UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF SOCIAL SCIENCE

SOCIAL STATISTICS

Doctor of Philosphy

MODELLING WITH AGE, PERIOD AND COHORT IN DEMOGRAPHY

by Walter Richard Gilks

Age, period and cohort are amongst the most fundamental of demographic variables, and many descriptive and theoretical models of aggregate level demographic data involving two or all three of these date-related variables have been proposed. Models which involve all three variables have been criticised as illogical (Goldstein, 1978, 1979) and statistically insupportable (Glenn, 1976) because of the logical relationship between age, period and cohort.

This thesis shows that simultaneous age, period, cohort models are not necessarily illogically conceived, although statistical supportability is a more serious problem. The extent to which theoretical considerations may usefully be incorporated into such models is examined, and methods for doing so are also explored. Particular attention is paid to the additive age-period-cohort model, the bimodel (derived from Gabriel's (1971) Biplot technique), and cohort-experience models (Hobcraft et al, 1979). Developments of these models lead to a technique for model generation, and to theoretically intriguing nuptiality and fertility models.

The use of date-related variables with individual-level data is explored using proportional hazards models (Cox, 1972) of World Fertility Survey data, and serious biasing mechanisms are found to be in operation in these circumstances. This analysis shows that age and the pace of previous fertility have a profound effect on current fertility, and finds evidence for a risk of infecundity following childbirth.

Acknowledgements

I gratefully acknowledge the generous and wise guidance of John Hobcraft of the World Fertility Survey, London, in his supervision of this research; and the comments and suggestions of Ian Diamond of the Social Statistics Department, Southampton University, on earlier drafts of these chapters.

My thanks are especially due to my wife, Anoma, whose patience, kindness and encouragement have sustained me in my efforts to create this thesis.

W. R. Gilks. Nov. 1982.

Contents | Page

Page

Tables

Page

Figures Page

Chapter 0. Introduction.

0.1 Age, Period and Cohort.

Age, period and cohort are the most fundamental of variables for demographers. Data on these variables are collected in virtually all social surveys and censuses, and also routinely via vital registration systems. Age-specific demographic data have been routinely collected since the eighteenth century in Sweden and since the nineteenth century in most countries of Western Europe. Moreover, age-specific data contained in parish burial registers go back much further and have enabled latter-day demographers to gain important insights into historical populations (see, for example: Jones, 1980). Perhaps the popularity of the three variables is due to the fact that they all are connected with dates, and consequently are well-defined and often reasonably easily measured.

Of the three variables, age might appear to be the least important, since it could be argued that there is little point in discussing age trends and differentials in abstract from the time-periods and cohorts from which the data are drawn. However, much of the work presented in this thesis is oriented towards doing just this; abstracting age variation from other sources of variation, in order to gain insights into time-invariant substantive processes underlying age-specific demographic data. Such processes include biological mechanisms, which have profound effects on fertility and mortality, and less direct effects on nuptiality and migration. For most types of demographic data, age differentials are much greater than period or cohort differentials, and this underlines the importance of the age dimension in demographic research.

Undoubtedly economic and social conditions in any time- period also have a powerful effect on demographic processes: within the last two centuries substantial reductions in mortality and fertility rates have been recorded for most countries of the world. However even within short periods of time large fluctuations in the rates of demographic processes can occur, as evidenced by the recent 'boom' and 'bust' in fertility in developed countries, in which fertility almost doubled, then halved, within the space of twenty years. Wars and famines are obvious examples of conditions which have a direct and immediate effect on mortality, and these conditions also affect levels of migration, fertility and nuptiality. Standards of education, nutrition, hygiene and medical provision vary over time, and have considerable impact on cross-sectional mortality and fertility levels in particular.

The cohort variable is of importance since it identifies groups of individuals immutably throughout their lifetimes. However, in general it is harder to conceive of factors underlying demographic processes which may be linked with cohorts, rather than with periods or ages, since events and conditions in the history of a birth-cohort may affect other birth-cohorts similarly, and may not affect all members of any birth-cohort equally, tending to destroy whatever homogeneity each birth-cohort may have initially possessed. Much of the ensuing development is motivated by the need for a careful evaluation of the usefulness of the cohort variable in demographic analysis.

In some situations it is useful to consider other types of 'age' variable, such as duration of marriage or motherhood; and also other types of 'cohort' variable, such as marriage or motherhood cohort. Further types of date-related variables include 'age-at-entry' variables such as age-at-marriage or motherhood. The usefulness of these variables is determined in the first place by whether or not the data permit their calculation.

0.2 Modelling with age, period and cohort.

There are in general three main reasons for modelling data: - to gain substantive insights, to validate existing substantive theory, and to form a basis for projections. Demographers are concerned with all three, although the ultimate goal is often to project. Models may be placed along a continuum ranging from the purely theoretical to the purely descriptive. A theoretical model is one which makes assertions about substantive phenomena thought to be underlying the data: these assertions may be derived from a combination of intuition, reasoning, previous results and possibly the data themselves. A descriptive model is one which asserts nothing substantively, and which aspires only to describe the variation in the data. For most purposes, a theoretical model which also describes the data well would be greatly preferred, provided that some reliance could be placed in its substantive assertions.

However this view is not universally accepted in connection with population projection: Brass et al (1968) and Brass (1971, 1974a, b) have developed a class of 'relational' model applicable to any demographic age distribution, which basically requires only the substitution of one standard age schedule to become context-specific. This class of model is essentially descriptive, and can be used to separately project the components of population change. Brass (1980) has taken the removal of theoretical considerations one stage further by demonstrating that population size may be projected without reference even to the separate components of population change. At the other extreme , Easterlin and Condran (1976) have explained recent movements in fertility in England and Wales and elsewhere in purely theoretical terms involving cohort sizes. Projection techniques are not considered in this thesis, although some of the models and techniques developed may have considerable potential for projection. This thesis does however examine in detail the questions of how and to what extent theoretical considerations may usefully be incorporated into models of demographic processes.

Some of the early attempts to incorporate age into a demographic model were motivated by the desire to correlate the over-all level in an event rate of a demographic process with the age-structure of the process, across time-periods or cohorts. Thus age-specific models were proposed, applicable either to cross-sectional or longitudinal data, the choice usually being determined by the numbers of reasonably complete periods and cohorts available in the data. Most of these models were descriptive rather than theoretical. The earliest of such models occur in the field of mortality and involve both empirical distributions, as in the case of the Breslau life-table of Halley in 1693 (see Smith and Keyfitz, 1977), and mathematical distributions due to Gompertz in 1825 and Makeham in 1867 (see Smith and Keyfitz, 1977). In the field of fertility the earliest age-specific model appears to be that of Tait in 1866 (see Yule, 1906) who represented legitimate fertility rates with a simple linear function of age.

The Gompertz and Makeham laws of mortality still continue to be widely used today. Mortality life-table relational models involving one or more empirically determined age distributions have also been developed (Ledermann and Breas, 1959; Bourgeois-Pichat, 1962; Coale and Demeny 1966; Brass et al (1968); Brass (1971, 1974a, b); Zaba, 1979; Le Bras, 1979; Hogan and McNeil, 1979; Hobcraft, 1979).

Recent attempts to find mathematical expressions to describe age-specific fertility rates include: polynomials (Brass, 1960; Brass et al, 1968); the Beta distribution (Mitra, 1967; Romanuik, 1973; Mitra and Romanuik, 1973); Johnson's (1949) S_B functions (Talwar, 1974); the Lognormal and Gamma distributions (Duchêne et al, 1974); a specialised non-linear form due to Mazur (1963); and the Gompertz and Makeham functions (Wunsch, 1966; Murphy and Nagnur, 1972; Murphy, 1982). As with mortality, a number of relational models of fertility involving empirically determined age distributions have also been proposed (Coale and Trussell, 1974; Brass, 1974b; McNeil and Tukey, 1975; Pittenger, 1980; Murphy, 1982). Hoem et al (1981) review and test several of these age-specific fertility models. Relational marriage-duration-specific fertility models have been developed by Farid (1973) and Page (1976).

Theoretically motivated models of age-specific nuptiality rates have been put forward by Hernes (1972) who uses a mathematical age distribution, and by Coale (1971) and Coale and McNeil (1972) who use a relational empirical formulation.

The question of whether age-specific models should be applied to cross-sectional or to longitudinal data has long been debated. Derrick (1927) has argued that cohorts provide a more consistent basis for projecting mortality than do periods, whilst the problems of projecting the most recent, incomplete, cohorts has led Brass (1974b) to take the opposite view. Kermack et al (1934), analysing age-specific mortality rates for Sweden and Scotland, found greater regularity within cohorts than within periods, although Cramer and Wold (1935) using very similar data could not find evidence to support this result. Frost (1939) and Springett (1950) in connection with tuberculosis mortality, and Case (1956a, b) and Townsend (1978) in connection with cancer mortality all favour, on empirical and theoretical grounds, the cohort perspective; but Osmond and Gardner (1982) find evidence for some cancer sites for a period perspective. In the context of fertility analysis, various arguments have been put forward in favour of the cohort or period perspectives: Easterlin (1968, 1973) suggests that 'relative cohort size' is responsible for changes in the age pattern of fertility, whilst Lee (1980) proposes a theory of 'target' fertility determined by factors operating cross-sectionally. Wunsch (1979) and Preston and

McDonald (1979) come to opposite conclusions about period and cohort perspectives in the analysis of divorce rates.

The difficulty of deciding on theoretical grounds between the period and cohort perspectives of age-specific models has led several demographers to attempt to combine age, period and cohort components into a single model. The simplest and most commonly used model of this form involves additive age, period and cohort parameters. This formulation has been used in a variety of applications in the social sciences, and in the demographic context has been used by Greenberg et al (1950) to model incidence rates of syphilis; by Sacher (1957, 1960, 1977) and Mason and Smith (1979) to model tuberculosis mortality rates; by Barrett (1973, 1978a, b), Beard (1963) and Osmond and Gardner (1982), to model site-specific cancer mortality; by Thurston (1979) for nuptiality rates; and by Sanderson (1979), Isaac et al (1979), Pullum, (1980) and Barrett (1979) with fertility data. Unfortunately this model possesses a vexing identification problem, and has also been criticised as illiogical (Goldstein, 1978, 1979) and statistically insupportable (Glenn, 1976).

Hobcraft et al (1979) have noted that in many situations the concept of constant cohort effects within a cohort, implied by the additive age, period, cohort model, is unrealistic, and suggest the use of a class of 'cohort-inversion' or 'cohort-experience' model in which cohort effects reflect the accumulated experience of each cohort. Hernes' (1972) nuptiality model in which a pressure-to-marry component depends on the proportion already married, and Lee's (1980) and Butz and Ward's (1979) target fertility models in which couples take into account existing children in order to achieve a target completed fertility, all represent cohort-experience type models. In the context of mortality, cohort-experience elements could include the selectivity effects of heterogenous susceptibility to death, investigated by Vaupel et al, (1979).

Hobcraft et al (1979) review in more detail several of the analyses cited above, and a number of these are also examined in the following chapters.

0.3 Objectives.

Several of the aims of this thesis have already been alluded to in the two sections above. The central objective is to contribute to age, period, cohort methodology by considering a number of more or less distinct issues.

In chapter 1, Goldstein's (1978, 1979) argument concerning the illogicality of age, period, cohort models is examined, and several techniques of dealing with the identification problem of the additive model are reviewed. Statistical tests for non-additivity are also developed.

Chapter 2 explores the possibility that simultaneous age, period and cohort factors are not necessary to describe the data variation if a sufficiently flexible and powerful two-factor model is developed. It is shown in chapter 2 that a period- (or alternatively cohort-) weighted sum of two empirically determined age-distributions represents such a model. This model, termed here the 'bimodel', is therefore of the relational type. Its algebraic structure is the same as that of a model of age specific fertility suggested by McNeil and Tukey (1975), and also belongs to the family of models considered by Hogan and McNeil (1979) in the context of mortality. It is equivalent to the algebraic basis of the 'Biplot' (Gabriel, 1971), a technique for graphical representation of multivariate data, and this connection provides a very efficient method of estimating the empirical distributions of the bimodel and a method by which the data may be revealingly displayed. The Biplot technique is itself closely connected with other multivariate techniques - in particular Principal Component Analysis and Correspondence Analysis (see, for example; Benzécri, 1976), the latter having been used in the demographic contect by Brouard (1980) and Garenne (1980). Chapter 2 below also develops some extensions and applications of the bimodel in demography.

Chapter 3 undertakes the task of comparing cross-sectionally and longitudinally applied age-specific models; additive age, period, cohort models; cohort-experience models and the bimodel on a variety of nuptiality, fertility and mortality[1] data. The objectives are primarily to assess the extent to which theoretical considerations may be usefully incorporated into models of demographic age, period, cohort data; and to evaluate strategies for such incorporation. As a part of this evaluation the case for cohort factors versus period factors is examined. Particular attention is paid to the development of cohort-experience type models.

The first three chapters are concerned with highly aggregated data, which have the advantage that they are often available for long time-series and for a number of populations. Individual-level survey data, however, generally have the advantage of a much greater depth of information, and chapter 4 explores the use of a number of date-related variables, including age, period and cohort, derived from individual-level maternity-history data obtained from nine World Fertility Survey countries. This research builds on, and extends, work done by Braun (1980) and Casterline and Hobcraft (1981) who also have modelled maternity-history data using date-related variables. The

---

1. Some of this work, the bulk of which was originated by this candidate, has appeared in Hobcraft and Gilks (1981).

methodology used here is that of proportional hazards modelling (Cox, 1972), which has also been used in the demographic context by Stoto and Menken (1977), Braun and Hoem (1979), and Menken et al (1981). The results of the analysis of chapter 4 demonstrate that, for individual-level data, date-related variables can by very powerful in comparison with other background variables.

Finally, in chapter 5, results and conclusions from chapters 1 to 4 are drawn together, and suggestions for further research are made.

Chapter 1 - Age, period, cohort models.

1.1 Introduction.

Age, period, cohort models have been in use for at least thirty years, and have appeared in the epidemiological, sociological and demographic literatures. Yet in recent years their use has been a source of controversy, involving criticisms of illogicality (Goldstein, 1978 and 1979) and statistical insupportability (Glenn, 1976). An identification problem associated with many such models has received varied and often misguided attempts at resolution in the literature, with frequent mis-interpretations. It is the purpose of this chapter to clear the confusion surrounding age, period, cohort models, and to contribute to age, period, cohort methodology.

Age, period, cohort models are applied to data which have been collected on individuals at various ages (or age-groups), and at various points in time (or time-periods). An individual aged a at time p is therefore a member of the cohort of individuals born at time c, where a, p and c are related as follows:

$$a - p + c = 0 \qquad (1.1)$$

Typically the data to be analysed relate to categories of age, period, cohort, rather than to exact points on these axes, and typically the categories of each dimension are evenly and equally spaced. In this typical situation, if a given observation belongs to the i^{th} age, j^{th} period and k^{th} cohort categories, then i, j and k are related as follows, by virtue of equation (1.1):

$$i - j + k = \ell \qquad (1.2)$$

where ℓ is constant over all the data.

With this arrangement of data, the usual age, period, cohort model is of the following form:

$$Y_{ijk} = \tau + \alpha_i + \beta_j + \gamma_k + \varepsilon_{ijk} \qquad (1.3)$$

where Y_{ijk} is the observation of the dependent variable in the i^{th} age, j^{th} period and k^{th} cohort categories, where τ, α_i, β_j and γ_k are model parameters, and where ε_{ijk} is an error term. The model may be estimated as a regression on dummy variates which represent each age, period and cohort category. Fienberg and Mason (1979) discuss this approach in the specific context of Maximum Likelihood Estimation of Logistic models of the form of (1.3), and compare it with the Iterative Proportional Fittingtechnique.

Now the parameter estimates of model (1.3) are not unique since, for any μ_1, μ_2, μ_3, and λ, model (1.3) may be rewritten, using (1.2), as follows:

$$Y_{ijk} = [\tau-\lambda\ell-\mu_1+\mu_2-\mu_3] + [\alpha_i + (\mu_1+\lambda i)] + [\beta_j - (\mu_2+\lambda j)] + [\gamma_k+(\mu_3+\lambda k)] + \varepsilon_{ijk} \quad (1.4)$$

and replacing the items in square brackets by starred parameters:

$$Y_{ijk} = \tau^* + \alpha_i^* + \beta_j^* + \gamma_k^* + \varepsilon_{ijk} \quad (1.5)$$

which is the same as model (1.3) except that the model parameters differ from those in (1.3) by arbitrary linear quantities, although the error term ε_{ijk} is unaltered. Thus model (1.3) lacks identification due to the four degrees of freedom in the parameters represented by μ_1, μ_2, μ_3 and λ in equation (1.4). The value of a first-difference in the parameters in equation (1.5) may be related to its value in (1.3) in the following manner:

$$\begin{aligned} \alpha^*_{i+h} - \alpha^*_i &= [\alpha_{i+h} + \mu_1 + \lambda(i+h)] - [\alpha_i + \mu_1 + \lambda i] \\ &= \alpha_{i+h} - \alpha_i + \lambda h \end{aligned} \quad (1.6)$$

showing that first-differences in parameters are also inestimable due only to the degree of freedom represented by λ. The value of a second-difference in the parameters in equation (1.5) may be related to its value in (1.3) in a similar manner:

$$\alpha^*_{i+2h} - 2\,{}^*_{i+h} + \alpha^*_i = \left[\alpha_{i+2h} + \mu_1 + \lambda(i+2h)\right] - 2\left[\alpha_{i+h} + \mu_1 + \lambda(i+h)\right]$$

$$+ \left[\alpha_i + \mu_1 + \lambda i\right] \qquad (1.7)$$

$$= \alpha_{i+2h} - 2\alpha_{i+h} + \alpha_i$$

showing that second-differences of the above form are estimable.

