the need for more than two.

For example, Table 1 gives the variance breakdown for the French male post-War dataset described above. The first column gives the percent of residual variance explained by each term. The fact that the additive term explains already 99.5% of the total variance confirms what was said earlier about the essentially additive structure of the matrix. Nevertheless, the fact that the first two multiplicative terms each explain, in their turn, over 60% of the remaining variance indicates that there are systematic deviations from additivity which may merit our attention. We then observe a notable falloff in explanatory power for the third multiplicative term, indicating that there is probably no clear rectangular structure left in the residuals after the second multiplicative term and that we should hence choose a model with one additive and two multiplicative terms. In fact, the improper addition of a third multiplicative term would distort the estimation of the diagonal effects, so it is important to stop at two.

Table 1. Percent of variance explained for various sub-models. French male mortality data, ages 0-89, years 1946-1981.

Final term:	% Residual variance	% Total variance	% Cumulative variance
Additive	99.496	99.496	99.496
Multiplicative 1	65.444	0.330	99.826
Multiplicative 2	62.528	0.109	99.935
Multiplicative 3	23.026	0.015	99.950

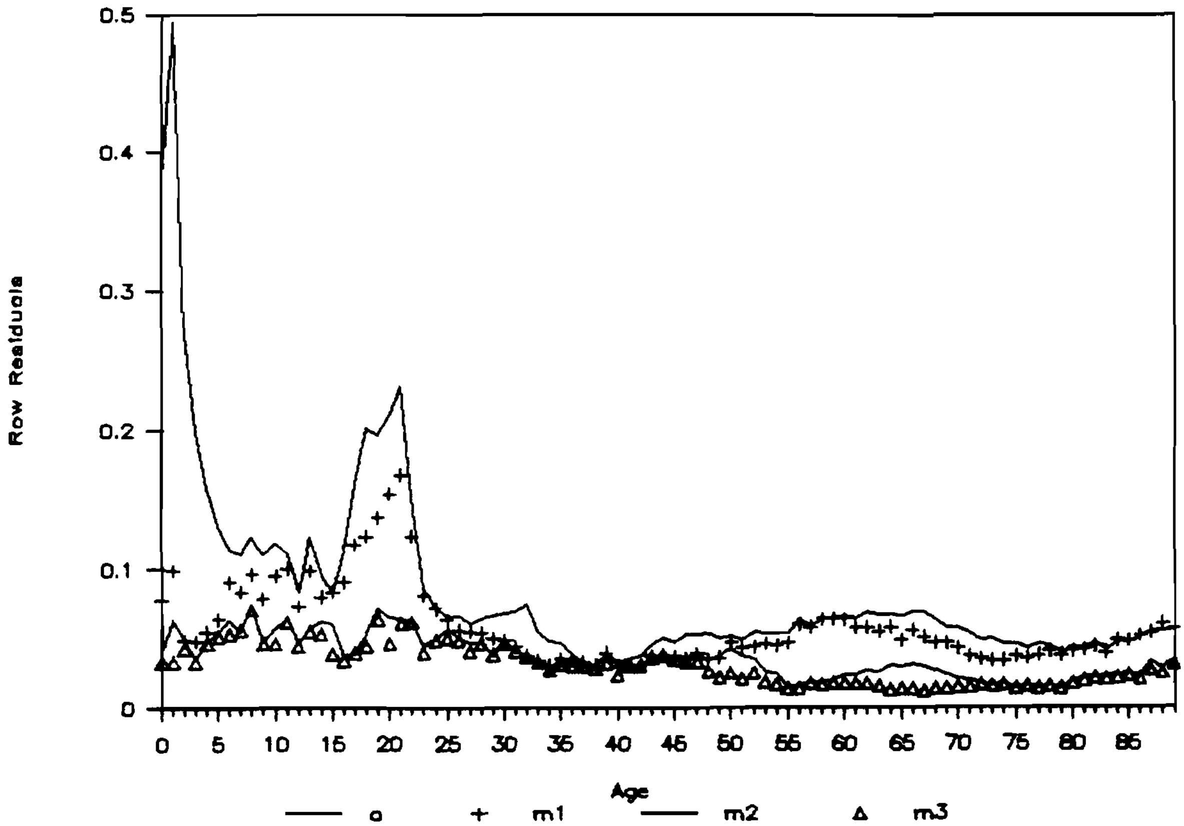
Another criterion for the choice of the model lends insight into the function of the multiplicative terms within the model, and we call it the "balanced residual" criterion. In short, it relies on the fact that a good descriptive model of the structure of the data matrix should show residuals which are of similar absolute magnitude for all rows and columns of the matrix. Calculating the mean absolute residual for every row and column, we add multiplicative terms until the model describes the situation in every row and column with similar accuracy. Figures 1a and 1b show the changes in the mean absolute residuals for the rows and columns of the matrix as we add successive terms. After the additive term, the model fits uniformly well over much of the age range with the notable exceptions of the first 22 or 23 years. The first multiplicative term seems to explain the non-additivity of the earliest years, while the second one picks up the non-additive structure in the late teens and early 20's. The years of time which seem the most heavily non-additive are the earliest (1946-1950) and the latest (1979-1981), but these inconsistencies likewise disappear with the addition of two multiplicative terms.

We may also consider the balanced residual criterion by examining the distribution of the mean absolute residuals for either the rows or the columns. For example, Figures 2a-c show stem-and-leaf displays (Tukey, 1977) of the mean absolute row residuals at successive steps in the fitting of the model. After the additive term (Figure 2a), the majority of the mean absolute residuals are concentrated around 0.03 to 0.07; nevertheless, there are several outliers above 0.10, and some extreme values which are listed below the main display. These extremely long tails seem to be controlled somewhat by the addition of one multiplicative term (Figure 2b), but it requires two such terms in order to eliminate the long upper-tail of the distribution (Figure 2c).