Model (1.3) is of the form of the 3-way ANOVA model for dimensions age, period and cohort. For any 3-way array, model (1.3) would be under-identified due to the three degrees of freedom in the parameters represented by μ_1, μ_2, and μ_3 in (1.4). For the age, period, cohort data a further lack of identification is present due to the one degree of freedom represented by λ in (1.4), and this is brought about because of the logical relationship between age, period and cohort expressed in (1.2). (Equation (1.2) effectively means that all the data lie in a 2-dimensional sub-table of the 3-way array.) Thus, because of equation (1.2), first-differences as in equation (1.6) are inestimable. The problem of constraining this one degree of freedom is sometimes referred to as the 'identification problem' of model (1.3).

Other age, period, cohort models appear in the literature; for example, Greenberg et al (1950) constrain the age parameters in a multiplicative version of model (1.3) to correspond to a Pearson type III distribution, which becomes a linear model upon taking logarithms. Age, period and cohort categories need not be evenly or equally spaced; Fienberg and Mason (1979) consider some of the problems which arise in this situation. It may also be noted that a 'cohort' may be more generally defined as a group of individuals all of whom experienced a given event at the same time, and 'age' may then refer to the time elapsed since the event. For example, Barrett (1979) analyses marital fertility rates with a marriage-duration, period, marriage-cohort model, which is formally equivalent to an age, period, cohort model.

1.2 Purposes of age, period, cohort modelling.

Some of the confusion which surrounds age, period, cohort modelling stems from the failure to distinguish two important purposes of such an activity: the descriptive and the theoretical purposes. For each purpose the approach to the modelling is quite different, particularly in relation to the identification problem of model (1.3), as will be seen below:

1.2.1 The descriptive purpose.

Faced with a typical age, period, cohort data set, the analyst may well be interested in measuring the variation in the dependent variable between ages, between periods and between cohorts purely as a means of describing the data, without making any assumptions, assertions, or drawing any conclusions about the nature of the causal or random mechanisms giving rise to the data. Model (1.3), being the 3-way ANOVA model, might seem the appropriate model to fit, and the identification problem described above does not prevent adjusted regression sums of squares from being calculated. (If the particular model fitting algorithm used does not permit under-identified models to be fitted, then model (1.3) may be fitted by arbitrarily constraining one first-difference in the parameters of one dimension. However, the submodels of model (1.3) containing parameters for at most two dimensions, should then be fitted without this constraint.) The identification problem does, however, frustrate attempts to measure age-effects, period-effects and cohort-effects since first differences in the parameters of each dimension are not identified, as shown in equation (1.6) above.

The difficulty in measuring variation in age, period, cohort data is, however, more pernicious than the above account would suggest. An age-effect, for example, is defined as the change induced in the dependent variable by a specified change in age, while period and cohort remain constant. Now Goldstein (1978 and 1979) has pointed out that, by virtue of equation (1.1) above, it is logically impossible to vary age whilst holding constant period and cohort; and consequently the separate, simultaneous effects of age, period and cohort are illogically conceived, not merely inestimable. It follows that even effects which are constructed as second-differences in the parameters of one dimension, such as in equation (1.7), are illogically conceived, even though they are estimable. Since the purpose of the model was to describe the variation in terms of simultaneous age, period and cohort effects, it seems that the whole approach is illogical, and thus Goldstein (1979) is led to assert that parameters for at most two of the three dimensions should be included in the model.

Goldstein's (1978 and 1979) argument is valid, yet second-differences in the parameters of one dimension are estimable and so, clearly, they do describe an aspect of the data; exactly which aspect of the data is shown by equation (1.8). From equation (1.3), for all i, j and k satisfying (1.2):

$$\left[\hat{Y}_{i+2h,j+h,k-h} - \hat{Y}_{i+h,j+h,k}\right] - \left[\hat{Y}_{i+h,j,k-h} - \hat{Y}_{i,j,k}\right]$$
$$= \alpha_{i+2h} - 2\alpha_{i+h} + \alpha_i \qquad (1.8)$$

where $\hat{Y}$ denotes the fitted value for Y. The age, period, cohort cells involved in this contrast are depicted in Figure 1.1. In this second-difference in fitted values, all three of age, period and cohort vary. Thus it is seen that the second-difference in age parameters in equation (1.7) describes (and is interpretable as) the typical second-difference between observations in the configuration of figure 1.1, over j (or k), fixing i and h. Second-differences in the parameters of the period and cohort dimensions describe analogous aspects of the data.

It should be noted that second-differences of the form in equation

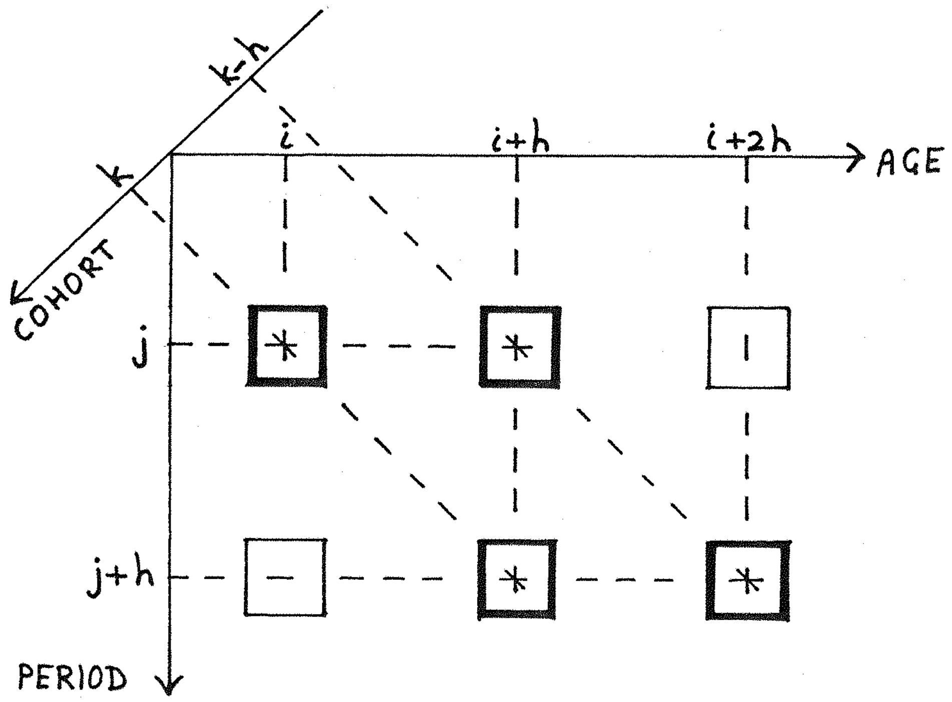


Figure 1.1

Showing (in bold lines) the age, period, cohort cells involved in the contrast in equation (1.8)

(1.7) are not the only form of estimable second-difference. Any estimable linear combination of model parameters relates to a set of linear combinations in the fitted values, and consequently describes the typical value within that set of linear combinations in the observations. Identifying model (1.3) in any particular way merely serves to produce parameter estimates which are equivalent to estimable second-differences in the unidentified parameters, and their interpretation is therefore also equivalent. First-differences in the unidentified parameters are inestimable, and therefore cannot be expressed in terms of the fitted values or observations. Consequently they do not describe any aspect of the data, and are of no interest.

The purpose of the above discussion is firstly to demonstrate that second-differences in the parameter estimates from model (1.3) <u>do</u> have some descriptive power, although the description is <u>not</u> as straighforward as might superficially be thought. Indeed the information about the data conveyed by these second-differences may be considered to be sufficiently obscure to justify abandoning the model as a descriptive tool, although with familiarity the method would possibly appear more attractive. The second purpose is to show that the inestimability of first-differences in the parameters simply means that first-differences have no power of data description, it does not mean that some aspects of the data are somehow indescribable. Consequently, failing to identify the model <u>loses nothing</u> in terms of data description.

1.2.2 The theoretical purpose.

In most circumstances the analyst would not be content merely to describe certain aspects of the variation in the data, but would rather go further to gain insight into the substantive processes giving rise to the observed variation. Model (1.3) is frequently used for this purpose, and the parameters of a given dimension are interpreted in terms of factors (specified or unspecified) associated with that dimension.

Before going on to discuss in detail the assumptions embedded in such a procedure, it is necessary to reconsider Goldstein's (1978 and 1979) remark that it is illogical to think of simultaneous age, period and cohort effects defined in the conventional way. Since it is illogical to conceive of these simultaneous effects is it then illogical to conceive of separate simultaneous age, period and cohort factors? Defining age factors as factors whose levels are indexed only by age, and defining period factors and cohort factors similarly, there is no logical relationship between the levels of age factors, period factors and cohort factors, as there is between age, period and cohort in the form of equation (1.1) or (1.2). Consequently it is perfectly conceivable that the level of the factors of one dimension may vary whilst the levels of the factors of the other two dimension remain constant, and thus the concept of separate simultaneous effects of age factors, period factors and cohort factors is not illogical. For example, Mason and Smith (1979) analyse mortality rates from tuberculosis, and hypothesise that age factors include exposure to the tubercle bacillus, that period factors include medical innovation and methods of classifying the disease, and that cohort factors include resistance to the tubercle bacillus. Now although it is logically impossible to vary age, period and cohort independently, it is not inconceivable that the levels of exposure to disease, medical innovation etc. and resistance to the disease can vary independently. In particular, if between two time-periods there is no new medical innovation nor any changes in the levels of other period factors, then between the two cells defined by these two time-periods and a given age group, only resistance to the disease and the level of the other cohort factors can vary, and thus the cohort factor effects are meaningful quantities.

In a nutshell, the concept of separate simultaneous effects of age, period and cohort is illogical, but the concept of separate simultaneous effects of age factors, period factors and cohort factors is perfectly logical.

Almost always, in practice, an age, period, cohort model is based on the very weakest a priori information or reasoning about the factors, their levels and their effects. Thus justification for the assumptions embodied in the model usually rests entirely on the goodness-of-fit of the model to the data. Furthermore, in general, not all aspects of the model can be validated by its goodness-of-fit, as will be seen below.

All age, period, cohort models involve at least two types of assumptions. Firstly it must be assumed, for each factor operating on the dependent variable, that its levels are indexed by only one dimension. In the tuberculosis study mentioned above, for example, it is conceivable that 'exposure to the tubercle bacillus' is not a pure age factor since its levels might well vary between different periods for the same age group. For most, if not all, age, period, cohort data sets, factors cannot be considered as treatments in an experiment. Consequently, discussion about which factors are in operation, and whether they can be considered as age, period or cohort factor effects, is usually speculative to a large extent.

The second type of assumption concerns the way in which the factors of each dimension interact. In model (1.3) for example, it is assumed that the effects of age, period and cohort factors are additive. Glenn (1976) points out that no assumption concerning the way in which age, period and cohort factors interact can be supported _by the data alone_, since other models always exist which fit the data _identically_, but which involve the factors of only two dimensions. For example, the cohort parameters in model (1.3) could be the result of a rather curious interaction between age and period factors, in which case the additivity assumption would be invalid. Thus, without suitable _a priori_ evidence, the effects of age, period and cohort factors are completely confounded. Mason et al (1976) reply to Glenn(1976), and defend the additivity assumption on the grounds that it is an overt simplification, and that models purposively simplify. Nevertheless, without a priori support for the additivity assumption, conclusions drawn from the analysis can be at best tentative. In section 1.3 below, tests for additivity against specific alternative hypotheses are described.

The third type of assumption relates specifically to model (1.3), and concerns its identification problem. Under the model, a first-difference in the parameters of a dimension represents the effect of a change in the levels of the factors indexed by that dimension. Thus a first-difference in the parameters is interpretable, yet inestimable, as shown in equation (1.6). Estimation of the effects of the factors of each dimension is therefore dependent upon a priori information or reasoning about factor levels and/or factor effects, sufficient to constrain the one degree of freedom represented by λ in equation (1.6). Various techniques of identification have appeared in the age, period, cohort literature, many of which are unreliable or misguided. Several of these are reviewed in section (1.4) below.

1.2.3 Comparing approaches.

It has been seen, under sections 1.2.1 and 1.2.2 above, that the two purposes of age, period, cohort modelling result in different approaches to modelling, each having quite different problem areas. Considering the disparity between the two approaches, it is scarcely surprising that controversy and confusion should arise amongst age, period, cohort analysts who have failed to perceive the distinction. The above discussion should alert analysts to two important differences between the two situations. Firstly, when the purpose is description, no assumptions are made and so none have to be justified; whereas when the purpose is theoretical several strong assumptions require a priori information as a basis for support, together with considerable discussion about factors thought to be in operation. Secondly, in the descriptive case, first-differences in the unidentified parameters contain no descriptive information about the data, and are therefore of no interest; whereas in the theoretical case, the unidentified first-differences in the parameters from model (1.3) represent the factor effects of central interest, and can only be estimated with the help of suitable a priori information on the factors.

The following two sections deal with issues arising when the purpose is theoretical.

1.3 Tests for Non-Additivity.

As stated in section 1.2.2 above, for the theoretical approach to age, period, cohort modelling, the assumption of additivity in the effects of age, period and cohort factors in model (1.3) must be at least partly supported by a priori information or reasoning about the factors involved, since without such a priori support other, non-additive, models could be devised which fit the data identically; for example the cohort parameters could represent a curiously constrained interaction between age and period factors. Thus, in order to have any confidence at all in the interpretations from model (1.3), at least some a priori support for the additivity assumption must be found.

In many situations a priori information might suggest that non-additivity between the factors of two or three dimensions should be in the form of a low order polynomial defined on those dimensions. In these circumstances the parameters of one dimension will be only partially confounded with non-additivity in the other two dimensions. To understand the confounding, consider the following sequence of models:

$$Y_{ijk} = \tau + \alpha_i + \beta_j + \varepsilon_{ij} \qquad (1.9)$$

$$Y_{ijk} = \tau + \alpha_i + \beta_j + \eta ij + \varepsilon_{ij} \qquad (1.10)$$

$$Y_{ijk} = \tau + \alpha_i + \beta_j + \gamma_k + \varepsilon_{ij} \qquad (1.11)$$

$$Y_{ijk} = \tau + \alpha_i + \beta_j + \gamma_k + \zeta i^2 j + \varepsilon_{ij} \qquad (1.12)$$

where η and ζ are model parameters. (It is assumed here, for ease of exposition, that age, period and cohort categories are equally and evenly spaced so that (1.2) holds. The following arguments are easily generalisable for an irregular table.) Model (1.10) represents the addition of a quadratic polynomial in age and period to the terms in model (1.9). All but the cross-product term of the polynomial are confounded with the terms in model (1.9). Model (1.11) is the same as model (1.3) and in fact model (1.10) is a submodel of this model, since by (1.2):

$$ij = (i^2 + j^2 - k^2)/2 \qquad (1.13)$$

and all of the terms on the right hand side of (1.13) are confounded with the terms in model (1.11). Moreover, equations (1.2) and (1.13) can be used to show that the quadratic component of γ_k is completely confounded with the terms in model (1.10). Model (1.12) represents the addition of a cubic polynomial in age and period to the terms in model (1.11). All the non-cross-product terms of the polynomial are confounded with the terms in model (1.11), as is the cross product term in ij; by virtue of equation (1.13). The two terms remaining, in i^2j and ij^2, represent only one additional degree of freedom, since by (1.2):

$$ij^2 = i^2j - (i^3 - j^3 + k^3)/3 \qquad (1.14)$$

and again the second term on the right hand side of (1.14) is completely confounded with the terms in model (1.11).

Finally, by (1.2), a polynomial in age and period can be reparameterised as a general polynomial of the same order in any two or all three of age, period and cohort. So, in summary, models (1.9) to (1.12) form a nested sequence. Quadratic polynomials in at least two dimensions are confounded in all but one degree of freedom with two-factor models such as (1.9), and are completely confounded with the three-factor model (1.11). Cubic polynomials in at least two dimensions are confounded in all but one degree of freedom with the three-factor model (1.11). Model (1.11) is distinguishable from model (1.10) only by cubic order terms in k.

Thus two tests may be constructed. Firstly it may be tested whether cubic order interactions are present, using model (1.11) as the null hypothesis versus model (1.12) as the alternative. Secondly, if the null hypothesis from the latter test is accepted, it may be tested whether, say, cohort factor effects in (1.11) are significantly different from a quadratic age-period interaction, using model (1.10) as the null hypothesis versus model (1.11) as the alternative.

The two tests may be applied to the tuberculosis data analysed by Mason and Smith (1979), referred to in section 1.2.2 above. This data relates to T.B. mortality of white males in Massachusetts, U.S.A., in ten-year age-groups, for every tenth calendar year between 1880 and 1970, giving 8 age categories and 10 period categories. The 80 cells in the age-by-period array encompass 17 cohort categories, each of width ten years, although most of these cohorts are not represented in the data for all of the above age categories. Previous analyses of T.B. mortality had usually reckoned only age and cohort factors to be operative. Consequently the age, period, cohort model (1.3) may be used to test for the presence of period factors; and the two tests outlined above may be used to test the assumption of additivity in model (1.3), and whether period factor effects are distinguishable from a quadratic age-cohort interaction, respectively. The analysis of variance of log T.B. mortality rates is given in table 1.1. (Four cells for which no deaths were recorded were each assigned 0.5 deaths in order to avoid zero mortality rates.) The error terms, ε_{ijk}, are assumed to be i.i.d. normal random variables.

The conventional test in the present circumstances would be of the null hypothesis of age and cohort factor effects, against the alternative of additonal period factor effects. This is test D in table 1.1 and the P-value in excess of .999 would strongly indicate adopting the full age, period, cohort specification. The folly of this step is demonstrated by test A, which tests for the presence of non-additivity in the form of a cubic-order interaction between any or all of age, period and cohort factors. The P-value for test A of approximately .997 strongly indicates that the assumption of additivity is inappro-

Source	SSq	d.f.	MSq Test A	MSq Test B	MSq Test C	MSq Test D
Grand Mean	4822.38	1				
Age factor effects (adjusted)	101.13	7				
Cohort factor effects (adjusted)	273.44	16				
Quadratic age-cohort interaction (adjusted)	2.32	1			2.320	0.490
Period factor effects (adjusted)	1.60	7*		0.229		
Cubic interaction (adjusted)	0.84	1	0.840	0.098	0.115	0.098
Residual	3.86	47	0.082			
Total	5205.59	80	Test A	Test B	Test C	Test D
	F-ratio		10.24	2.34	20.17	5.00
	P-value		∿ .997	∿.95	> .999	>.999

Table 1.1 Analysis of Variance of log mortality rates from tuberculosis. Each component of variance is adjusted for the sources of variation listed previously in the table.
*Note that, although there are 10 period categories, the adjusted period factor effects represent only 7 degrees of freedom, since 1 degree of freedom is confounded with the grand mean, 1 corresponds to the linear identification problem of model (1.3) described in section 1.1 and 1 is confounded with the quadratic age-cohort interaction.

priate, and since in the present circumstances no other a priori information exists with which to reliably unconfound the effects of age, period and cohort factors, this task must remain unattainable. Tests B and C demonstrate that the high P-value for test D could be largely accounted for by a quadratic interaction between age and cohort factors (which, as noted above, is completely confounded with the additional period factor effects.)

One of the earliest applications of age, period, cohort models appearing in the literature is of incidence rates of syphilis amongst black females in the area of North Carolina, U.S.A. (Greenberg et al, 1950). Previous analyses of syphilis incidence rates had been in terms of age and period factors only. It is interesting, therefore, to examine the evidence for an age, period, cohort model of this syphilis data, using the tests described above. The analysis of variance for the log of the incidence rates is given in table 1.2. (All rates are incremented by one incident per thousand to avoid zero rates, for comparability with the analysis of Greenberg et al, 1950). The data matrix consists of 15 single year age groups from 15 to 29 years for each of 7 time-periods between 1941 and 1947, and therefore contains 21 single-year width incomplete cohorts.

Test A in table 1.2 indicates that if age, period and cohort factors are operative, then they are not additive in their effects. (Even if the additive model was accepted then test B shows that the cohort effects are indistinguishable from a quadratic-order age-period factor interaction. However, test C shows there is no evidence for such an interaction, the additive model in age and period factor effects being quite adequate. The conventional test D also indicates that there are no additive cohort factor effects.)

It is interesting to note that in neither of the above examples

Source	SSq	d.f.	MSq Test A	MSq Test B	MSq Test C	MSq Test D
Grand Mean	508.647	1				
Age factor effects (adjusted)	51.194	14				
Period factor effects (adjusted)	8.985	6				
Quadratic age-period interaction (adjusted)	0.033	1			0.033	
Cohort factor effects (adjusted)	5.383	18*		0.299		0.285
Cubic interaction (adjusted)	1.037	1	1.037			
Residual	15.992	64	0.250	0.262	0.270	0.262
Total	591.271	105	Test A	Test B	Test C	Test D
F-ratio			4.15	1.14	0.12	1.09
P-value			∿.95	<.75	<.75	<.75

Table 1.2. Analysis of Variance of log incidence rates of syphilis. Each component of variance is adjusted for the sources of variation listed previously in the table.
*Note that the 21 adjusted cohort factor effects only represent 18 degrees of freedom, since 1 degree of freedom in confounded with the grand mean, 1 corresponds to the linear identification problem of model (1.3) described in section 1.1, and 1 is confounded with the quadratic age-period interaction.

does the analysis of variance point to an additive model in age, period and cohort factor effects.

1.4 A review of methods of resolving the linear identification problem.

Various techniques for constraining the single degree of freedom in the identification problem of model (1.3) have appeared in the literature. As stated in section 1.2 above, if the objectives of the analysis are purely descriptive then the unidentified first-differences in the parameters are of no importance. If the objectives, however, are theoretical then the unidentified first-differences are of central importance, and can only be identified with the use of suitable a priori information on the factors. Many of the identification techniques appearing in the literature are either pointless or misguided, depending upon the objectives of the analysis, which are not always made abundantly clear. Several of these techniques are reviewed below, and may be discussed under six headings. It is assumed throughout that the analyses reviewed are of the theoretical rather than descriptive type.

1.4.1 Internal procedures.

Mason et al (1973) and Pullum (1978) both suggest procedures for constraining the one degree of freedom in the identification problem by purely internal means. It is clear that purely internal techniques are bound to fail in the task of producing for each dimension parameter estimates which are reliably interpretable in terms of the effects of the factors indexed by that dimension; although of course such techniques will undoubtedly succeed in constraining the problematical degree of freedom, but to no purpose.

Mason et al (1973), noting that different sets of just-identifying restrictions on the parameters of the model cannot be assessed on the basis of the fit of the model to the data, but that different sets of over-identifying restrictions may be assessed in this manner, suggest that "a clearer picture of the 'true' effects in a given set of cohort data might be obtained by comparing the results from several distinct models making more than the minimum assumptions needed for estimability", and that such a procedure "might provide clues about the nature of ageing, cohort and period effects for the analyst unable or unwilling to make a priori constraints on the cohort model". They suggest that "if the general nature of the estimates is similar for all models ... then it is probably safe to interpret these findings in substantive terms", and suggest alternative procedures if the various submodels do not yield similar parameter estimates.

The technique of Mason et al (1973) could be of value if a priori information on the levels and effects of factors indicates the particular submodels of model (1.3) which are to be compared. However, without such a priori information, even a well-fitting model which omits, say, all cohort parameters could be reparameterised to contain a linear dependence on cohort, without altering the fitted values. (This is an important point, and it complements the fact that, without suitable a priori information, a full age, period, cohort model may be reparameterised in terms of a constrained interaction between just two of the sets of factors, without altering the fitted values.)

Pullum (1978) proposes a fully automated identification procedure which, loosely speaking, results in the regression slopes in the parameter estimates of any given dimension being related to the degree of non-linearity amongst the parameter estimates of that dimension. Thus, for example, if the parameter estimates for the cohort dimension fall on a straight line, then the procedure would make all cohort parameters equal, and thus the cohort dimension would effectively be removed from the analysis.

The latter property of Pullum's (1978) technique appears attractive and the technique also seems to represent a way of 'hedging one's bets' as to the 'correct' identification - avoiding the possibility of assigning large linear trends to the effects of factors which show few other signs of being operative. Nevertheless, no reliable interpretations can arise through the use of this technique. The age factor effects resulting from the application of a version of Pullum's (1978) identification technique to the tuberculosis data of Mason and Smith (1979) are shown in figure 1.2.

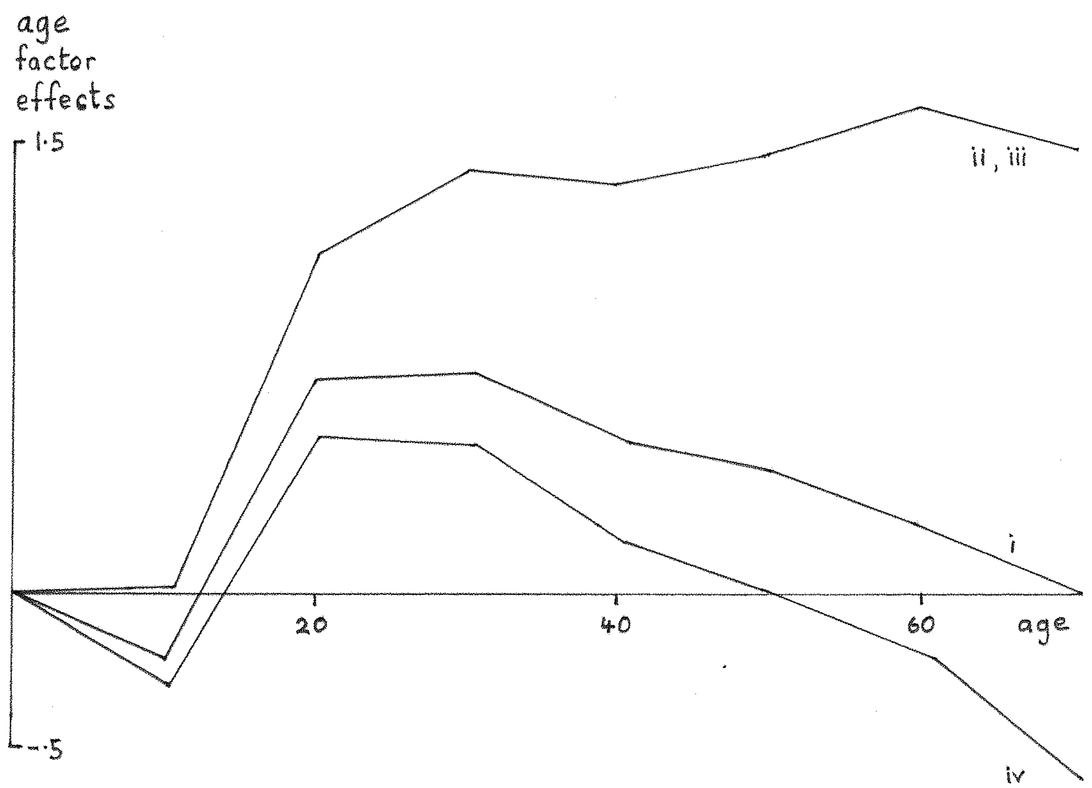


Figure 1.2

Age factor effects from model (1.3) applied to log mortality rates from tuberculosis, for four different identifications: (i) a version of Pullum's (1978) automatic identification; (ii) no linear trend in the first six cohort parameters, as in Sacher's (1960) analysis; (iii) equality in the first two age parameters, as in Mason and Smith (1979); (iv) equality in the first six period parameters, as in Mason and Smith (1979).

The two procedures outlined above are of a general nature. Particular cases of identification using internal 'evidence' may also be given: both Sacher (1960) and Mason and Smith (1979) have identified in this manner. Interestingly, and usefully, Sacher (1960) analyses the same data as Mason and Smith (1979), that is the tuberculosis data for white Massachussetts males, (although Sacher's (1960) data extends only up to 1940). Sacher (1960) constrains the regression line through the first six cohort parameters to have zero slope, so that his model is just-identified. This he does on the grounds that "the internal evidence clearly indicates that conditions were stationary during the first half of the 19th century". Mason and Smith (1979) constrain the first two age parameters to be equal on the grounds that tuberculosis mortality is usually about the same for these two age groups. Of course neither of these two identifications are justified since the same observations could be made if the age, period and cohort factor effects were altered by appropriate linear quantities. The age parameters from both these identifications are given in figure 1.2.

1.4.2 Grouping categories.

In an age-by-period array, a cohort may be defined as the set of cells which correspond to births in a given interval of time. Now there is no necessity to define the width of the cohort intervals to be the same as the width of the age and period intervals. For example, both Greenberg et al (1950) and Isaac et al (1979) define each cohort to be three times the width of each age and period category. Having defined cohorts in this way model (1.3) may still be fitted to the data, although now there will be fewer cohort parameters than would normally be the case. However, no relationship such as equation (1.2) holds in these circumstances, and consequently model (1.3) does not have an identification problem. In both of these examples referred to above, the decision to define cohort categories in this way seems to be purely for computational convenience, and in neither case is reference made to the identification problem of the more usual framework.

Superficially this seems an ideal way to dodge the identification problem. However this is a most dangerous illusion. The model in both of the above examples implies that, across the width of each of the defined cohorts, cohort factor effects are constant. There can be no internal evidence to support this, since equal linear trends in cohorts factor effects across the width of each of the defined cohorts could yield the same fitted values. That is, only a priori information can suggest that, within each of the defined cohorts, cohort factor effects are constant rather than linear. Consequently, for a given dimension, the parameter estimates from such a model cannot be reliably interpreted purely in terms of the effects of factors indexed by that dimension.

The above discussion demonstrates that age, period, cohort modelling is something of a statistical 'minefield', since it is quite possible that a researcher can innocently decide to limit the number of parameters in model (1.3) by grouping the categories of one dimension in the above manner, and neither suspecting nor finding an identification problem, proceed to mis-interpret the resulting parameter estimates.

Presumably, in the two examples quoted above, the reasons advanced for grouping cohort categories into threes could equally well have been advanced in favour of grouping period categories into three instead. Figure 1.3 contains the age parameters resulting from both over-identifications of model (1.3), when applied to the syphilis data of Greenberg et al (1950).

1.4.3 Intuition

Occasionally in the literature model (1.3) is identified by processes which might best be described as intuitive. For example, Sanderson (1979),

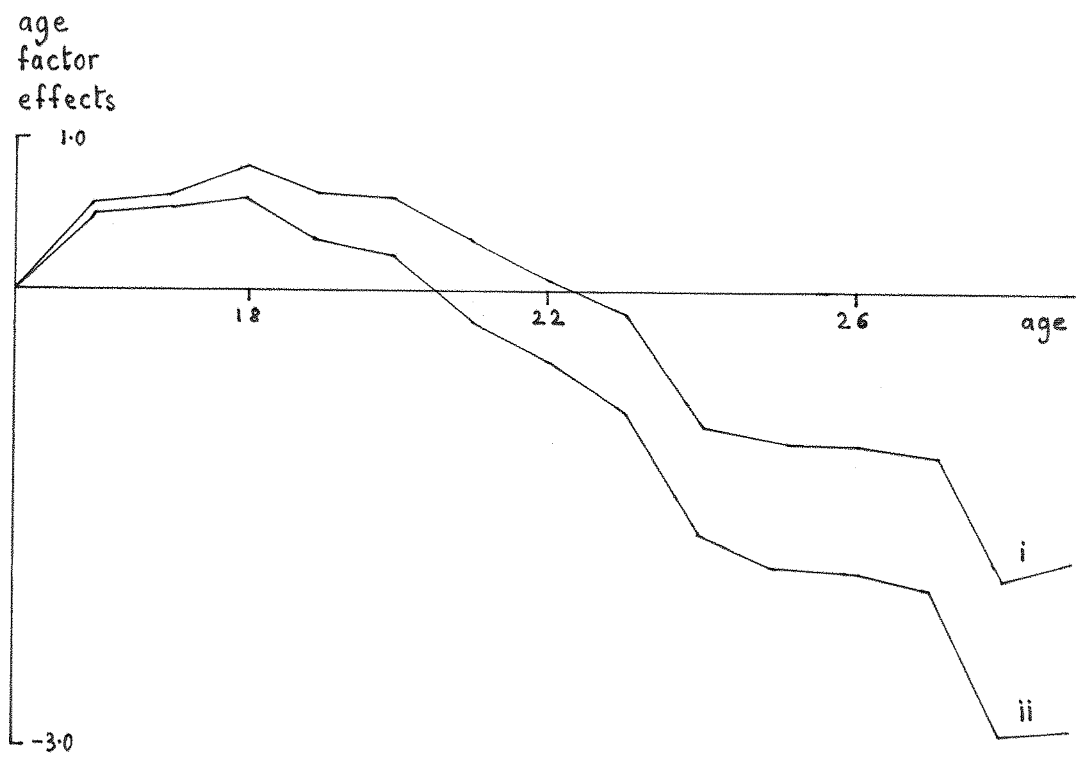


Figure 1.3

Age-factor effects from model (1.3), applied to log incidence rates of syphilis, for two alternative over-identifications obtained by grouping into threes the categories of (i) the cohort dimensions; and (ii) the period dimensions.

who analyses parity-specific birth probabilities for white females in the U.S.A., being faced with a total lack of a priori information on the factor levels and their effects, decides to identify in such a way that the slopes of the regression lines through the period and cohort factor effects are equal. This he justifies with the reasoning that some long-run factors could be expected to influence both period and cohort factors. However, without a clear idea of exactly what these factors are, and which period and cohort factors they operate on, and how they do so, this reasoning is highly speculative. Indeed the net result of all factors could well produce opposite trends in period and cohort factor effects, even if some factors do operate in the above-mentioned fashion.

Barrett (1973), in an analysis of mortality from cancer of the cervix, obtains an initial identification of model (1.3) by, arbitrarily, constraining the last two cohort factor effects to be equal. Observing that this produces an "unreasonably large positive trend" in the period factor effects, he re-identifies, constraining instead the last two period factor effects to be equal. The lack of reliable a priori information in this identification will produce a corresponding lack of reliability in the resulting parameter estimates.

1.4.4 A priori reasoning

Reliable a priori reasoning can represent a proper means of identifying model (1.3). For example Fienberg and Mason (1979) analyse educational attainment for white males in the U.S.A.; for each stage of formal education they model the logarithm of the odds of continuing to the next stage. They note that, for the vast majority of the population, formal education is complete by age 30 years, and that after age 60 years biases creep into the data due to mortality-age-education differentials and recall accuracy. Over-identification is obtained by equating age factor effects between the ages of 30 and 60 years. This over-identification is selected on the basis that it does not constrain period and cohort contributions which are of central interest in this study, and it does not constrain the regions of the age dimensions for which there are "a priori grounds for interpreting age effects". The constancy in age factor levels between the ages of 30 and 60 years (which seems reasonable to assume from the above discussion) represents a reasonable basis for (over-) identifying model (1.3).