Hence, both criteria seem to indicate the need for two multiplicative terms in order to account for the rectangular structure of the matrix. Clearly though, our discussion has avoided the question of how to obtain the estimates of the row and column effects at each step of the procedure. Likewise, we have given no consideration to the question of op-

Figure 1. Residuals from various sub-models: additive plus zero, one, two, or three multiplicative terms.

a) Row effects by age



b) Column effects by period

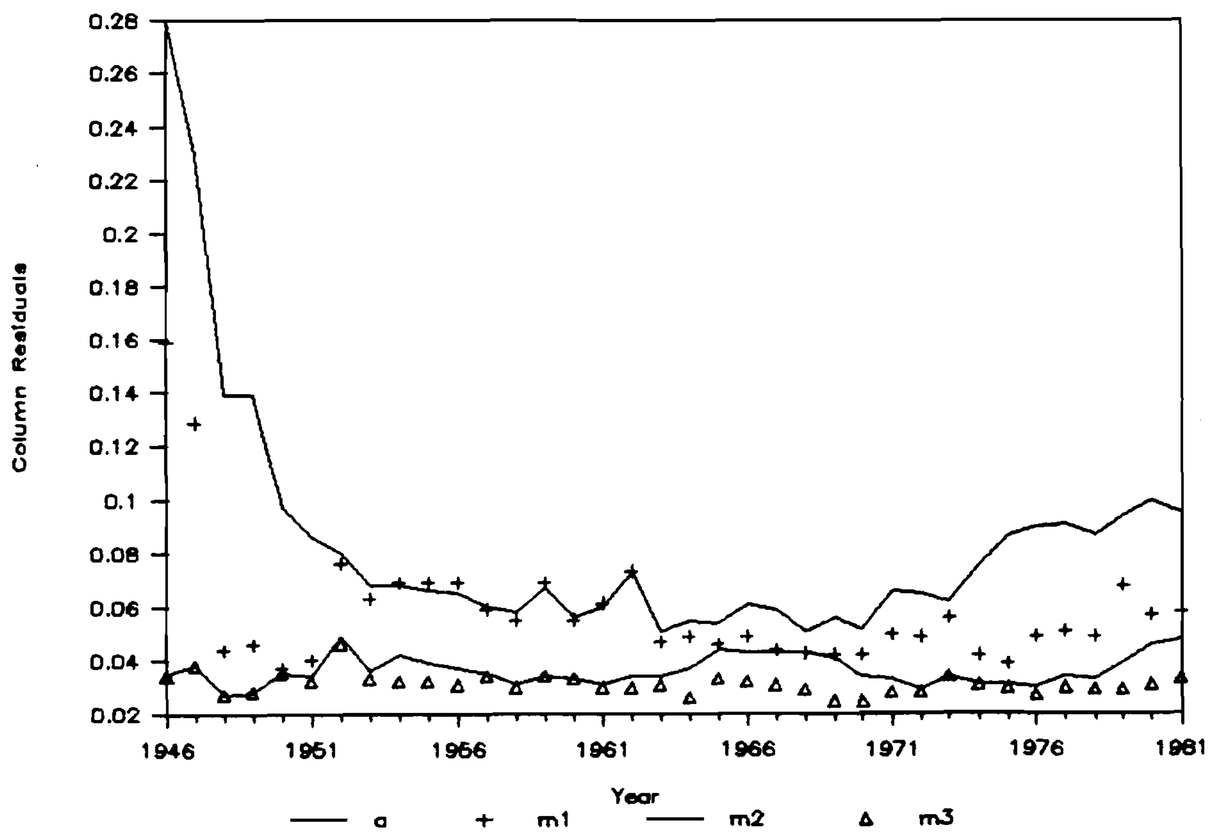


Figure 2. Stem-and-leaf display of row residuals for sub-models.

a) Simple additive model

Depths (n=90)		Residuals (unit=0.01)
5	03	13367
25	04	00012345566677788899
45	05	0011233444445556688 9
45	06	0344444566667789999
26	07	25
24	08	446
21	09	5
	10	
20	11	001348
14	12	33
12	13	1
	14	
11	15	18
9	16	0
	17	
	18	
8	19	69
6	20	19
	21	
	22	
4	23	1
	HI	.269, .380, .494

b) Additive-plus-one model

Depths (n=90)		Residuals (unit=0.01)
25	03	0000134445555566777788889
(23)	04	00112334445567777889999
42	05	01245556678888
28	06	0134444
21	07	2389
17	08	0144
13	09	115799
7	10	0
6	11	7
5	12	33
3	13	7
	14	
2	15	4
1	16	7

c) Additive-plus-two model

Depths (n=90)		Residuals (unit=0.01)
13	01	4445666677889
31	02	001223344445566779
(28)	03	0000001223344455566777888889
31	04	0011257778889
18	05	00455788
10	06	02233445
2	07	25

timality in our choice of a fitting method. Our next task then is to discuss briefly the means of fitting the additive, multiplicative, and diagonal terms of the model. As work on robust and resistant fitting techniques for this model is still in its early stages, our current fits rely on very classical methods. Furthermore, we discuss here only some very general considerations on the question of fitting, with a more technical discussion being relegated to the Appendix.

A traditional criterion for fitting a model to data is that of minimizing the residual variance. If we ignore the diagonal term, it is quite simple to obtain least squares estimates of the additive and multiplicative terms by fitting each additional term to the residuals of previous sub-model. This procedure not only minimizes the residual variance at each step, but also yields coefficients which minimize the residual variance for the complete model (excepting the diagonal term).

The addition of a diagonal term, however, lends a complication to the fitting procedure which has not yet been thoroughly resolved. There would seem to be two choices: either make some kind of optimal (for now, least squares) fit of the diagonal effects to the residuals of the other terms, or find an iterative procedure which minimizes the residual variance of the complete model through the simultaneous consideration of all terms. Obviously, the former approach is much simpler and is the one which we have adopted in this paper, although the latter possesses a certain theoretical appeal.

A theoretically correct iterative solution to the second fitting approach involves adjusting the original data for the diagonal effect, then recalculating the additive and multiplicative terms on the "corrected" data, then recalculating the diagonal effect, recorrecting the original data, recalculating the additive and multiplicative terms, and so on until convergence. In practice, this procedure has proved to be unstable when applied to real data, although when applied to simulated data it appears little different from (if not slightly inferior to) the simpler approach described above. For this reason, and since the iterative procedure proves to be quite slow to converge, we have chosen for the time being to fit the diagonal effects to the residuals of the additive-multiplicative model, as described in the Appendix. The interpretation of the diagonal effects is thus that they are residual elements which remain even after we have accounted for the dominant rectangular structure of the data array.

IV. Presentation and Interpretation of Results

The rectangular structure of the matrix, as described by the additive and multiplicative terms, recalls elements of the characteristics and evolution of post-War mortality which are already well-known. Figure 3a, for example, shows the additive row effects, which represent the average age-pattern of mortality over the period considered, while Figure 3b depicts the additive column effects, which demonstrate the average pattern of decline from 1946 to 1981.

Rectangular deviations from additivity are characterized by the combination of the two multiplicative terms. The individual coefficients of these two terms are often difficult to interpret if presented separately (as we have done above for the additive coefficients). Instead, in order to show the multiplicative interaction of the coefficients, we present in Figure 4 a two-level contour map of the combined effect of the two terms. The axes of the map consist of the column (x-axis) and row (y-axis) indices. The level of the map at each point is determined by the values

$$c_i d_j + g_i h_j \quad . \tag{3}$$

At those points (j,i) where the combined multiplicative effects are positive, the map is colored dark; where negative, the map is light.

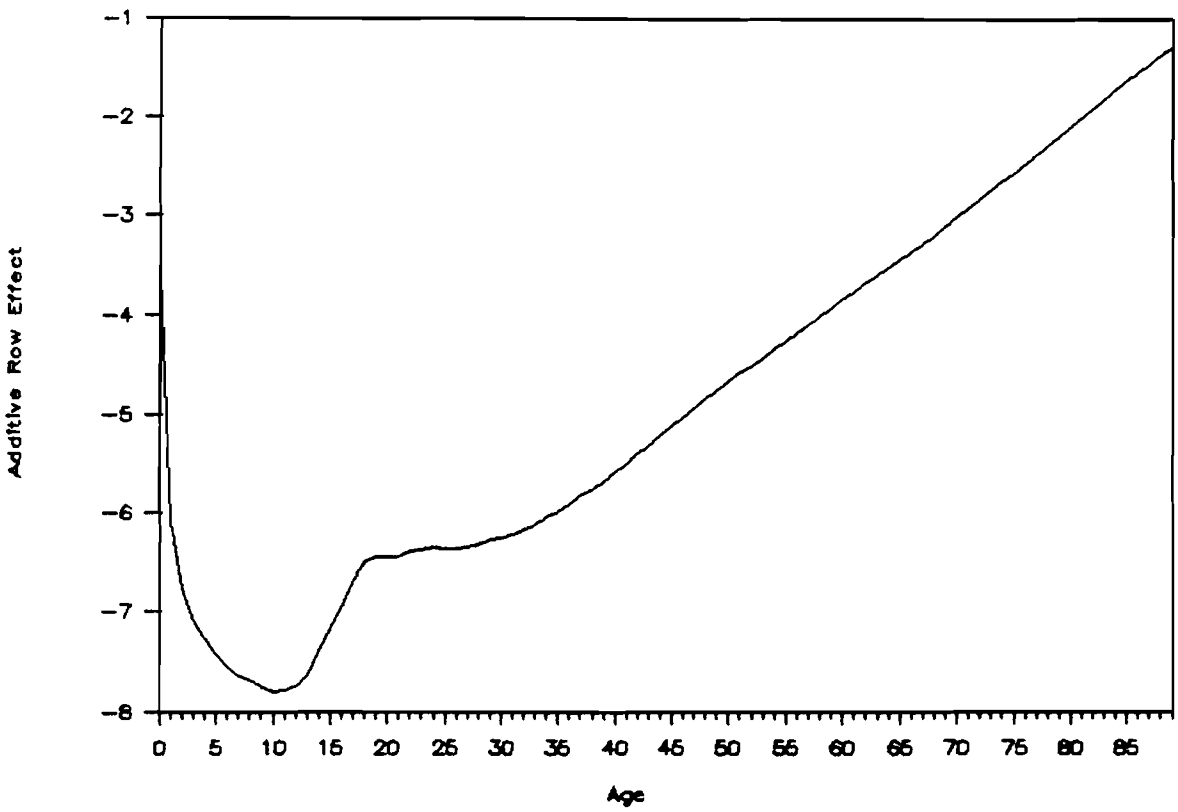
The dark (light) points on the map thus isolate ages and years at which, according to the multiplicative terms, mortality was actually higher (lower) than that predicted by a simple additive model. Of course, the multiplicative terms do not capture all elements of non-additivity in the matrix, but they do indicate a general tendency for levels of mortality which are higher or lower than foreseen by the additive model.

There are two equivalent ways in which we may interpret the positive and negative portions of the multiplicative terms. As we have said, the multiplicative portion of the model corrects partially for non-additivity in the observed data. This non-additivity may be observed by studying the rows or the columns of the matrix.

In the first case (non-additive rows), we consider the speed of mortality decline at certain ages in comparison with the average rate of decrease. For example, at the youngest ages we observe positive combined multiplicative effects in the early years and negative ones in the later years. This illustrates a well-known characteristic of the evolution of infant and childhood mortality, which has consistently experienced stronger than average declines. On the contrary, slower than average declines have generally been observed in the age groups above 40 years. Slightly more complicated has been the change in mortality in the age group 15-25, which also saw relatively fast progress in the early years, followed by relative (and sometimes absolute) increases since the late 1960's (undoubtedly

Figure 3. Coefficients from simple additive model. French male mortality.