Another example of the use of a priori reasoning in identifying model (1.3) is provided by the tuberculosis study of Mason and Smith (1979) who check their identification, described in section 1.4.1 above, by equating the first six period parameters instead, on the grounds that there is no substantive reason for letting them vary. The validity of the latter over-identification depends on the validity of their assertion about the period factor effects. The age parameters from the over-identification are shown in figure 1.2.

1.4.5 A priori data.

Beard (1963) analyses mortality from cancer of the lung, and associates cohort factors with the proportion of smokers in each cohort, period factors with the level of consumption of cigarettes in each time-period, and age factors with resistance to the disease at each age. The model is over identified using external data on the proportion of smokers in each cohort and the level of cigarette consumption in each time-period.

Farkas (1977) analyses employment rates for white females in the U.S.A. Period factors are taken to be business cycle fluctuations, and these are measured using unemployment rates for white females. Model (1.3) is over-identified by equating period factor effects with the unemployment rates. Apparently the relationship between employment and unemployment rates is not tautological, and Farkas (1977) claims that any one of a number of highly correlated macro-economic variables would serve the purpose of measuring business cycle fluctuations.

Both of the above examples demonstrate how a priori data can provide a basis for identification or over-identification of model (1.3). The reliability of the resulting parameter estimates depends on the reliability of the asserted connections with the a priori data.

1.4.6 No identification.

Some researchers, recognising their inability to reliably identify model (1.3) have contented themselves with estimating and interpreting second-differences in the parameter estimates, of the form of expression (1.7). Pullum (1980) defines a 'relative difference' which is a form of second difference, and Sanderson (1979) examines trends in first-differences between categories of one dimension a fixed distance apart. (Although first-differences are identification dependent, such trends in first-differences are identification independent). Barrett (1978b) confines his attention to peaks in the structures of parameter estimates, but strictly the positions of such peaks are not identification independent.

Sometimes there is no need to obtain a reliable identification of model (1.3). For example Thurston (1979), in analysing Swedish nuptiality, and Pullum (1980) both use model (1.3) as a basis for projections. Projecting period and cohort factor effects separately, and recombining them with the age factor effects, yields projections of the dependent variable which are independent of the choice of identification, provided that no attempt is made to compromise the linear trend in the period or cohort factor effects. Unfortunately Thurston (1979) does just that; by placing bounds on the projected period factor effects the projections become identification dependent. As another example Sanderson (1979), wishing only to detect the presence of cohort factor effects in parity-specific birth probabilities, in order to test an aspect of Easterlin's (1973) work on the economics of fertility, does not need to examine the parameter estimates themselves.

1.4.7 Comparison of identification techniques.

Several techniques for identifying model (1.3) have been described above, many of them unreliable or misguided. In particular, figure 1.2 shows various identifications of the tuberculosis data, which could lead to markedly differing conclusions concerning the effects of age factors. The period and cohort factor effects also differ between identifications, and to a similar extent. The two solutions due to Mason and Smith (1979) are vastly different, yet Mason and Smith (1979), who actually modelled logit death probabilities by Maximum Likelihood Estimation, found the solutions to be similar. The present analysis of this data is of log death rates estimated by least squares, and corresponds to the analysis of Sacher (1960). Thus the precarious nature of these identifications is emphasised. Figure 1.3 shows alternative over-identifications for the syphilis data. Although the solutions here do not vary so dramatically as those in figure 1.2, there is nonetheless a substantial difference in trend between the two sets of age parameters, the gradients in one set being over 1.5 times the corresponding gradients in the other. Again, similar differences would be observed for the period and cohort dimensions.

Analyses of variance for different identifications also differ: for example, if there is thought to be no age factor effect between the first two age groups, then submodels involving age, as well as the full age, period, cohort model, should carry the constraint of equality between the first two age parameters. For the submodels the constraint will actually affect their fit to the data. In the case of the tuberculosis data, analyses of variance for the various solutions differ to only a small extent, and the same is true for the syphilis data also. However, in principle, different identifications could produce quite different analyses of variance.

1.5 Conclusions

The distinction between the descriptive and theoretical purposes of age, period, cohort modelling brings to light some very substantial differences between the corresponding approaches to modelling. When the purpose is description, no assumptions are made; it is simply required to decompose the variance in the data along axes of interest. When the purpose is theoretical however, several strong assumptions must be made, concerning the presence of true age, period and cohort factors (and no other factors) and about the way these factors interact.

The two purposes also differ in respect of the treatment of the identification problem of model (1.3). In the descriptive case, unidentified contrasts mean nothing in terms of data description, and hence are of no interest. Different identifications merely serve to produce parameter estimates which describe the same aspects of data as certain second-differences in the unidentified parameters. In the theoretical case, however, the unidentified first-differences in the parameters of model (1.3) are of central interest since, hypothetically at least, they reflect the effects of the factors of each dimension, separately.

Both purposes encounter difficulties. In the descriptive case, the aspects of the data described by estimable contrasts in the parameters are somewhat less straightforward than might at first be imagined. In the theoretical case, the analyst is seldom able to find reliable a priori support for the necessary assumptions, and without any such information the data is incapable of distinguishing between models of widely different construction. Furthermore, even if model (1.3) can be reliably assumed, further a priori information on the factor levels and effects is required in order to estimate the separate factor effects, otherwise the estimable second order factor effects will have to suffice.

A proper approach to theoretical age, period, cohort modelling should include a full discussion of the factors thought to be in operation, together with a priori support for the assumptions embedded in the model.

Faced with these difficulties the analyst may well decide to abandon the use of age, period, cohort models, deciding instead to use models involving at most two of the three dimensions. In the theoretical case this could actually be a mistake if it is thought that age, period and cohort factors might be in operation, since even a perfectly fitting two factor model can be re-expressed in terms of factors from all three dimensions, unless a priori information indicates otherwise. Most of the difficulties associated with the theoretical case could be overcome with sufficient a priori information. The validity of the assumptions embedded in model (1.3) can be assessed with a little such information, as demonstrated in section 1.3.

Confusion about age, period, cohort modelling is evident in the literature, and exists in the minds of many who have considered using the technique. Much of the confusion arises from failure to identify the aims of the analysis: descriptive or theoretical. The above discussion draws attention to the importance of this distinction.

Chapter 2. Using the biplot to extract patterns in age, period, cohort data.

2.1 Introduction.

Tables of demographic rates or proportions frequently exhibit a high degree of regularity, especially across 'age' dimensions such as age or duration of marriage, within periods or cohorts. As will be seen in Chapter 3, simultaneous age, period, cohort models, apart from their problems of confounding discussed in Chapter 1, also often have the disadvantage that they fail to capture these age patterns efficiently. The biplot (Gabriel, 1971), a graphical technique of multivariate analysis designed to extract and display the main components of the structure of any two-way data matrix, however, is perfectly suited to capturing these age patterns; and this chapter and chapter 3 convincingly demonstrate this.

In this chapter the algebraic basis of the biplot is presented (section 2.2), and its ability to extract and display trends in demographic tables is exemplified in section 2.3. Summary measures of 'level' (e.g. total fertility rate) and 'structure' (e.g. mean age at childbearing) across the age dimensions may be conveniently represented on the biplot graph by means of level and structure axes, as is shown in section 2.4. Section 2.5 shows how cubic splines may be used to construct on the biplot smooth curves related to the age dimensions, for the purpose of interpolation. Finally, section 2.6 describes the generation of models of demographic schedules using the biplot decomposition, and the method is illustrated for age-specific legitimate live-birth rates, and compared with the model of Coale and Trussell (1974).

In the general context, the biplot may be used as a diagnostic tool to detect outliers in the data; to detect clusters of individuals or variates; to suggest regression models possibly for submatrices of the data, (Bradu and Gabriel, 1978); and as a model in its own right (McNeil and Tukey, 1975; Hobcraft and Gilks, 1981). The biplot technique has a connection with principal component analysis (Gabriel, 1971) and with correspondence analysis (Benzécri, 1976).

2.2 The Biplot.

2.2.1 Definition.

Any data matrix Y_r having m rows and n columns, and of rank r, may be decomposed as a sum of r components, each component being an mxn matrix of rank 1. This decomposition may be written:

$$Y_r = \sum_{\ell=1}^{r} \underline{u}_\ell \underline{v}_\ell' \tag{2.1}$$

where $\underline{u}_\ell$ and $\underline{v}_\ell$ are vectors of order m and n respectively. The decomposition (2.1) is not unique, and this problem is partly resolved with the requirement that the vectors $\{\underline{u}_\ell, \underline{v}_\ell ; \ell = 1,\ldots,r\}$ should be chosen so that the mxn rank k matrix:

$$Y_k = \sum_{\ell=1}^{k} \underline{u}_\ell \underline{v}_\ell' \tag{2.2}$$

is the best rank-k approximation to Y_r in the least-squares sense, for each choice of k in the range $[1,r]$.

The decomposition (2.2) of the approximation Y_k forms the basis for a useful graphical representation of Y_r. For the i^{th} row $(i = 1,\ldots,m)$ and for the j^{th} column $(j = 1,\ldots,n)$ of Y_r, vectors $\underline{g}_i$ and $\underline{h}_j$ of order r may be defined so that the ℓ^{th} elements of $\underline{g}_i$ and $\underline{h}_j$ are the i^{th} element of $\underline{u}_\ell$ and the j^{th} element of $\underline{v}_\ell$, respectively. Thus the ℓ^{th} elements of $\underline{g}_i$ and $\underline{h}_j$ are drawn exclusively from the ℓ^{th} component of Y_r. The $\underline{g}_i$ are termed row markers and the $\underline{h}_j$ are termed column markers. The first element of each of the (m + n) markers may be plotted against the second; on a separate graph the second element of each of the (m + n) markers may be plotted against the third; and so on. Thus (r - 1) graphs may be produced representing the position vectors of each of the row and column markers. This is termed the exact biplot of Y_r. The purpose of the biplot, however, is to portray only the most important information contained in Y_r, and this is achieved by plotting only the first k elements of each of the row and column markers; that is, only the first (k - 1) graphs of the exact biplot of Y_r. This is termed the rank-k approximate biplot of Y_r, and is equivalent to the exact biplot of Y_k. Typically k = 2 or 3 is sufficient to capture all the variation of interest.

Table 2.1 contains a decomposition of a matrix of single-year age-specific fertility rates for all women in England and Wales for each time period from 1938 to 1979; (O.P.C.S., 1979). The $\underline{u}$ vectors are the columns of panel (a); the $\underline{v}$ vectors are the columns of panel (b); the first two components of the $\underline{g}$ vectors are the rows of panel (a); the first two components of the $\underline{h}$ vectors are the rows of panel (b). Figure 2.1 contains the rank-2 approximate biplot of this data. For clarity the row (age) markers have been joined up, as have been the column (period) markers.

The decomposition (2.1) is still not unique despite the requirement that Y_k, in equation (2.2) should be the least-squares rank k approximation to Y_r, for all k. For example simultaneous multiplication of $\underline{u}_\ell$ by a scalar c_ℓ and division of $\underline{v}_\ell$ by c_ℓ would not alter the ℓ^{th} component, although of course the resulting biplot would be affected. This lack of identification may be removed by scaling the $\underline{u}_\ell$ and $\underline{v}_\ell$ vectors so that both row and column markers are well dispersed over the biplot. Gabriel (1971) shows that any method of removing this lack of identification confers certain distance properties on the row and column markers, although for present purposes these properties are not particularly useful.

2.2.2 Calculation.

The decomposition (2.2) may be calculated using the Singular Value Decomposition (S.V.D.) of Y_r. Algorithms for calculating the S.V.D. are extremely rapid, having almost cubic powers of convergence, and do not

Table 2.1 The first two components of the decomposition of the age by period array of fertility rates for all women in England and Wales (O.P.C.S., 1979). The rows of panels (a) and (b) form the first two components of the row(age) and column(period) markers respectively. Panel (c) contains the values of the R_k^2 statistic (equation (2.11)) for the first two components.

(a)

AGE (completed years)	$\underline{u}_1$	$\underline{u}_2$
15	-.007	.009
16	-.028	.054
17	-.082	.114
18	-.156	.152
19	-.241	.158
20	-.328	.141
21	-.408	.113
22	-.472	.082
23	-.514	.060
24	-.532	.046
25	-.531	.035
26	-.515	.024
27	-.486	.007
28	-.450	-.019
29	-.408	-.051
30	-.364	-.084
31	-.321	-.114
32	-.281	-.136
33	-.246	-.146
34	-.216	-.146
35	-.189	-.139
36	-.163	-.127
37	-.138	-.112
38	-.113	-.096
39	-.089	-.080
40	-.067	-.066
41	-.049	-.054
42	-.037	-.045
43	-.029	-.041
44	-.022	-.044

(b)

PERIOD	$\underline{v}_1$	$\underline{v}_2$
1938	-.234	-.129
1939	-.234	-.119
1940	-.224	-.099
1941	-.222	-.088
1942	-.250	-.118
1943	-.259	-.145
1944	-.282	-.191
1945	-.255	-.152
1946	-.315	-.204
1947	-.350	-.156
1948	-.310	-.095
1949	-.297	-.065
1950	-.287	-.063
1951	-.283	-.054
1952	-.286	-.043
1953	-.296	-.034
1954	-.296	-.028
1955	-.299	-.017
1956	-.316	-.002
1957	-.332	.003
1958	-.340	.014
1959	-.348	.020
1960	-.360	.019
1961	-.373	.026
1962	-.382	.040
1963	-.389	.044
1964	-.396	.043
1965	-.384	.057
1966	-.372	.071
1967	-.357	.079
1968	-.347	.090
1969	-.333	.101
1970	-.325	.115
1971	-.322	.127
1972	-.300	.126
1973	-.275	.127
1974	-.261	.122
1975	-.246	.105
1976	-.239	.090
1977	-.232	.073
1978	-.243	.066
1979	-.259	.055

(c)

R_1^2	R_2^2
.972	.997

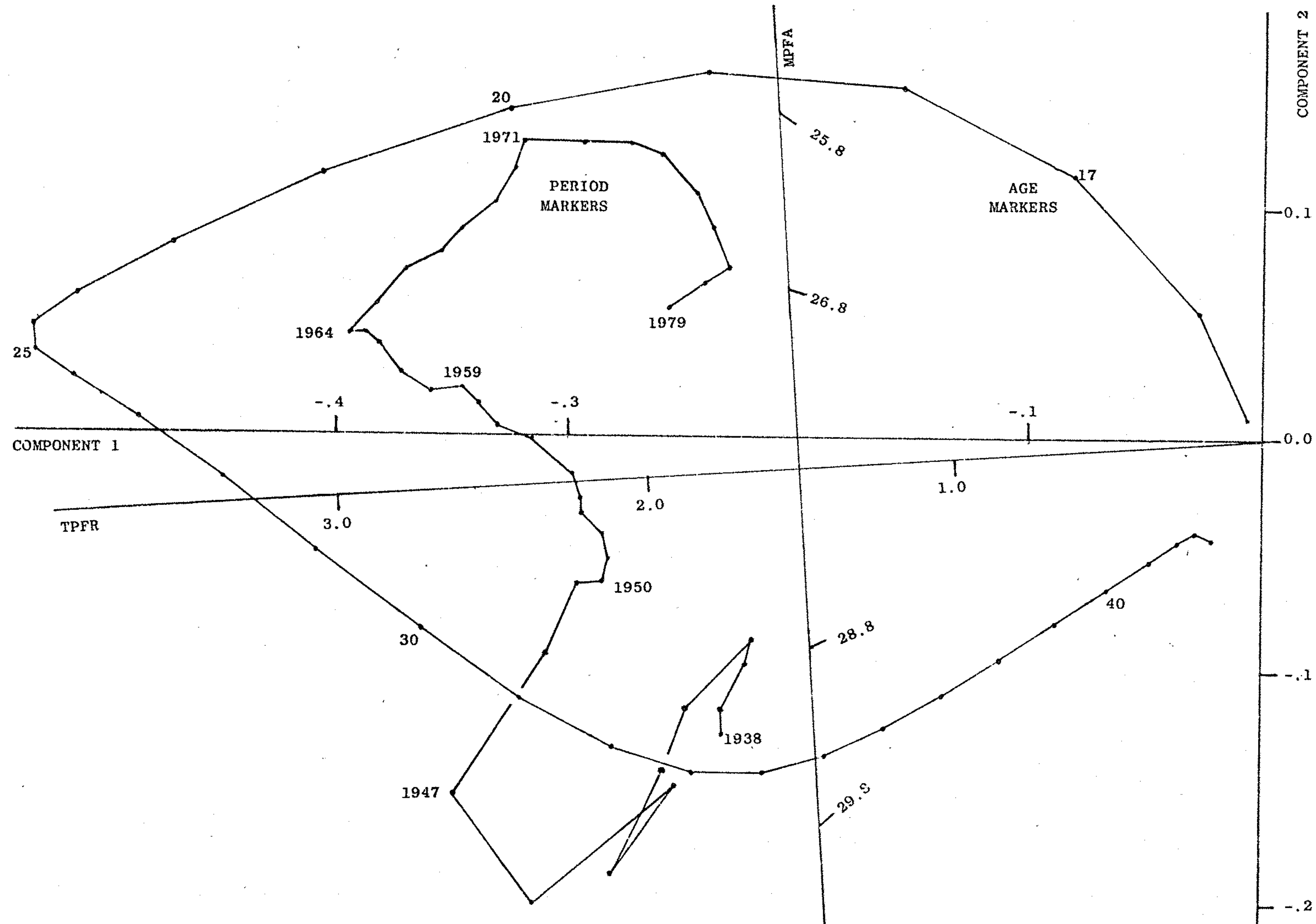


Figure 2.1 Biplot of fertility rates for all women by single years of age and time-period for England and Wales (O.P.C.S., 1979) also showing the T.P.F.R. and M.P.F.A. axes (see section 2.4). Obtained from the decomposition in table 2.1.

require user-supplied starting values. More importantly, these algorithms are widely available, being contained in most, if not all, scientific software libraries.