a) Row effects by age



b) Column effects by period

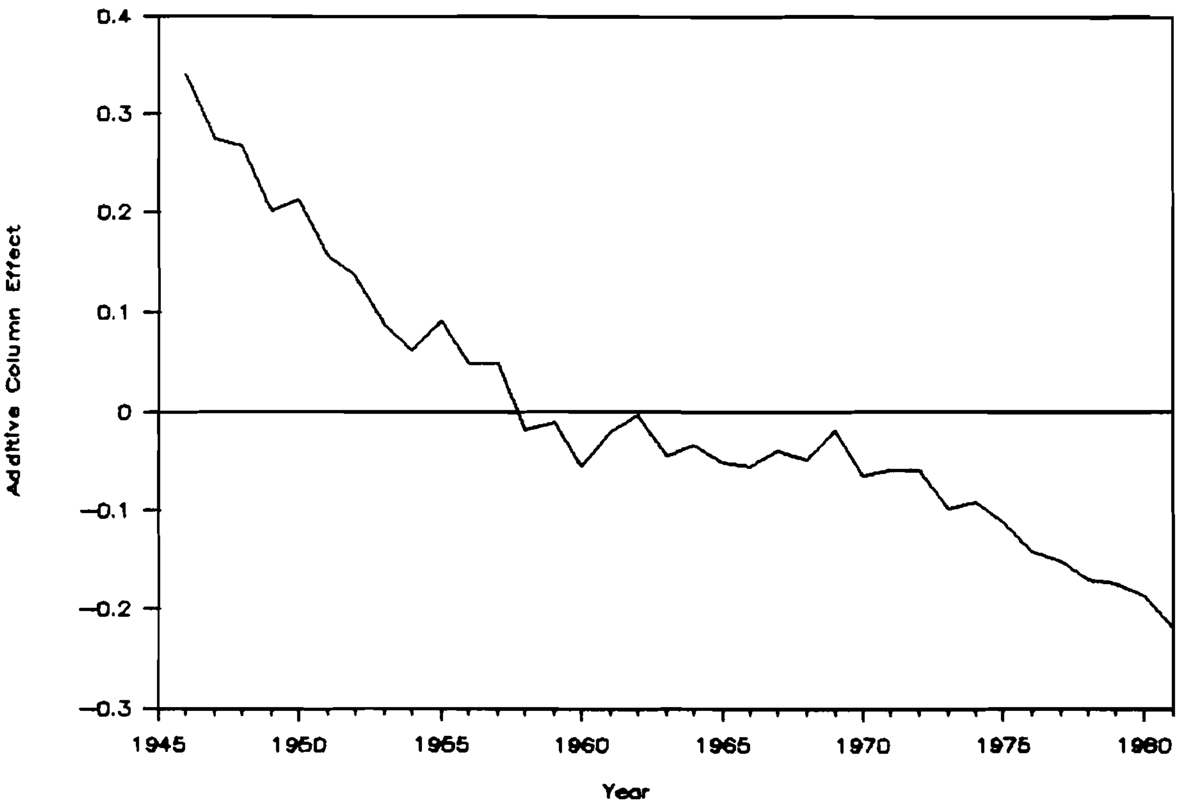
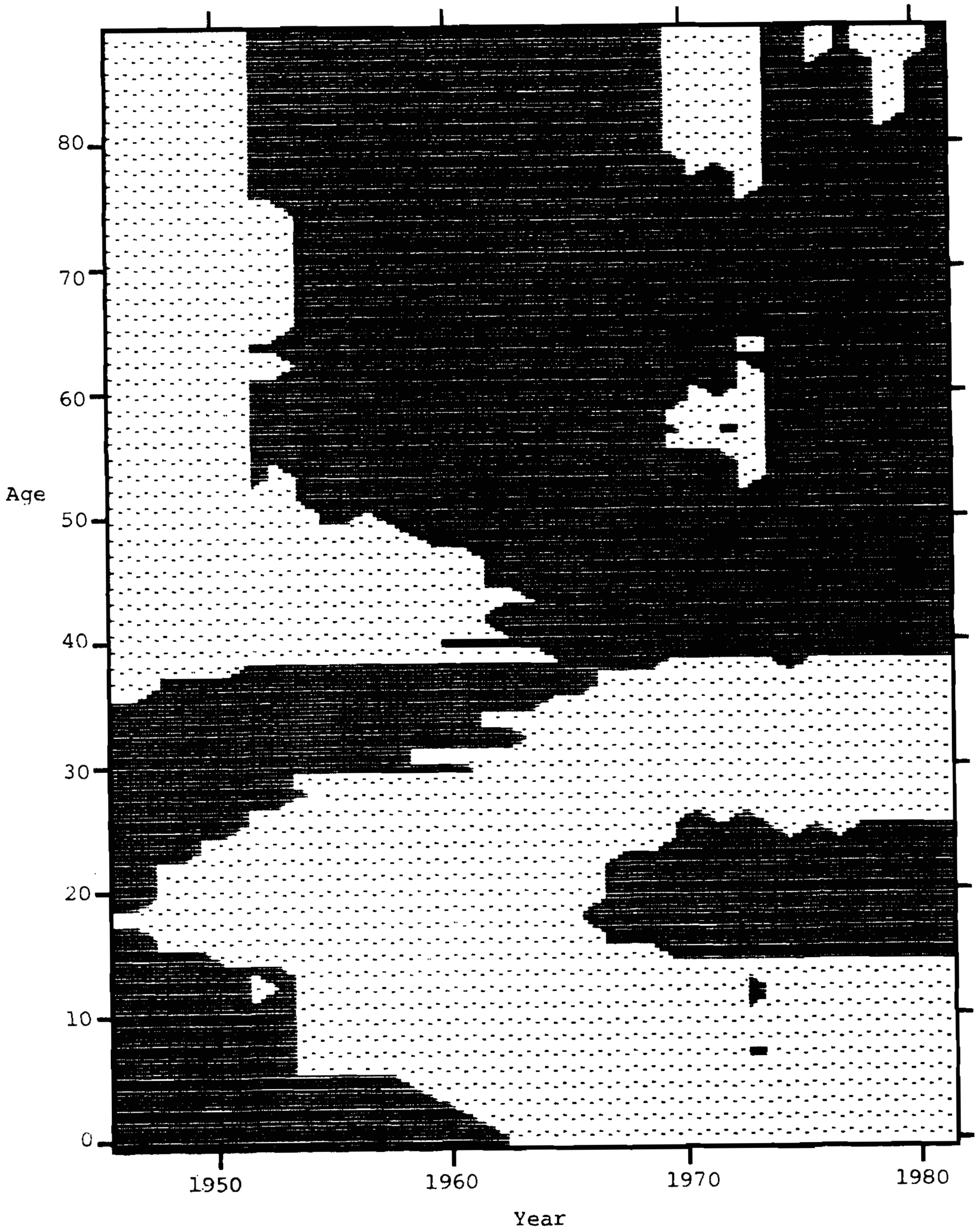


Figure 4. Two-level contour map of combined multiplicative terms (dark = positive, light = negative). Non-additivity in French male mortality.



due to increasing accident mortality).