Any mxn matrix Y_r of rank r may be decomposed into three matrices as follows:

$$\mathbf{Y_r} = \mathbf{P} \Lambda \mathbf{Q'} \tag{2.3}$$

where P is an mxr matrix, Λ is an rxr diagonal matrix and Q is an nxr matrix; and where also:

$$P'P = I_r \tag{2.4}$$

$$Q'Q = I_r \tag{2.5}$$

$$\text{and} \qquad \lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_r > 0 \tag{2.6}$$

where I_r is the rxr identity matrix and λ_ℓ denotes the ℓ^{th} diagonal element of Λ. This is the S.V.D. of Y_r.

Any decomposition satisfying (2.2) may be obtained from the S.V.D. simply by setting:

$$\mathbf{G} = \mathbf{P} \Lambda_1 \tag{2.7}$$

$$\mathbf{H} = \mathbf{Q} \Lambda_2 \tag{2.8}$$

where G is the mxr matrix whose i^{th} row is $\underline{g}_i'$, i = 1, ..., n; where H is the nxr matrix whose j^{th} row is $\underline{h}_j'$, j = 1, ..., n; and where Λ_1 and Λ_2 are rxr diagonal matrices satisfying:

$$\Lambda_1 \Lambda_2 = \Lambda \tag{2.9}$$

In general, setting

$$\left.\begin{aligned} \Lambda_1 &= \left(\frac{m}{n}\right)^{\frac{1}{4}} \Lambda^{\frac{1}{2}} \\ \Lambda_2 &= \left(\frac{n}{m}\right)^{\frac{1}{4}} \Lambda^{\frac{1}{2}} \end{aligned}\right\} \tag{2.10}$$

will achieve a reasonable dispersion of the markers over the biplot. It may be verified, when equations (2.10) hold, that on any axis, the average squared score for the row markers is equal to that for the column markers. The biplots in figures 2.1 and 2.6 have been identified using (2.10).

The S.V.D. also enables measures of goodness-of-fit of Y_k to Y_r to be constructed. It may be shown that the sum of squared elements of Y_k is given by $(\lambda_1^2 + \lambda_2^2 + \ldots + \lambda_k^2)$ from which the proportion of variance in Y_r 'explained' by Y_k may be calculated:

$$R_k^2 = 1 - \sum_{\ell=k+1}^{r} \lambda_\ell^2 \bigg/ \left(\sum_{\ell=1}^{r} \lambda_\ell^2 - mn\bar{y}^2 \right) \qquad (2.11)$$

where $\bar{y}$ is the mean of the elements of Y_r. R_k^2 for the first two components of the decomposition of the England and Wales fertility data (O.P.C.S., 1979) is given in panel (c) of table 2.1.

Missing values in the data matrix Y_r can be dealt with by omitting the rows or columns which contain them, and analysing only the remaining sub-matrix. In general the biplot is not greatly influenced by this loss of information provided that the proportion of omitted data is not large. Alternative a weighted analysis may be performed (Gabriel and Zamir, 1979), although the S.V.D. cannot then be used.

2.2.3 Interpretation.

The uniqueness of the biplot in figure 2.1 was obtained by making both the 'least-squares' requirement of equation 2.2 and the 'dispersion' requirement of equation (2.10). Had either of these requirements been different then a different biplot would have been produced. Consequently these requirements should constantly be borne in mind when interpreting features of the biplot. Any alternative to the 'least-squares' requirement would correspond to a series of reflections in and rotations about the biplot axes of the row and column markers. Any alternative to the 'dispersion' requirement would correspond to scaling the row and column markers in the direction of the biplot axes, such that for any given axis the scaling of the row markers is the inverse of the scaling of the column markers.

It is useful to note a few features of the biplot which remain invariant under reflections, rotations and scalings about the biplot axes. Such features may be interpreted without regard to the particular identification requirements made. They are:

i) linearity within sets of row markers and within sets of column markers;

ii) orthogonality between sets of row markers and sets of column markers;

iii) inner products between row and column markers (that is $\underline{g}_i' \underline{h}_j$).

Inner products between row and column markers are useful because they reproduce the data, since it follows from (2.1) that:

$$y_{ij} = \underline{g}_j' \underline{h}_j \qquad (2.12)$$

where y_{ij} is the $(i, j)^{th}$ element of Y_r. Inner products can be constructed on the biplot, as is shown in figure 2.2. Dropping the perpendicular from the row marker $\underline{g}_i$ onto the position vector of column marker $\underline{h}_j$ gives the projected length of $\underline{g}_i$ on $\underline{h}_j$ labelled d in figure 2.2. If $\underline{g}_i$ projects onto the negative direction of $\underline{h}_j$ then d should be negated. The length of position vector $\underline{h}_j$ multiplied by d gives the inner product $\underline{g}_i' \underline{h}_j$. (Of course the same value results if the length of $\underline{g}_i$ is multiplied by the projected length of $\underline{h}_j$ on $\underline{g}_i$). This technique is particularly useful for <u>comparing</u> elements of Y_r in the same row or column, since no multiplication need then be carried out. For example, from figure 2.2 it is clearly seen that $y_{i'j} < y_{ij}$ since the projected length of $\underline{g}_{i'}$ on $\underline{h}_j$ is less than that of $\underline{g}_i$ on $\underline{h}_j$.

Another useful property of the biplot is that linear combinations of the elements in a column of Y_r can be represented on the biplot since by (2.12):

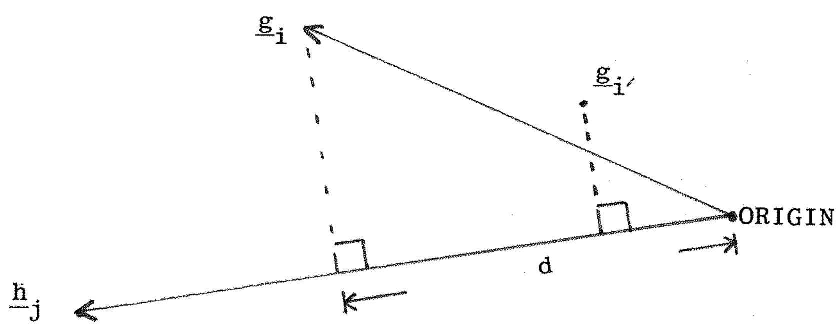


Figure 2.2

$$\sum_{i=1}^{m} \alpha_i \, y_{ij} = \sum_{i=1}^{m} \alpha_i \, \underline{g}_i' \, \underline{h}_j$$

$$= \underline{g}_\alpha' \, \underline{h}_j \qquad (2.13)$$

where $\underline{g}_\alpha$ denotes $\sum_{i=1}^{m} \alpha_i \underline{g}_i$, and may be plotted as a row marker. Linear combinations of the elements in a row of Y_r can be treated similarly.

Equations (2.12) and (2.13) apply exactly if all r components of the markers are used. However, when utilising only the first k components, as in the rank-k approximate biplot, (2.12) and (2.13) apply exactly to Y_k and only approximately to Y_r.

The R_k^2 statistic given in equation (2.11) may assist in deciding upon the appropriate number of components, k. To this end it is also helpful to examine residuals and higher order components in order to detect systematic or meaningful aspects of the data which are not included in the first few components. When the first three components fail to capture all the important variation in the data, it is sometimes helpful to partition the data matrix, and biplot each part separately. In such cases the biplot of the first few components would indicate how the partitioning should be arranged.

A detailed examination of the biplot in figure 2.1 of the fertility data for England and Wales (O.P.C.S., 1979) is included in the following section.

2.3 Explicating Trends in a Demographic Table.

Regular trends across the dimensions of a demographic table are only implicitly represented in the table since they have to be constructed by eye. Plotting the raw data in some suitable way may assist the eye in its task of discerning trends, but when the table is large this is likely to lead to a confusing tangle of lines. It is therefore desirable to find a way of explicitly representing the trends across each dimension, ideally by means of a single plotted line for each dimension. The biplot explicates such trends in precisely this way.

The biplot in figure 2.1 of the decomposition given in table 2.1 presents explicitly the trends in the age and period dimensions of the fertility data for England and Wales (O.P.C.S. 1979). The raw data themselves are only implicitly represented in the biplot. It is important, however, that the raw data may be reconstructed from the biplot in order that the trends in the biplot may be interpreted in terms of the raw data. Strictly the first two components of the biplot only contain information on a rank 2 approximation to the data matrix; however the high R_2^2 value in table 2.1 indicates that little information is lost by restricting attention to the first two components. Two techniques for interpreting trends in the biplot are now illustrated using figure 2.1.

Firstly, the technique of data reconstruction given in section 2.2 above may be utilised. For example, the age structure of fertility in 1944 may be visualised by dropping the perpendiculars from each of the age-markers onto the position vector of the 1944 period marker, in the manner of figure 2.2 above. It will be seen in particular that for 1944 the modal age of fertility is at 25 years, that between ages 18 and 22 there is a very rapid rise in fertility, and only a slight decline between ages 25 and 32 years. Repeating the procedure for 1974 reveals a modal age at 24 years, a rapid rise in fertility between ages 15 and 21, and now a quite marked decline between ages 25 and 34 years. Repeating for other time-periods will confirm the impression that moving between period markers in a clockwise direction about the origin describes a gradual shift in the fertility distribution away from the higher ages and towards the lower ages. Thus it may be seen that the period markers describe trends towards younger fertility from 1938 to 1941, and from 1944 to 1974, and trends towards older fertility from 1941 to 1944 and from 1974 to 1979. Between 1941 and 1946 the trend is somewhat erratic, and between 1946 and 1949 the trend is quite rapid, but elsewhere period trends are moderate and remarkably regular.

The distance from the origin of a period marker gives an indication of the overall level of fertility in that time-period. For example, since the 1967 and 1979 period markers lie approximately on a straight line through the origin, they have approximately the same age-structure of fertility. The distance of the 1967 marker from the origin is about 1.4 times that of 1979 marker, which indicates, as discussed in connection with figure 2.2 above, that for each age group fertility in 1967 was about 1.4 times that for 1979. Comparing time-periods with different age-structures of fertility requires some summary measure of fertility such as the Total Period Fertility Rate (T.P.F.R.). The use of the T.P.F.R. in conjunction with the biplot will be discussed in the next section.

The same information may be viewed from a different perspective by projecting each period marker onto the position vector of a given age marker, in order to give an impression of the 'period structure' of fertility for that age. Doing this, for example, for each age up to 19 years reveals, for these ages, that at any time after 1953 fertility was higher than at any time before 1953. Repeating for each age after 30 years reveals that fertility was higher in 1947 than at any other time for these ages. It is interesting to note that, excluding age 15, there are no two age markers forming a straight line through the origin; consequently none of these ages have the same period structure.

A second technique for interpreting biplot trends, which may be utilised when the 'least squares' criterion (2.2) holds is to examine each biplot component in turn. From table 2.1 or figure 2.1 it is seen that the basic age-structure of fertility is one of rapidly increasing fertility up to age 24 years followed by a more gradual decline. The second age component may be viewed as an adjustment to this basic age-structure, being a contrast between those ages before 27 years and those after, with the greatest weights being given to ages around 19 and 33 years. The first period component shows a basically low level of fertility before 1946 and after 1971, and a high level of fertility between 1960 and 1967. The second period component shows that the adjustments to the basic age-structure were most extreme in an absolute sense in 1946 and 1971.

2.4 Axes for level and structure.

Demographic data classified by an age variable are frequently summarised using measures of level and structure across ages. For example, for the data analysed in section 2.3, these might be the Total Period Fertility Rate (T.P.F.R.) and the Mean Period Fertility Age (M.P.F.A.) It would be useful to be able to represent such summary measures on the biplot, and it is shown in this section that axes may be drawn on the biplot allowing these measures to be read off. The development below is conducted in terms of a measure of level (LEV_j) and structure (STR_j) for the j^{th} column constructed analogously to the T.P.F.R. and M.P.F.A. respectively. Similar measures for rows rather than columns may be dealt with by transposing the roles of the rows and columns. Techniques for graphical representation of other summary measures which are simple functions of linear combinations of the elements in rows or columns may be devised.

As shown in equation (2.13), linear combinations of the elements in a row or column may be represented on the biplot. Define:

$$LEV_j = \sum_i y_{ij}$$

By (2.12) this is

$$= \sum_i \underline{g}_i' \underline{h}_j$$

$$= \underline{g}_{\cdot}' \underline{h}_j \qquad (2.14)$$

where $\underline{g}_{\cdot}$ denotes the row marker $\sum_i \underline{g}_i$. Drawing $\underline{g}_{\cdot}$ on the biplot would enable the values of LEV_j for each j to be reconstructed, using the technique described in section 2.2 in connection with figure 2.2. This technique involves a certain amount of mental arithmetic; however this may be done away with by means of a slight development of the technique. The vectors $\underline{g}_{\cdot}$ and $\underline{h}_j$ have been drawn in figure 2.3. The position vector of the foot of the perpendicular from $\underline{h}_j$ to $\underline{g}_{\cdot}$, denoted by $\underline{h}_{j\cdot}$ is given by:

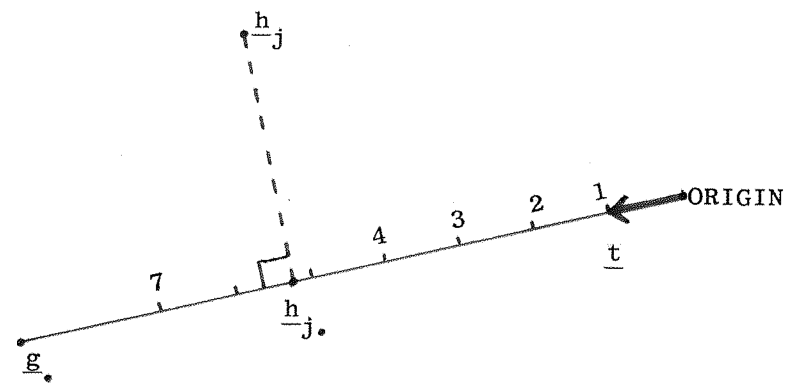


Figure 2.3

$$\underline{h}_{j.} = \underline{g}_.'\underline{h}_j\ (\underline{g}_.'\underline{g}_.)^{-1}\ \underline{g}_.$$
$$= \text{LEV}_j\ \underline{t} \qquad (2.15)$$

using (2.14), where:

$$\underline{t} = (\underline{g}_.'\underline{g}_.)^{-1}\ \underline{g}_. \qquad (2.16)$$

Thus, from (2.15), if an axis is drawn through the origin in the direction of $\underline{g}_.$, and is scaled in units equal to the length of $\underline{t}$, with the value 0.0 appearing at the origin, then the value of LEV_j may be read off this axis at the foot of its perpendicular from $\underline{h}_j$.

Define

$$\textbf{STR}_j = \sum_i \textbf{a}_i\ \textbf{y}_{ij} / \sum_i \textbf{y}_{ij}$$

where a_i denotes the value of the interval-level variable defining the row categories, at the mid-point of the i^{th} row,

$$= \sum_i \textbf{a}_i\ \underline{\textbf{g}}_i'\ \underline{\textbf{h}}_j/\textbf{LEV}_j$$

using (2.12) and (2.14)

$$= \underline{g}_*'\underline{h}_j/\text{LEV}_j \qquad (2.17)$$

where $\underline{g}_*$ denotes the row marker $\sum_i a_i\ \underline{g}_i$. Again, drawing $\underline{g}_*$ on the biplot would enable STR_j for each j to be reconstructed using the technique described in section 2.2, but a development of the technique allows the mental arithmetic involved to be done away with. Figure 2.4 contains the vectors $\underline{g}_.$ and $\underline{g}_*$. In general, the period marker $\underline{h}_j$ will not be in the plane of $\underline{g}_.$ and $\underline{g}_*$, and so its projection $\underline{h}_{j.*}$ onto that plane is drawn in figure 2.4. The position vector of $\underline{h}_{j.*}$ is given by:

$$\underline{h}_{j.*} = A(A'A)^{-1}A'\underline{h}_j$$

where A is the rx2 matrix $\left[\underline{g}_.,\ \underline{g}_*\right]$

$$= A(A'A)^{-1} \begin{bmatrix} 1 \\ \text{STR}_j \end{bmatrix} \text{LEV}_j$$

using (2.14) and (2.17),

$$= \frac{\text{LEV}_j}{c}\ (\underline{w} + \text{STR}_j\ \underline{s}) \qquad (2.18)$$

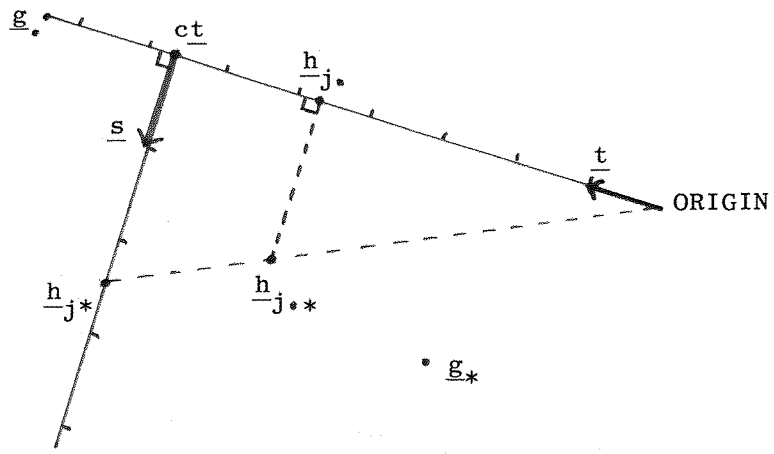


Figure 2.4

where

$$[\underline{w}\ ,\ \underline{s}] = cA(A'A)^{-1} \tag{2.19}$$

for some arbitrary scalar value c. It may easily be verified from (2.19) that $\underline{s}$ is orthogonal to $\underline{g}_.$. Now let $\underline{h}_{j*}$ be the point lying on the radius from the origin through $\underline{h}_{j.*}$ which scores a value c on the $\underline{t}$-axis (see figure 2.4). Then the position vector of $\underline{h}_{j*}$ is given by:

$$\underline{h}_{j*} = \underline{w} + STR_j\ \underline{s}$$

using (2.18)

$$= c\underline{t} + (STR_j - \underline{g}_*{}'\ \underline{t})\underline{s} \tag{2.20}$$

using (2.16) and (2.19). Thus, from (2.20), if an axis is drawn through the point $c\underline{t}$ in the direction of $\underline{s}$, and is scaled in units equal to the length of $\underline{s}$, with the value $(\underline{g}_*{}'\ \underline{t})$ appearing at the point $c\underline{t}$, then the value of STR_j may be read off this axis where it meets the radius from the origin through $\underline{h}_{j.*}$.