In the second case (non-additive columns), we examine the differing shapes of the various age-curves of mortality. For example, in the early years (say 1946-1950) we observe positive combined multiplicative effects at ages 0-35 or 36, with the exception of a few ages around age 18 or 19. After age 40, the multiplicative terms have a combined negative effect. Nearly the opposite seems to be true in the period after 1975. This indicates that the observed age-curve of mortality for 1946-1950 had a form which was relatively higher at the youngest ages, and lower around age 18 and at the more advanced ages, than the curve estimated by the additive model. Again, the opposite was true for the years after 1975.

In fact, the mortality curve estimated by the simple additive model for a year j is simply the curve shown in Figure 3a plus the column effect b_j. We may compare this estimated curve with the observed one to see the changes in the shape of the age-curve with respect to the average one. In Figures 5a and 5b are shown the comparison, for 1946-1950 and 1976-1980, between the estimated and observed curves (where we have calculated the simple averages of the estimated and observed curves over the five year periods). This comparison confirms our expectations on the inadequacies of the simple additive model, and demonstrates the importance of the multiplicative terms for removing rectangular non-additivity.

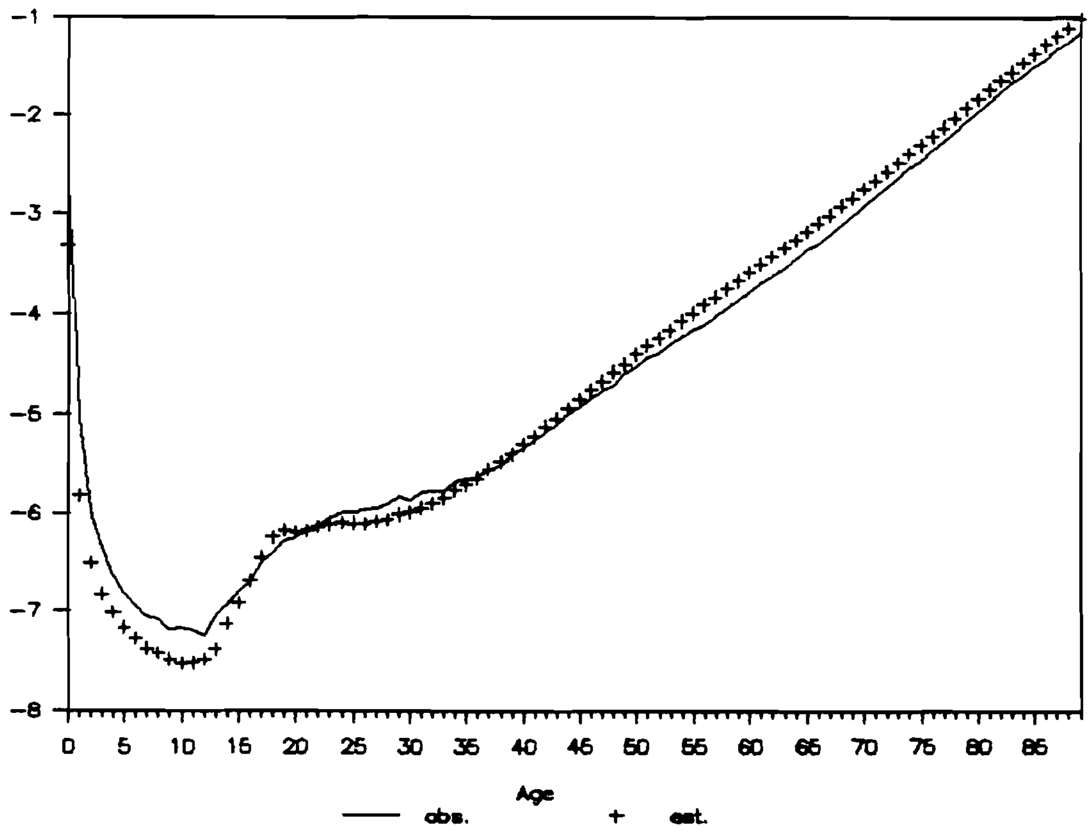
In examining the diagonal coefficients of the model, we see further non-additivity in the original data which seems not to be explained by the multiplicative part of the model. Figure 6 shows the estimated diagonal effects for all cohorts included in the study, along with a band of significance (approximatedly 95%) around zero. What comes out clearly is an apparent "cycle" in these diagonal effects, beginning around 1896 and continuing until around 1960. After and before these two dates, no significant or consistent patterns seem to be identifiable.

We use the term "cycle" with some hesitation, since in fact we are identifying groups of cohorts which we may call "peculiar" due to their seemingly anomalous mortality experience. The fact that one group of cohorts with positive diagonal effects tends to be followed by another with negative ones has led us to the convenient use of the term "cycle". It should be understood, however, that this is merely a cycle of relative cohort mortality patterns within a given period of time.