The $\underline{t}$- and $\underline{s}$- axes defined above are based on r-dimensional vectors $\underline{g}_.$ and $\underline{g}_*$, and in principle permit values of LEV_j and STR_j to be read from the exact biplot of Y_r. For the rank-k approximate biplot of Y_r it is a natural extension of the technique to construct axes which relate to values of LEV_j and STR_j corresponding to Y_k rather than Y_r. This may be achieved by using only the first k components of $\underline{g}_.$ and $\underline{g}_*$ in the calculation of vectors $\underline{t}$ and $\underline{s}$ in equations (2.16) and (2.19). It is important to note that these k-component versions of $\underline{t}$ and $\underline{s}$ are not in general the same as the first k components of the r-component versions of $\underline{t}$ and $\underline{s}$. In practice the $\underline{t}$- and $\underline{s}$- axes are most likely to be of use in the case when k = 2, since in the higher dimensions it is difficult to visually construct orthogonal projections. For k = 2, the points $\underline{h}_j$ and $\underline{h}_{j.*}$ are identical; this means that radii from the origin are taken through the column markers themselves when reading off the values of STR_j.

The 2-component versions of $\underline{g}_.$, $\underline{g}_*$, $\underline{t}$ and $\underline{s}$ for the fertility data for

England and Wales (O.P.C.S., 1979) are shown in table 2.2, and the $\underline{t}$- and $\underline{s}$- axes have been drawn in figure 2.1. The value of the scalar c in equation (2.19) has been set to 1.5, as this has the effect of placing the $\underline{s}$-axis in a convenient position in this biplot. It may be seen from figure 2.1, to a rank-2 approximation, for 1946 for example, that the T.P.F.R. is 2.45 births per woman, and the M.P.F.A. is 29.2 years.

Table 2.2. The 2-component versions of vectors $\underline{g}_{\bullet}$, $\underline{g}_*$, $\underline{t}$ and $\underline{s}$ for the decomposition in table 2.1

$\underline{g}_{\bullet}$	$\underline{g}_*$	$\underline{t}$	$\underline{s}$
-7.48364	-206.592	-.133017	.0052122
-0.50607	- 33.434	-.008995	.0770676

2.5 Spline interpolation of biplot markers.

Although the age markers in figure 2.1 have been connected using straight lines, the overall impression is that of a smooth curve. Age dimensions in demographic tables of rates or proportions typically exhibit this property of smoothness when the number of age categories is large. When the number of age categories is small, however, as for example with five-year age-specific fertility rates for ages 15 to 50 years, it may be desirable to connect the corresponding markers with a smooth line rather than a series of straight lines. It is shown below that this may be conveniently done using a set of cubic splines.

Any set of points $\{(a_i, y_i): i = 1, \ldots, m\}$ where the $\{a_i\}$ are strictly increasing, may be smoothly connected with a cubic-spline function, $y(a)$, as in figure 2.5. The $\{a_i : i = 1, \ldots, m\}$ are termed the 'knots' of the spline, and the spline is derived so that for each interval between adjacent knots $y(a)$ is a cubic polynomial, and for each open interval ($a \lesssim a_1$ or $a \gtrsim a_m$) $y(a)$ is a straight line. The coefficients of these $(m - 1)$ cubic polynomials and two straight lines are chosen so as to ensure continuity in $y(a)$ and $y'(a)$ at each of the knots, and also continuity in $y''(a)$ at all but the first and last knots. These requirements are sufficient to uniquely define the cubic spline. McNeil et al (1977) describe how the spline may be calculated, and algorithms for doing so are commonly available in scientific software libraries.

Let $u_\ell(a)$ be the cubic spline passing through the points $\{(a_i, u_i) : i = 1, \ldots, m\}$ where a_i is the value at the mid-point of the i^{th} row of the variable defining row categories, and where $u_{i\ell}$ is the i^{th} element of $\underline{u}_\ell$ in equation (2.1). Denote:

$$y_j(a) = \sum_{\ell=1}^{r} u_\ell(a) v_{j\ell} \tag{2.21}$$

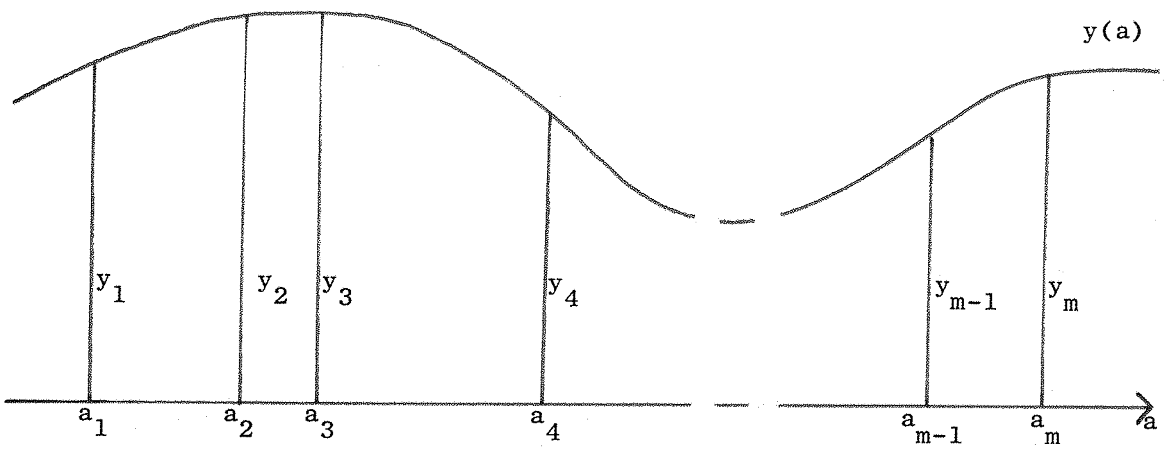


Figure 2.5 A cubic spline

where $v_{j\ell}$ is the j^{th} element of $\underline{v}_\ell$ in equation (2.1). Since, for each i, $u_\ell(a_i)$ equals $u_{i\ell}$ then $y_j(a_i)$ equals y_{ij} by equation (2.1). Moreover, the continuity and polynomial properties of the $\{u_\ell\ (a)\}$ are conferred on $y_j(a)$ by virtue of (2.21). Therefore $y_j(a)$ is the cubic spline passing through the points $\{(a_i, y_{ij}) : i = 1, \ldots, m\}$.

Now, for any a, a vector $\underline{g}_a$ of order r may be defined so that its ℓ^{th} element is equal to $u_\ell\ (a)$. Then from (2.21):

$$y_j(a) = \underline{g}_a{}'\ \underline{h}_j \tag{2.22}$$

using the definition of $\underline{h}_j$. Thus the inner product of a period marker with $\underline{g}_a$ gives the cubic spline interpolated value at a in the j^{th} column of Y_r. Moreover, inner products involving only the first k elements of $\underline{g}_a$ give interpolated values for Y_k rather than Y_r. Thus $\underline{g}_a$ is a row marker, and $\underline{g}_a$ may be plotted on the biplot for each of a large number of values of a. The $u_\ell\ (a)$ splines from which the $\underline{g}_a$ markers are derived may be calculated as described by McNeil et al (1977).

Column markers $\underline{h}_b$ may be defined, calculated and biplotted similarly for each of several values of b, the variable defining column categories. The inner product $\underline{g}_a{}'\ \underline{h}_b$ then produces the value of the bi-cubic spline surface y(a,b) which passes through the elements of Y_r (or Y_k if only the first k components are used) at the cell mid-points.

It is interesting to note that, although a cubic-spline is a single-valued function of its argument (a), the resulting $\underline{g}_a$ or $\underline{h}_b$ is not necessarily single-valued on any of the biplot component axes. For example, for figure 2.1, neither $\underline{g}_a$ nor $\underline{h}_b$ would be single valued on either of the first two component axes.

Table 2.3 contains the biplot markers for five-year age-specific legitimate fertility rates for several populations (United Nations, 1965 and 1975). According to the theory above, in order to connect the seven age markers on the 2-component biplot with a smooth line, two cubic

Table 2.3 The first two components of the decomposition of an array of five-year age-spcific legitimate live-birth rates for various populations (United Nations, 1965 and 1975). The rows of panels (a) and (b) form the first two components of the row (age) and column (period x country) markers, respectively. Panel (c) contains the values of the R_k^2-statistic (equation (2.11)) for the first 2 components.

(a)

Age (completed years)	$\underline{u}_1$	$\underline{u}_2$
15-19	-.910	.335
20-24	-.670	-.133
25-29	-.478	-.226
30-34	-.293	-.227
35-39	-.172	-.199
40-44	-.066	-.099
45-49	-.011	-.026

(b)

Country x	Period	$\underline{v}_1$	$\underline{v}_2$	Group
Australia	1961	-.562	.123	A
	1971	-.463	.094	A
Canada	1961	-.573	.056	A
	1971	-.351	.086	A
England & Wales	1964	-.477	.116	A
	1973	-.344	.127	A
Holland	1963	-.542	.037	A
	1974	-.292	.014	A
Scotland	1964	-.514	.091	A
	1974	-.365	.165	A
Belgium	1961	-.425	.042	B
	1970	-.405	.155	B
France	1963	-.489	.042	B
	1972	-.465	.168	B
Luxembourg	1960	-.412	.065	B
	1970	-.380	.163	B
Denmark	1963	-.524	.265	C
	1973	-.337	.087	C
Sweden	1963	-.473	.216	C
	1974	-.381	.104	C

Table 2.3 - (b) continued ....

Country x	Period	$\underline{v}_1$	$\underline{v}_2$	Group
Austria	1961	-.543	.174	D
	1974	-.414	.266	
Finland	1963	-.507	.133	
	1973	-.373	.236	
Switzerland	1960	-.569	.191	
	1970	-.459	.198	
Czechoslovakia	1961	-.440	.215	E
	1970	-.436	.289	
Norway	1960	-.471	.027	
	1974	-.461	.284	
Poland	1960	-.505	.128	
	1974	-.518	.276	
Cyprus	1960	-.440	-.176	F
	1973	-.500	.083	
Portugal	1960	-.491	-.096	
	1973	-.424	-.042	
Spain	1960	-.497	-.183	
	1970	-.507	-.096	
El Salvador	1962	-.607	-.364	G
	1971	-.561	-.191	
Panama	1960	-.561	-.115	
	1969	-.537	-.096	
Venezuela	1961	-.634	-.321	
	1971	-.559	-.147	
Greenland	1960	-.646	-.394	H
	1970	-.399	-.029	
Hungary	1963	-.304	.111	
	1974	-.394	.159	
Ireland	1961	-.734	-.222	
	1971	-.740	-.020	
Japan	1960	-.398	-.083	
	1970	-.364	-.183	
Macau	1960	-.580	-.690	
	1970	-.280	-.262	
Phillippines	1960	-.360	-.280	
	1970	-.446	-.269	

(c)

R_1^2	R_2^2
.927	.991

splines should be calculated, one for each column of panel (a) of table 2.3. At ages, say, one year apart both splines may be evaluated, and the two values which result for each age may be plotted as an age marker on the biplot, and the resulting age markers joined up. Inner products of these single-year age markers with a column marker would produce estimates of single-year age-specific legitimate fertility rates for the corresponding population. However, these may not be the best available estimates of single-year rates, since it is reasonable to insist that the average of the single-year rate estimates within each five-year age group should equal the corresponding five-year rate estimate, which implies that the average of the single-year age markers within each five-year age group should equal the corresponding five-year age marker, and in general this will not be the case for the spline-estimated age markers. To remedy this, the five-year age markers should be incremented by the amount which they exceed the average of the corresponding spline-estimated single-year age markers, and the splines should then be recalculated to pass through these new five-year age markers. The process should continue iteratively, at each iteration incrementing the <u>current</u> five-year age markers by the amount which the <u>original</u> five-year age markers exceed the average of the current estimates of the corresponding single-year age markers, until convergence. For the age markers in panel (a) of table 2.3, the process took just 3 iterations. The resulting single-year age markers are given in table 2.4, and it may be verified that although these markers do not pass through those in panel (a) of table 2.3, (for example, for age-group 30 - 34 the marker is (-.293, -.227) whereas for age 32 the marker is (-.289, - .227) they do possess the desired averaging property. Figure 2.6 contains the biplot corresponding to table 2.3, in which the single-year age markers are plotted.

Table 2.4. Cubic splines $u_1(a)$ and $u_2(a)$ based on the $\underline{u}_1$ and $\underline{u}_2$ vectors in panel (a) of table 2.3, evaluated at single year intervals of age. The rows in the table form the first two components of the $\underline{g}_a$ row (age) markers.

Age (a) (completed years)	$u_1(a)$	$u_2(a)$
15	-1.012	0.568
16	-0.961	0.454
17	-0.910	0.335
18	-0.858	0.217
19	-0.807	0.103
20	-0.757	-0.001
21	-0.710	-0.088
22	-0.667	-0.156
23	-0.626	-0.199
24	-0.589	-0.222
25	-0.553	-0.230
26	-0.517	-0.230
27	-0.480	-0.226
28	-0.440	-0.224
29	-0.400	-0.223
30	-0.360	-0.224
31	-0.322	-0.226
32	-0.289	-0.227
33	-0.261	-0.228
34	-0.236	-0.227
35	-0.215	-0.224
36	-0.194	-0.216
37	-0.173	-0.204
38	-0.150	-0.187
39	-0.126	-0.165
40	-0.103	-0.141
41	-0.081	-0.118
42	-0.062	-0.096
43	-0.047	-0.077
44	-0.035	-0.062
45	-0.026	-0.049
46	-0.018	-0.037
47	-0.011	-0.026
48	-0.004	-0.015
49	0.004	-0.003

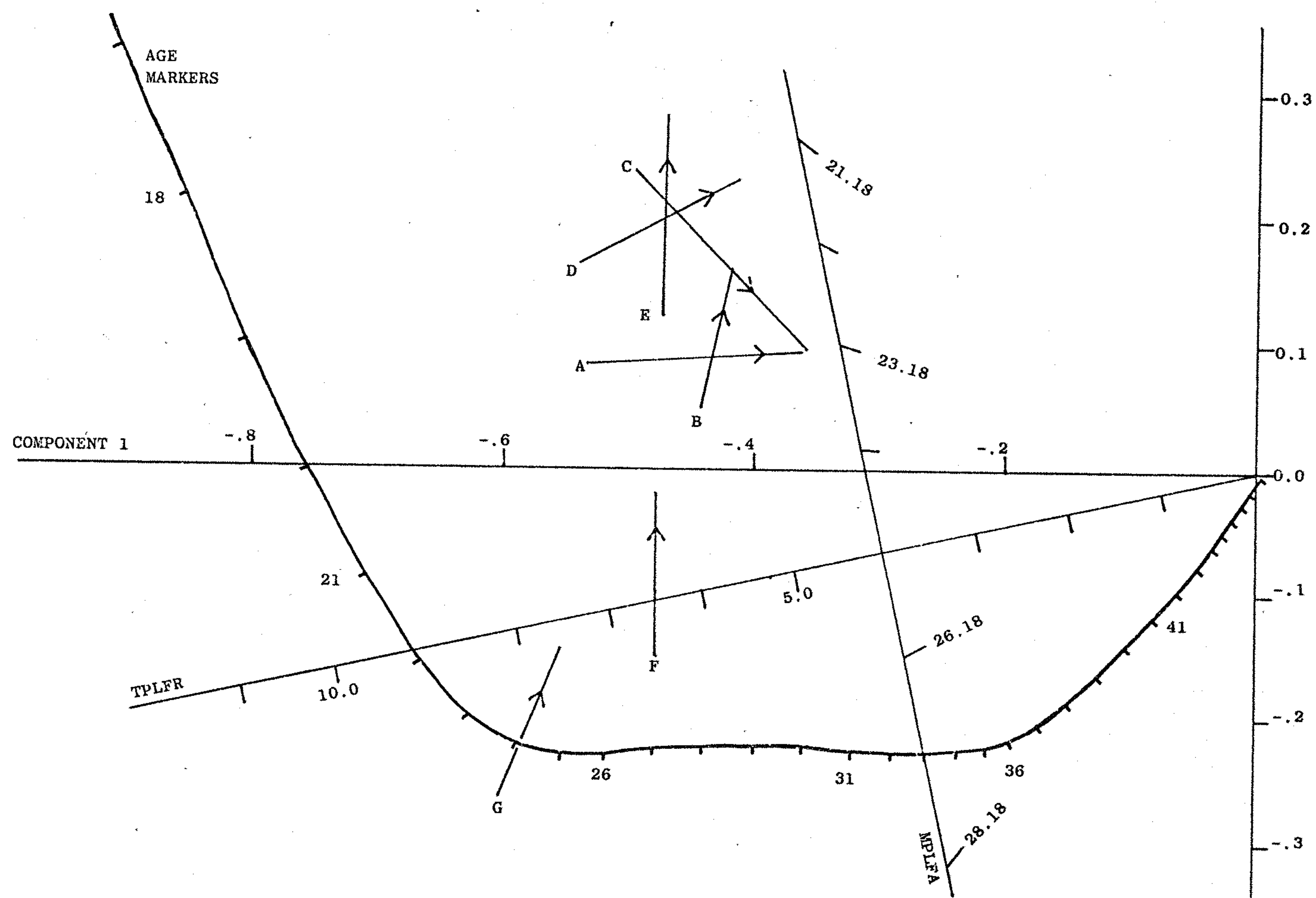


Figure 2.6 Biplot of five-year age-specific legitimate live birth rates for various populations, (United Nations, 1965 and 1975) showing the movement between two time-periods for each group of countries defined in table 2.3(b), and the spline-estimated single-year age markers (see section 2.5). Also showing the T.P.L.F.R. and M.P.L.F.A. axes (see section 2.6). Obtained from the decomposition in table 2.3.