Those groups with positive diagonal effects include the cohorts 1896-1909, 1925-1935, and 1950-1959. These groups have demonstrated mortality levels over the period 1946-1981 which are relatively higher than those of the neighboring cohorts. On the con-

Figure 5. Comparison of mortality age-curves for two periods: observed vs. estimated from simple additive model (transformed scale – see equation (1)).

a) 1946–50



b) 1976–80

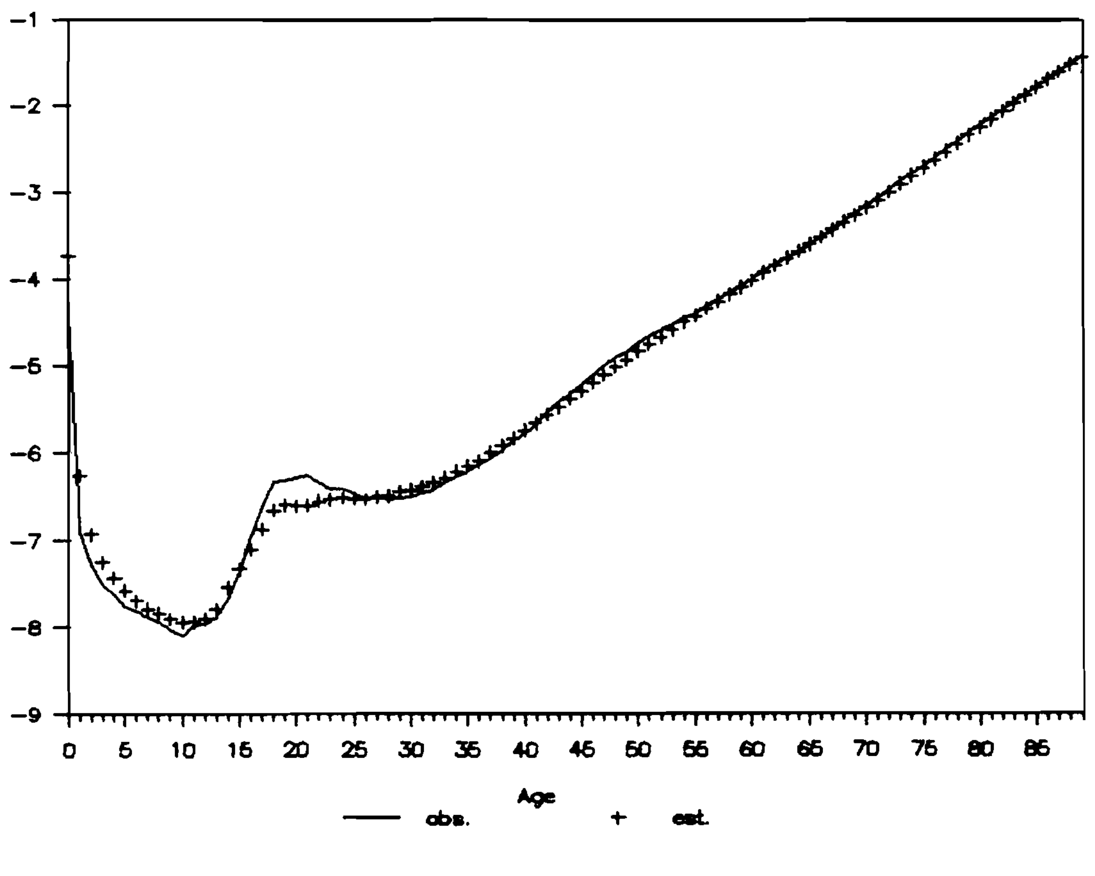
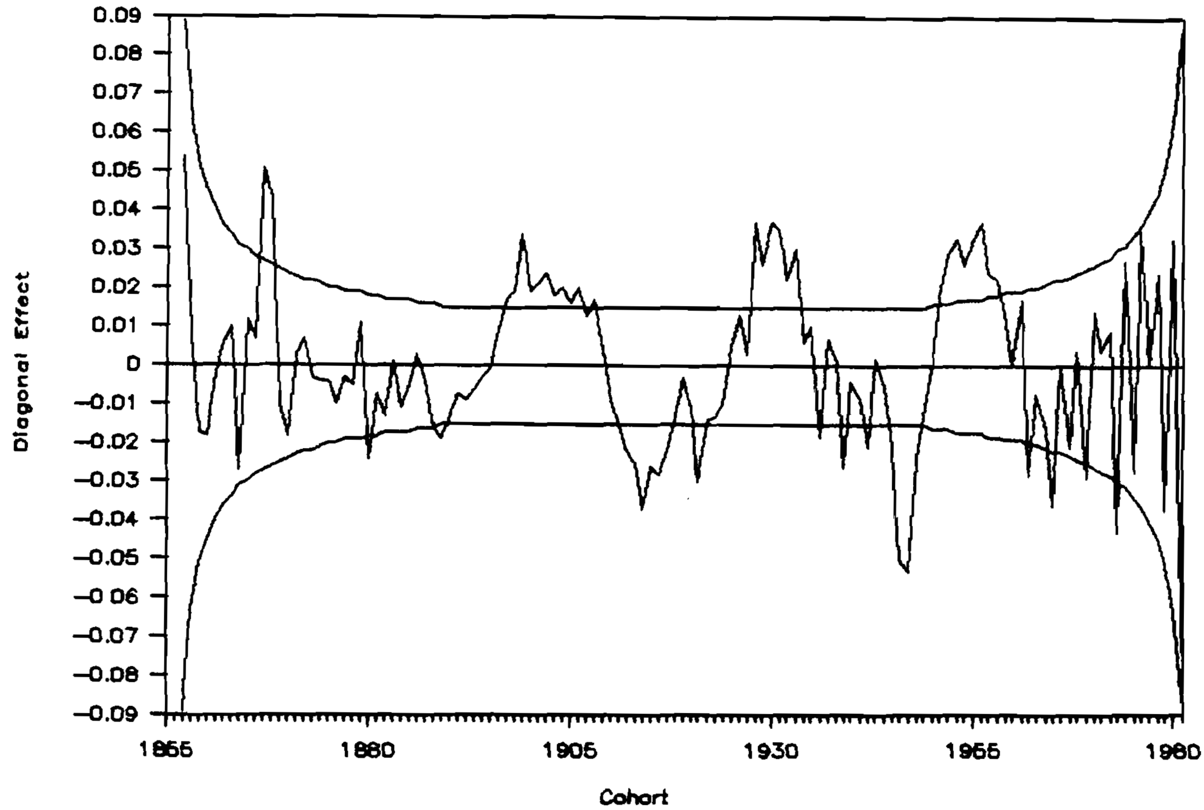


Figure 6. Residual diagonal effects by cohort. Identification of cohorts with a peculiar mortality experience. French male post-War mortality.



trary, the groups of cohorts with negative diagonal effects include 1910-1924 and 1936-1949 (if we ignore the few, very small positive peeks in the last group—see Figure 6). These are the groups with relatively lower mortality levels during the post-War period with respect to the other cohorts alive during these years. It may be said, furthermore, that the level of the diagonal effect corresponds approximately to the percent of deviation of the observed level of mortality (in terms of the force of mortality) from that which would be predicted by a model containing only additive and multiplicative terms. Thus, a diagonal effect of 0.02 indicates an "excess mortality" of around 2%, where clearly the "excess" is relative to the level of mortality expected given the age and period of time considered.

The identification of diagonal effects which in many cases appear to be significantly different from zero may lead one immediately to consider the influence of the life history of the cohorts concerned. As noted earlier, many authors have attempted to analyze large matrices of mortality data (as well as fertility data, etc.) in terms of "age, period, and cohort effects". And although it is our opinion that the precise meaning of the three terms have seldom, if ever, been clearly laid out, we are convinced that any complete definition of the term "cohort effect" in mortality must include consideration of peculiar cohorts of the kind we identify in this paper. We would thus say that surely the diagonal effects which we document here are some kind of "cohort effect", but they are not necessarily the only type which might be possible in theory or observable in practice.

Hence, it seems appropriate to attempt some kind of interpretation of the mortality experience of these peculiar cohorts during the post-War period in light of their life histories. In particular we may consider those cohorts which appear to be relatively disadvantaged (recalling, 1896-1909, 1925-35, and 1950-59). The first two of these groups consist primarily of those cohorts who were involved in the two World Wars at young ages or who went through adolescence near the end of the wars. This is especially true if we limit our discussion to those cohorts in the middle of these two groups, who tend to have quite elevated values for their diagonal effects (in all cases, the end cohorts in each group have values which are not significantly different from zero, and seem to be points of transition from the positive to the negative groups).

The case of the cohort group from 1950-59 suggests no immediate explanation. It would seem to be a suspicious result that these cohorts should have such strongly positive diagonal effects, since we observe them for a relatively brief period and at relatively young ages, at which, we have already observed, changes in mortality levels tend to be abrupt and unpredictable. In any case, we thought it appropriate to consider whether the positive effect for these cohorts depended heavily on one age group or another. However, in an examination of the residuals from the additive and multiplicative model (from which the diagonal effects are calculated), we observed no ages which seemed to dominate in the calculation of the diagonal effects. In fact, the residuals for these cohorts tend to be positive over the entire age range observed.

The cohorts which prove to be relatively advantaged during the post-War period (1910-24 and 1936-49, with the same considerations for the end cohorts) are those born during or around the two World Wars, many of whom suffered high mortality in infancy or early childhood. If those in the first group did indeed participate in the Second World War, it was at a more advanced age than those in the group 1925-35, and hence farther from the apparently fragile ages of adolescence.

V. Discussion

Several brief comments can be made on the validity, generality, and interpretation of these results. A first question of validity concerns the quality of the data being analyzed. After careful examination of the methods of calculating the age-specific probabilities for France in the post-War period, we are convinced that there exists no possibility for artifactual findings as a result of systematic biases in the data. In particular, the denominators used in the calculation of these probabilities have the advantage of being reconstructed from frequent, unevenly-spaced national censuses (taken every 6, 7, or 8 years during this period). This uneven timing of the censuses minimizes the possibility that the same cohorts would be systematically over- or under-counted in successive censuses due to some phenomenon of age-heaping, a problem which seems to affect the results for France during the pre-War period and for other countries studied.

A test of the validity of the method of analysis has been carried out using simulated mortality data, based on Coale-Demeny model life tables. The results have indicated that, in the case of "clean" data (that is, simulated data which should theoretically show no diagonal effects), the method in no way indicated the presence of unexpected diagonal patterns. After introducing artificial diagonal structure to the simulated data, the method proved capable of reproducing that structure, even in the presence of random or contaminated noise.

Early examinations of the post-War mortality experience of two other countries, Italy and Holland, have given support to the generality of these results for males. In all the cases, the cohorts affected and the manner in which they are affected are similar. In general, we have not been able to identify similar cohort "cycles" for females in these countries.

These results also seem to serve as a confirmation of the findings of other authors on the long-term effects of the World Wars on survivors (for example, Vallin, 1973, and Horiuchi, 1983), particularly those males who participated directly in the wars or who had their adolescence during or near the end of the wars. We would add, however, that these cohorts experiencing "excess" mortality seem actually to be part of a larger pattern, tending to follow, or to be followed by, other cohorts showing mortality levels which are relatively lower than expected. It is thus difficult, in examining any single dataset, to specify whether it is the former group which has been debilitated, or whether the latter has been selected. Clearly, our real interest is in determining whether the mortality experience in the post-War period of both groups of cohorts is different from what it would have been in the absence of the long-term effects of the two conflicts. However, in an analysis of the type which we are proposing (and, we claim, in the other analyses mentioned above), it remains unclear whether the long-term effect has been to increase the mortality levels of some cohorts or to decrease the levels of others, or if some combination of both has occurred.

One step toward the resolution of this difficulty may be the comparison of similar patterns, as we find them, for both sexes and for various countries. The fact that some countries were more or less affected by the two wars, or that they were affected in slightly different time periods, may throw some light on the underlying causes. Furthermore, it is clear that women in general should not show the effects of having participated in the conflicts, although they may have still suffered malnutrition and other adverse effects.

A further development of this work needs to include an identification of the mechanisms which produce the observed cohort differentials. In particular, we need to isolate the ages at which these differentials are the most pronounced and the causes of death which seem to contribute. In any case, a more complete understanding of these results will require a large measure of interdisciplinary collaboration between demographers, biologists, and other medical professionals.

Appendix

Method of Fitting

The purpose of this appendix is to provide a brief overview of the fitting methods employed, including their optimality properties and the means of calculation. The development is not intended to be comprehensive, and interested readers should consult various references for further discussion (in particular, Mandel, 1971, McNeil and Tukey, 1975, and Emerson and Wong, 1985).