2.6 The Biplot and Model Schedules.

Models of demographic schedules, such as the Coale-Tussell model of age-specific fertility (Coale and Trussell,1974), are of considerable value to the demographer, particularly in connection with incomplete or unreliable data. Fitting models to observed schedules requires only a suitable computer program, but at the time of writing there appears to be no systematic approach reported in the demographic literature for deriving these models.

Typically, model schedules are required in only one or two continuous dependent variables, and are based on census and registration data from a wide variety of populations. There are good reasons for not attempting to build a model from elaborate substantive hypotheses, since, as may be seen in chapter 3, the data would in general be incapable of properly verifying it, and a simple formulation would be as likely to fit the observed schedules well. Moreover, a large data base and underlying continuity in the dependent variables ensure a high degree of regularity in the observed schedules; which means that very close approximations to the observed schedules should be attainable by the model. Thus a good model should be simple in structure, unambitious in its substantive interpretation and provide a very close fit to the data.

A number of different models of demographic schedules have been reported in the literature, but each of these have been developed for a particular type of schedule, and there is no reason to suppose that any of them would be transportable to other types. The decomposition (2.2) of the biplot, however, provides a general method of generating models for all types of schedule, as is illustrated below for the case of age-specific legitimate fertility. Such a method might be criticised for being atheoretical, that is, paying no attention to substantive issues. However, as argued above, there is little point in constructing sophisticated theoretical models, and there is no guarantee that the parameters in a crudely constructed theoretical model may be validly interpreted in the intended way. On the other hand, the parameter estimates from a well fitting atheoretical model may be interpreted flexibly and powerfully in conjunction with supplementary information about the populations concerned.

If a variety of observed schedules of a given type are arranged in columns, with the rows corresponding to levels of the independent variable(s), then a model schedule $\underline{y}_k$ of this type may be given by:

$$\underline{y}_k = \beta_1 \underline{u}_1 + \beta_2 \underline{u}_2 \cdots + \beta_k \underline{u}_k \tag{2.23}$$

where the vectors $\underline{u}_1$ to $\underline{u}_k$ are obtained from the decomposition (2.2) of the resultant data matrix, Y_r, and where $\underline{\beta}$ is a set of k model parameters. Model (2.23) may be fitted by least-squares to any observed schedule $\underline{y}$ which need not be amongst those in Y_r: if, however, $\underline{y}$ is the j^{th} column of Y_r then it follows from (2.2) that the estimated $\underline{y}_k$ is the j^{th} column of Y_k and that the estimated $\underline{\beta}$ are the first k components of the j^{th} column marker. Note that the regression model (2.23) does not contain an 'intercept' term. Now it follows from equations (2.4) and (2.7) that the $\underline{u}$ vectors are mutually orthogonal, and so the usual regression formula for calculating parameter estimates simplifies to:-

$$\hat{\beta}_\ell = \underline{u}_\ell{}'\underline{y}/\underline{u}_\ell{}'\underline{u}_\ell \tag{2.24}$$

for $\ell = 1, \ldots, k$, which is simple enough to permit calculation by hand. The estimated vector $\hat{\underline{\beta}}$ may be plotted as a column marker on the k-component biplot. The R^2 value from the regression model (2.23) on an observed schedule $\underline{y}$ would in general be expected to be slightly lower than the R_k^2 value from the decomposition (2.2) of Y_r, since the regression model (2.23) contains no row parameters.

Use of the biplot decomposition to obtain model schedules is now illustrated for the case of five-year age-specific legitimate fertility rates. Observed schedules from 28 countries each at two time-periods were selected (United Nations, 1965 and 1975) to provide a basis for the calculations of the standard $\underline{u}$ vectors. The observed schedules were arranged in columns and the decomposition (2.2) of the resultant data matrix is given in table 2.3, and the biplot in figure 2.6. To avoid presenting a multiplicity of column markers on the biplot, countries having similar pairs of markers have been grouped together as indicated in panel (b) of table 2.3, and only the mean markers for the earlier and later time-periods have been plotted (joined with an arrow to indicate the direction of time) for each country group A to G. Countries which could not be grouped in this way constitute group H. The biplot also contains Total Period Legitimate Fertility Rate (T.P.L.F.R.) and Mean Period Legitimate Fertility Age (M.P.L.F.A.) axes, calculated as described in section 2.4, although the scale on the T.P.L.F.R. axis reflects the fact that the T.P.L.F.R. is <u>five times</u> the sum of the five-year age-specific rates. As described in section 2.5, the biplot also includes spline estimates of single-year of age markers.

The high R_2^2 value of .991 in panel (c) of table 2.3 suggests that model schedules which approximate well a wide variety of observed schedules may be generated using only $\underline{u}_1$ and $\underline{u}_2$ in equation (2.23), whose values are given in panel (a) of table 2.3. To test this, a further eight schedules of five-year age-specific legitimate fertility (United Nations, 1965) together with the standard schedule of 'natural fertility' (Henry, 1961; Coale and Trussell, 1974) were regressed on the $\underline{u}$ vectors in panel (a) of table 2.3, with the results given in table 2.6. The overall R^2 value is .982, which is encouraging, although the natural fertility schedule is not fitted well. The $\underline{\beta}$ estimates in table 2.6

Table 2.5. The 2-component versions of vectors $\underline{g}_{.}$, $\underline{g}_{*}$, $\underline{t}$ and $\underline{s}$ for the decomposition in table 2.3.

$\underline{g}_{.}$	$\underline{g}_{*}$	$\underline{t}$	$\underline{s}$
-2.59942	-63.426	-.366763	.018440
-0.57487	-23.622	-.081111	-.083380

Table 2.6. Parameters estimates from models (2.23) and (2.25) of 8 schedules of five-year age-specific legitimate fertility rates (United Nations, 1965) and the standard schedule of natural fertility (Henry, 1961; Coale and Trussell, 1974, appendix A).

		Model (2.23) Regression on $\underline{u}$ vectors in table 2.3(a)			Model (2.25) Coale-Trussell model of legitimate fertility		
Country	Period	β_1	β_2	R^2	γ_1	γ_2	R^2
New Zealand	1961	-.627	.062	.998	1.222	1.823	.945
Greece	1961	-.373	-.127	.975	.685	1.142	.996
West Germany	1963	-.510	.198	.993	1.053	2.642	.854
Bulgaria	1956	-.294	-.006	.949	.596	1.963	.999
Chile	1960	-.697	-.085	.985	1.232	1.085	.897
Peru	1961	-.503	-.259	.980	.832	.576	.965
S. Africa (White)	1960	-.543	.055	.999	1.058	1.833	.941
S. Africa (Negro)	1961	-.680	-.297	.989	1.149	.704	.965
Natural Fertility		-.667	-.695	.893	1.000	.000	1.000

may be plotted as column markers on figure 2.6. Estimates of single-year age-specific legitimate fertility rates for these schedules may be obtained as inner products of the corresponding column markers and the spline-estimated single-year age markers.

It is interesting to compare these results with those for the Coale-Trussell model of age-specific legitimate fertility (Coale and Trussell, 1974) on the same data. This model may be written:

$$y_{CTi} = \gamma_1 w_{1i} \exp(\gamma_2 w_{2i}) \tag{2.25}$$

where y_{CTi} denotes a Coale-Trussell model value for the legitimate fertility rate for the i^{th} five-year age group, w_{1i}, and w_{2i} are the standard values of natural fertility and departure from natural fertility for the i^{th} five-year group, provided in Coale and Trussell (1974) appendix A, and γ_1 and γ_2 are the model parameters. The results from fitting model (2.25) by least-squares to the nine schedules in table 2.6 are also given in table 2.6, and the overall R^2 is .947. The models fitted in table 2.6 have the same number of degrees of freedom, so, for these data at least, model (2.23) is much more successful than the Coale-Trussell model (2.25), overall accounting for 66% of the amount of variation which the Coale-Trussell model fails to explain. The success of model (2.23) is probably due to the fact that both standard age-structures, $\underline{u}_1$ and $\underline{u}_2$, reflect a variety of legitimate fertility age schedules, whereas in the Coale-Trussell model (2.25) one of the standard age-structures (w_{1i}) reflects only one, rather extreme, legitimate fertility age schedule.

It is also interesting to explore the algebraic links between models (2.23) and (2.25). Taking only the first two terms in the Taylor expansion of $\exp(\gamma_2 w_{2i})$ gives:

$$y_{CTi} \simeq \gamma_1 w_{1i} (1 + \gamma_2 w_{2i})$$

whence

$$\underline{y}_{CT} \simeq \gamma_1 \underline{w}_1 + \gamma_{12} \underline{w}_{12} \tag{2.26}$$

where $\underline{y}_{CT}$, $\underline{w}_1$ and $\underline{w}_{12}$ are vectors whose i^{th} elements are y_{CTi}, w_{1i} and $w_{1i}.w_{2i}$ respectively, and where $\gamma_{12} = \gamma_1\gamma_2$. Comparing expressions (2.23) and (2.26) it is seen that the two models are approximately of the same algebraic structure.

2.7 Conclusions.

The biplot can be used to provide a graphical summary of any two-way data table. When the data are demographic rates or proportions, the biplot brings out the underlying regularity which is often present in such data; in particular 'age' dimensions usually appear as smooth curves in the biplot, as in figure 2.1. This regularity in the data also leads to very good approximations to the data with just two biplot components, although further components may still contain systematic trends of interest. Tables of greater than two dimensions can be dealt with by combining two or more dimensions: for example the biplot in figure 2.6 was produced by combining the time-period and country dimensions to form the dimension listed in table 2.3(b).

When the data table contains age dimensions it is helpful to construct axes for 'level' and 'structure' on the biplot as detailed in section 2.4. Smooth curves representing the age dimensions may be constructed using cubic-splines, as described in section 2.5, and used for interpolation.

A particularly useful application of the biplot is in generating models of demographic schedules, using the decomposition from which the biplot is constructed. It is expected that the technique would work well for any 'age' structure of demographic rates or proportions, owing to the high regularity of such data, and to the flexibility of the standard age structures. In particular, in section 2.6, the technique is shown to work better than the model of Coale and Trussell(1974) for age-specific legitimate fertility rates. Deriving the standard age structures is easy, being a by-product of the biplot, and fitting the resulting model to further schedules is trivial, requiring only hand-calculation. The graphical representation of the results enhances this technique of model generation; calculation and plotting of level and structure axes, and of spline curves, increases the utility of this model generation technique still further.

Chapter 3 - Modelling vital rates.

3.1 Introduction

3.1.1 Objectives.

For the purposes of understanding and projecting demographic processes, tables of vital rates (or proportions) generally possess both advantages and disadvantages over individual-level survey data. On the one hand, published vital rates often extend over long periods of time (for example, the Swedish mortality data analysed in section 3.5 below), permitting the impact of a variety of social and economic conditions to be studied. On the other hand, long series of vital rates are frequently accompanied by only one cross-classifying dimension (usually age), and consequently the depth of reliable interpretation might be rather limited. The purpose of this chapter is to examine various models or classes of model in relation to a variety of demographic data, not merely to discover for each data set a suitable model, but principally to learn what to look for in a model, what to expect from a model, and how to approach the task of model building for highly aggregated demographic data.

As noted in Chapter 1, two quite different approaches to modelling, the theoretical and the descriptive, may be distinguished. The theoretical approach seeks to build a model from previous results, general experience , intuition and reasoning about factors thought to be underlying the data. The descriptive approach seeks only to provide a succinct algebraic description of the data, from which substantive insights may possibly result. In practice, a combination of these two approaches might well be adopted, alternately looking to the data for substantive clues, and combining theoretical considerations in a way which seems to correspond to the variation in the data. Conceivably it is possible that all approaches would lead to just one model,the extent to which this actually happens is one aspect of the present investigation.

With a limited depth of information available from the data, there seems little or no prospect of validating an elaborate theoretical model; equally there is no guarantee that a simple valid theoretical explanation exists. This suggests that a simple descriptive model might be more useful. However, the case against the existance of a simple valid theoretical model is not proven, and the extent to which theoretical considerations may usefully be incorporated into a model is a central concern of the development below.

Also of central importance is the way in which theoretical considerations may be expressed in a model. To explore this issue, four different types of model are examined in this chapter: model age-schedules, the additive age-period-cohort model, 'cohort-experience' models, and the 'bimodel'. These four types of model are discussed individually below. Some of these are more theoretical than others; that is, the strength of their substantive assertions varies considerably. Moreover, they represent a variety of techniques for incorporating substantive elements. Of course this set of types of model is not exhaustive; nevertheless, it is hoped that they are sufficiently diverse to enable general conclusions to be drawn. Equally, it is hoped that the data examined in this chapter (age-specific nuptiality rates, age-specific fertility rates, and age-by-marriage-duration-specific legitimate fertility rates for recent time-periods for England and Wales; and age-specific mortality rates for Sweden for a long time-series up to 1930) are also sufficiently diverse for present purposes.

3.1.2 Models

As noted in Chapter 0 above, the usual strategy for accommodating a number of unmeasured or unknown causal mechanisms into a model of vital rates is via period specific or cohort specific parameters in a model age-schedule derived either empirically (e.g. the marriage model of Coale and McNeil, 1972; or the relational models of Brass, 1974 a, b) or mathematically (e.g. the Gompertz function). Whether parameters in a model age-schedule should be made period-specific or cohort-specific has been a matter for some debate (Hobcraft et al , 1979; see also Chapter 0 above). Model age-schedules of nuptiality, fertility and mortality are considered below in order to shed light on this issue, and also to explore the possibility of assigning some parameters in a model age-schedule to periods and others to cohorts.

From a theoretical point of view it may be thought that both period-specific and cohort-specific mechanisms are operating in the data (in addition to age-related factors). The simplest way of combining all three types of factor is by means of additive age-period-cohort model (1.3) of Chapter 1. This model asserts that cohort factor effects are constant over all ages within each cohort; that period factor effects are constant over all ages within each period; and that age factor effects are constant over all time-periods (or cohorts) within each age group. Unfortunately the identification problem associated with this model severely hampers the interpretation of parameter estimates resulting from this model (see Chapter 1). This model may be generalised into a six-factor model for age-by-marriage-duration-by-period-specific legitimate fertility rates, as will be seen in section 3.4 below.

The identification problem of the additive age-period-cohort model is not its only drawback. Theoretical arguments might suggest that constant cohort or period factor effects over age are not realistic. Hobcraft et al (1979) advocate the use of models in which cohort factor effects are adjusted to reflect the cumulative experience of the cohort at each age. These models are termed 'cohort-inversion' (Hobcraft et al , 1979) or 'cohort-experience' (Hobcraft and Gilks , 1981) models. At least four types of cohort-experience mechanism may be distinguished. Firstly, for non-renewable events such as death or first-marriage, when there is heterogenous susceptibility to experiencing the event of interest, there is a tendency for the more susceptible individuals to be removed from the population earlier than the less susceptible, giving rise to a selection effect at the higher ages (Vaupel et al , 1979). This selection effect would vary with the amount of selection which has taken place, which in turn would vary between cohorts for any given age. Secondly, social and economic conditions in the history of a cohort may have an accumulating effect on the cohort's event rate; for example, epidemics and famines may permanently impair a cohort's vitality, giving rise subsequently to increased levels of mortality within the affected cohorts. Equally, some epidemics could actually strengthen a cohort's resistance to disease, producing subsequent decreases in its mortality. Thirdly, in the particular context of nuptiality, the proportion in a cohort already married could have a direct effect on the desire to marry amongst the non-married through their fear of being 'left on the shelf': this has been suggested by Hernes (1972). Fourthly, in the context of fertility, cohorts may aim at a 'target' completed fertility, so that the fertility at each age for a cohort would depend on the additional number of children required to reach the target, and on the favourability of current economic conditions towards childbearing (Lee, 1980). Cohort-experience models do not have a pre-defined algebraic form, unlike the additive age-period-cohort model, although a general framework for such models has been suggested by Hobcraft et al (1981). The principle is that theoretical considerations should dictate the form of the model.

Tables of vital rates often exhibit a transition from one cross-sectional age structure to another. Such a transition may be expressed algebraically as a sum of two age-schedules weighted by two period parameters: these two age-schedules may be fixed mathematical or empirical curves, although if the emphasis is on capturing the transition itself, there is no reason to fix these age schedules in advance. The resulting model, given in section 3.2.4 below, is equivalent to the algebraic structure of the biplot technique (Gabriel, 1971), discussed in Chapter 2 above. This model, termed the 'bimodel', is not merely descriptive since it asserts that, in addition to age factors, only time-period factors are in operation, via the two period weights in the model. A cohort factor bimodel may be constructed similarly, by replacing the two period weights with cohort parameters. The bimodel, however, cannot accommodate age, period and cohort factors simultaneously. The bimodel can be adapted to correspond to data arranged in more than two dimensions as will be seen in section 3.4 below.