As stated in the text, our optimality criterion has been the minimization of the squared residuals for the additive-plus-multiplicative model (where the number of multiplicative terms may be changed freely), followed by a second minimization of the squared errors in fitting the diagonal effects to the first set of residuals. We thus describe the calculation of the coefficients in the specific case of the add-plus-two model and then the fitting of the diagonal effects to the residuals.

The add-plus-two model,

$$f_{ij} = a_i + b_j + c_i d_j + g_i h_j + e_{ij} \quad , \tag{A.1}$$

requires the identifiability constraints,

$$\sum_j b_j = \sum_i c_i = \sum_j d_j = \sum_i g_i = \sum_j h_j = 0$$

$$\text{and} \quad \sum_j d_j^2 = \sum_j h_j^2 = 1$$

where $i = 1,\ldots,I$ and $j = 1,\ldots,J$. It is easy to verify that the least squares solutions for a_i and b_j are given by

$$\hat{a}_i = \frac{1}{J} \sum_j f_{ij} \quad \text{and} \quad \hat{b}_j = \frac{1}{I} \sum_i (f_{ij} - \hat{a}_i) \quad . \tag{A.2}$$

Then, defining

$$f_{ij}^{(1)} = f_{ij} - \hat{a}_i - \hat{b}_j \quad ,$$

the least squares solutions for c_i and d_j must satisfy then equations

$$\hat{c}_i = \sum_j \hat{d}_j f_{ij}^{(1)} \quad \text{and} \quad \hat{d}_j = \frac{\sum_i \hat{c}_i f_{ij}^{(1)}}{\|\hat{c}\|^2} \quad , \tag{A.3}$$

where

$$\|c\| = \sqrt{\sum_i c_i^2} \quad .$$

Likewise, defining $f_{ij}^{(2)} = f_{ij}^{(1)} - \hat{c}_i \hat{d}_j$, the coefficients $\hat{g}_i$ and $\hat{h}_j$ must satisfy

$$\hat{g}_i = \sum_j \hat{h}_j f_{ij}^{(2)} \quad \text{and} \quad \hat{h}_j = \frac{\sum_i \hat{g}_i f_{ij}^{(2)}}{\|\hat{g}\|^2} \quad . \tag{A.4}$$

The least squares solutions for the add-plus-two model may thus best be calculated in steps, where the first step (the additive portion) is quite straightforward. For the two multiplicative terms, however, an iterative solution is required. One routine which seems to work fairly well is the following:

1) $$c_i^{(n+1)} \leftarrow \sum_j d_j^{(n)} f_{ij}^{(1)} \quad ,$$

2) $$d_j^{(n+1)} \leftarrow \sum_i c_i^{(n+1)} f_{ij}^{(1)} \; / \; \|c^{(n+1)}\|^2 \quad ,$$

3) $$d_j^{(n+1)} \leftarrow d_j^{(n+1)} \; / \; \|d^{(n+1)}\| \quad ;$$

where we choose as starting values:

$$d_1^{(0)} \leftarrow -\frac{1}{\sqrt{2}} \; , \; d_2^{(0)} \leftarrow \frac{1}{\sqrt{2}} \; , \text{ and } \; d_j^{(0)} \leftarrow 0 \; \text{ for } \; j = 3,\ldots,J \quad .$$

Clearly, the third multiplicative term is calculated in a similar manner, using the residuals from the add-plus-one model.

We may define the residuals from the add-plus-two model as

$$f_{ij}^{(3)} = f_{ij}^{(2)} - \hat{g}_i \hat{h}_j \quad ,$$

on which we will calculate the diagonal terms, o_k, where $k = j - i$. The least squares solution for o_k in this case is merely the mean value of the residuals along the k^{th} diagonal.

REFERENCES

Boleslawski, Lech (1985) Differences in mortality among generations as a result of the world wars [Roznice w umieralnosci miedzy generacjami jako skutek wojen swiatowych]. *Studia Demograficzne*, No. 4/82, pp. 51-71.

Caselli, G. (1986) "The influence of cohort effects on differentials and trends in mortality", International Union for the Scientific Study of Population, Seminar on comparative studies of mortality and morbidity: Old and new approaches to measurement and analysis, Siena, Italy.

Caselli, G. and Egidi, V. (1981) New trends in European mortality, *Population Studies*, No. 5. Strasbourg: Council of Europe.

Emerson, J.D. and Wong, G.Y. (1985) Resistant nonadditive fits for two-way tables, in Hoaglin, D., Mosteller, F., and Tukey, J. (eds.), *Exploring Data Tables, Trends, and Shapes*. New York: John Wiley and Sons.

Hobcraft, J., Menken, J., and Preston, S. (1982) Age, period and cohort effects in demography: a review, *Population Index* 48(1):4-43.

Horiuchi, S. (1983) The long-term impact of war on mortality: old-age mortality of First World War survivors in the Federal Republic of Germany, *United Nations Population Bulletin*, No. 15, pp. 80-92.

Mandel, J. (1971) A new analysis of variance model for non-additive data, *Technometrics* 13:1-18.

McNeil, D.R. and Tukey, J.W. (1975) Higher-order diagnosis of two-way tables, illustrated on two sets of demographic empirical distributions, *Biometrics* 31:487-510.

Okoro, S. (1981) "Increase in mortality of middle-aged males in Japan", Nupri Research Paper, Series No. 3, June.

Tukey, J.W. (1977) *Exploratory Data Analysis*. Reading, MA: Addison-Wesley.

Vallin, J. (1973) "La mortalité par generation en France, depuis 1899", *Travaux et Documents, Cahier No. 63*, Presse Universitaires de France, Paris.

Vallin, J. (1984) "Tables de mortalité du moment et par generation 1899-1981: mise a jour provisoire des tables annexes du cahier 63", INED, Paris.

Vaupel, J.W., Gambill, B.T., Yashin, A.I., and Bernstein, A.J. (1985) *Contour Maps of Demographic Surfaces*, WP-85-33, International Institute for Applied Systems Analysis, Laxenburg, Austria.