The models described above represent a breadth of modelling strategies. The mathematically derived model age-schedules and the bimodel are more descriptive than the additive age-period-cohort and cohort-experience models, since they only assert the presence of period (or cohort) factors and do not explain why such factors should interact with age in the way specified by the model. All but the cohort-experience models have a pre-defined algebraic structure. All but the additive age-period-cohort model posit non-additive period or cohort factor effects. The model age-schedules and the bimodel do not accommodate both period and cohort factor effects simultaneously. These distinctions will be utilised in section 3.6 below in attempting to draw from the ensuing analysis general conclusions concerning modelling strategies.

3.1.3 Comparing models.

Before commencing the analyses of models and data, it is convenient to consider how models of the same data set might be compared.

The sample base for published tables of vital rates or proportions is usually the entire population. This has important consequences for the stochastic component of any model. Clearly, discrepancies between model and data cannot conveniently be attributed to sampling error; and conceiving of the population as a sample from a superpopulation is unlikely to account for anything but the most minute discrepancies owing to the generally very large 'sample' size. It is desirable however to have some way of conceiving of these discrepancies, if only to provide a systematic basis for estimating and comparing models, and in many situations it might be reasonable to assume that, in addition to the factors specified by the model, unknown factors act independently between cells and uniformly on individuals within cells of the table. Thus an error component may be added to the non-stochastic part of the model and assumed to be independently distributed across cells of the table. It is advisable to first transform rates using the logarithmic transformation, or proportions using the logit transformation, before adding the error term, in order to stabilise error variances across the table, and to avoid the possibility of estimating negative rates, or proportions outside the range (0, 1). Error terms may then be assumed to be independently and identically distributed normal random variables, and with these assumptions models may be fitted on the transformed scale by least-squares. This may be done using the Newton-Raphson procedure (Bock, 1975), which converges in one iteration for least-squares fitting of linear models, but which requires the provision of good starting values for rapid convergence in other situations. Least-squares fitting of the bimodel may be achieved more efficiently using the Singular Value Decomposition provided that the data matrix contains no missing values (see Chapter 2 above).

With the above error assumptions, the appropriate goodness-of-fit measure is the R^2 statistic, which expresses the proportion of variance in the transformed data accounted for by the model. For nested models, R^2 values may be used to constructe an F-test. However most of the models considered below are non-nested, and consequently comparisons of fit must be less formal. The adjusted R^2 value, R^2_{adj}, is useful for comparing models with different numbers of free parameters:

$$R^2_{adj} = 1 - \frac{N - 1}{N - n} (1 - R^2) \qquad (3.1)$$

where N is the number of non-missing rates or proportions in the table, and n is the number of degrees of freedom in the model.

Goodness-of-fit is not the only criterion by which models should be assessed. Interpretability is equally important: that is, parameter estimates should be consistent with the theoretical assertions of the model, which in turn should be consistent with general experience and reasoning about substantive processes underlying the data.

3.1.4 The sequel

The following four sections contain, respectively, the analyses of the nuptiality data, the fertility data and the legitimate fertility data for England and Wales, and the mortality data for Sweden, mentioned in section 3.1.1 above. Each section contains subsections corresponding to the four types of model discussed above, together with a summary of results. Finally, section 3.6 contains the conclusions from the analyses.

3.2 Nuptiality.

The data analysed here are male and female nuptiality rates classified by single years of age and time-periods from 1938 to 1976, omitting data prior to the 1920 birth cohort, for England and Wales (O.P.C.S., 1977a). Figure 3.1 displays age structures for selected periods for both males and females for these data, from which it is apparent that female nuptiality is more concentrated at the younger ages than male nuptiality, but for both sexes there is a clear trend towards younger marriage throughout the series, this trend being reversed briefly in the immediate post war years, presumably through marriages delayed by the war. The post-war female cross-sectional age structures are characterised by a discontinuity at age 21, which is also slightly evident in the male nuptiality data. The 1970 Family Law Reform Act, which lowered the age of majority from 21 to 18 years, has presumably been responsible for the disappearance of this phenomenon in recent years. Similar graphs of cohort age structures show marked irregularities around the war years.

3.2.1 The Coale-McNeil Model.

A model of the age structure of female cohort nuptiality is provided by Coale (1971), which is based on schedules of proportions ever-married at each age for female cohorts from a wide variety of countries, and is of the following form:

$$G(a) = \Psi G_s \left(\frac{a - \alpha}{\beta}\right) \qquad (3.2)$$

where G(a) denotes the proportion of females in the cohort at age a who are ever-married; ψ denotes the proportion at birth who will be exposed to risk of marriage (the proportion marriageable); G_s is a standard function of proportions ever-married based on Swedish

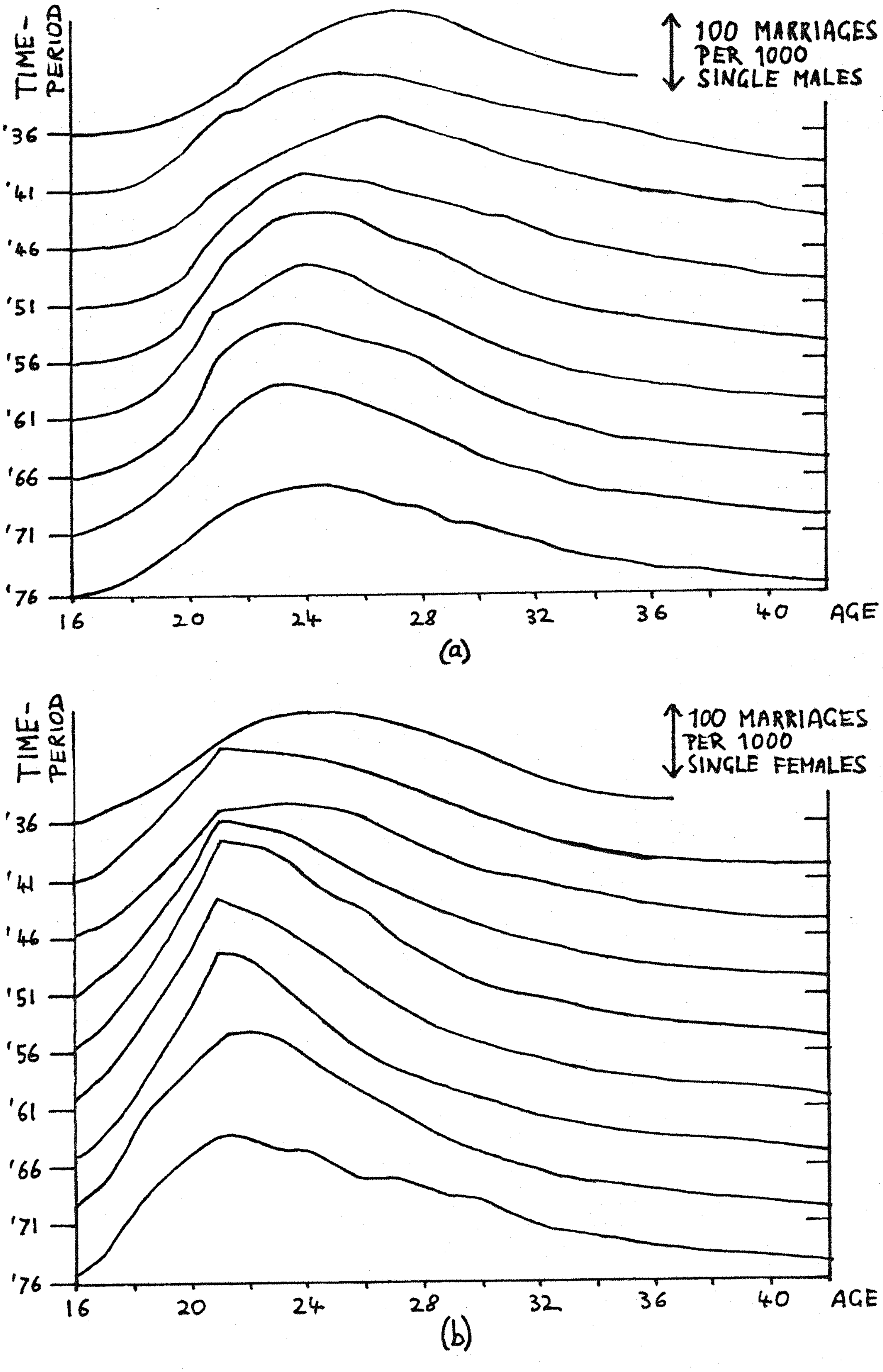


Figure 3.1 Age-structures of nuptiality rates for selected time-periods between 1936 and 1976 for England and Wales, for a) males and b) females. The structures are stacked with that for 1936 uppermost. For each time-period nuptiality rates are measured from the corresponding scale-mark on the vertical axis.

experience; and α and β specify, respectively, the initial age at, and pace of, nuptiality in the cohort. Coale and McNeil (1972) show that the shape of the empirical G_s function is consistent with a normally distributed age at entry into the marriage market followed by a sequence of exponentially distributed delays into subsequent pre-nuptial stages. Unfortunately a closed form expression for G_s has not been found, although Coale and McNeil (1972) provide a closed form expression for the standard function of first marriage frequencies, g_s, as follows:

$$g_s(x) = .1946 \exp\{-.174(x - 6.06) - \exp[-.2881(x - 6.06)]\} \qquad (3.3)$$

where

$$g_s(x) = \frac{d\, G_s(x)}{dx} \qquad (3.4)$$

Now the nuptiality rate y(a) at age a is given by:

$$y(a) = \frac{d\, G(a)}{da} \,/\, \{1 - G(a)\} \qquad (3.5)$$

$$= \frac{\psi}{\beta}\, g_s\left(\frac{a - \alpha}{\beta}\right) / \{1 - G(a)\}$$

using (3.2) and (3.4). Substituting g_s from equation (3.3) into this expression, taking logs and replacing functions of ψ, α and β with new parameters ϕ, γ and δ gives a formulation of the original model on the scale of log nuptiality rates:

$$\ln y(a) = \phi + \gamma . 0.604a - \exp(\gamma a + \delta) - \ell n\,(1 - G(a)) \qquad (3.6)$$

To apply (3.6) to cohorts in an age by period array of nuptiality rates, subscrips i, j, and k may be introduced to denote age, period and cohort categories, and an error term may be added, as follows:

$$\ell n\, y_{ijk} = \phi_K + \gamma_K . 0.604 a_i - \exp(\gamma_K . a_i + \delta_K) - \ell n(1 - G_{ijk}) + \varepsilon_{ijk} \qquad (3.7)$$

where a_i denotes the mid-point of age group i and G_{ijk} denotes the proportion of females ever-married by age a_i in cohort k.

This formulation of the marriage model of Coale and McNeil (1972) is somewhat less than ideal because it involves the empirical proportions ever-married G_{ijk} which may be calculated from the data using life-table techniques; this might be excusable if the calculation of G_{ijk} only involves the nuptiality rates prior to age i for cohort k, but a good estimate of G_{ijk} actually involves y_{ijk} itself, and thus y_{ijk} implicitly appears on both sides of equation (3.7). An alternative procedure would be to replace G(a) in equation (3.6) using (3.2) but this leads to an expression which is virtually intractible computationally. Another alternative would be to estimate model (3.2) directly, on the scale of porportions ever married, as do Coale and McNeil (1972), but the solution would then be non-optimal on the scale of log nuptiality rates, and this would exacerbate model comparisons. Moreover, it would then become computationally very difficult to generalise the model into a mixed age, period, cohort model as is done below. In these circumstances it seems preferable to stick with the relatively minor imperfections of model (3.7).

Model (3.7) is therefore applied to female nuptiality rates for England and Wales, with the goodness-of-fit reported in the first line of table 3.1 . Apart from those for the most recent (and therefore most incomplete) cohorts, the parameter estimates are reasonable, and indicate for successive female cohorts a slightly decreasing initial age-at-marriage, a substantially decreasing pace of marriage, and no discernable trend in proportions marriageable. The decreasing pace of marriage could reflect a trend towards longer courtships and pre-nuptial cohabitation. However, time-period influences due to the war and its aftermath, which are clearly evident in the data, are not accounted for by model (3.7) and could be partly responsible for the

Equation Number	Model	Sex	Degrees of Freedom Model	Data	R^2	R^2_{adj}
	Coale-McNeil					
3.7	Cohorts	F	180	1513	.911	.899
		M	180	1513	.980	.977
3.8	Periods	F	180	1513	.862	.843
		M	180	1513	.966	.961
3.9	Mixed	F	123	1513	.969	.966
		M	123	1513	.985	.984
	Additive age, period, cohort.					
3.10	Age + period + cohort	F	153	1513	.973	.970
		M	153	1513	.982	.980
3.11	Age + period	F	94	1513	.957	.954
		M	94	1513	.949	.946
3.12	Age + cohort	F	94	1513	.933	.929
		M	94	1513	.950	.947
	Cohort experience: Hernes' model.					
3.14	Cohorts	F	121	1513	.962	.959
		M	121	1513	.992	.991
3.15	Periods	F	121	1513	.968	.965
		M	121	1513	.988	.987
3.16	Periods with quadratic age	F	180	1513	.996	.995
		M	180	1513	.999	.999
	Bimodel					
3.17	Periods	F	121	840	.991	.989
		M	121	840	.997	.996
3.18	Periods	M + F	161	1680	.995	.994

Table 3.1: Goodness of fit for models fitted to nuptiality data for England and Wales. R^2 is the proportion of variance in log nuptiality rates explained by the model. R^2_{adj} adjusts for degrees of freedom in model and data, (equation (3.1)).

above cohort differentials.

To accommodate time-period influences, the marriage model of Coale and McNeil (1972), equation (3.2), is often applied cross-sectionally rather than longitudinally, although strictly this is illogical because G(a) may not be monotonically increasing with age within a cross-section. A cross-sectional version of equation (3.7):

$$\ell n\ y_{ijk} = \phi_j + \gamma_j .\ 0.604a_i - \exp(\gamma_j . a_i + \delta_j) - \ell n(1-G_{ijk}) + \varepsilon_{ijk} \qquad (3.8)$$

does not suffer from this inapplicability since the empirical proportions married are being used, but it is difficult to interpret this model which permits the initial age-at-marriage for a cohort to vary with age. Model (3.8) does not fit the female data as well as model (3.7), (table 3.1), which is surprising considering the time-period fluctuations in the data around the war-years. The poorer fit of (3.8) may be due to the ability of the incomplete cohorts in model (3.7) to accommodate recent changes in nuptiality patterns.

A more interpretable way of adapting (3.7) to accommodate time-period influences might be through the addition of a period parameter, to give the age, period, cohort model:

$$\ell n\ y_{ijk} = \theta_j + \phi_k + \gamma_k .\ 0.604a_i - \exp(\gamma_j . a_i + \delta_k) - \ell n(1-G_{ijk}) + \varepsilon_{ijk} \qquad (3.9)$$

but it is still difficult to interpret model (3.9) since ideally some allowance in the third and fourth terms of this model (which, from (3.5), are related to first-marriage <u>frequencies</u>) should be made for period fluctuations in nuptiality prior to time-period j. To economise on degrees of freedom in (3.9), cohort parameters were assigned to three-year width cohort categories. The resulting fit to the female data is reasonable (table 3.1), although the parameter estimates are now unstable and too extreme to be interpretable in terms of proportions marriageable etc. (Age, period, cohort, formulations of the model of Coale and McNeil, 1972, are also considered by Bloom, 1980).

Ewbank (1974) has demonstrated the applicability of the model of Coale and McNeil (1972) to Swedish male nuptiality rates. Fitting models (3.7) to (3.9) to the male nuptiality rates for England and Wales produces results similar to those for the female data, except that the fit tends to be better (table 3.1), and this is partly due to the greater smoothness of the male profiles.

3.2.2 Additive age, period and cohort factors effects.

The age, period, cohort model (3.9) above, derived from the marriage model of Coale and McNeil (1972), is difficult to interpret, and complex in construction. By comparison, the additive age, period, cohort model of equation (1.3) of Chapter 1:

$$\ell n \; y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_k + \varepsilon_{ijk} \qquad (3.10)$$

is attractively simple. Age factors in model (3.10) may be taken to include the length of the pre-nuptial stages, as suggested by Coale and McNeil (1972), and also length of employment which affects an individual's economic readiness for marriage. Macro-economic conditions, which also affect individuals' economic readiness for marriage, and wars, represent time-period factors. Attitudes to marriage formed during adolescence, and educational attainment (which affects economic readiness for marriage through earnings) represent cohort factors.

Model (3.10) fits the male and female data no better than model (3.9), (table 3.1), and interpretation of parameter estimates is hampered by the lack of identification of the model, discussed in Chapter 1. The age-period and age-cohort sub-models of (3.10):

$$\ell n \; y_{ijk} = \mu + \alpha_i + \beta_j + \varepsilon_{ijk} \qquad (3.11)$$

and

$$\ell n\ y_{ijk} = \mu + \alpha_i + \gamma_k + \varepsilon_{ijk} \qquad (3.12)$$

do not fit at all well.

The additive construction of (3.10) does not seem realistic for several reasons. Firstly, ecnomic readiness for marriage is likely to be an important factor only at the younger ages when individuals are most insecure financially; consequently macro-economic factors are likely to have their greatest impact at the younger ages. Secondly, cohorts with high educational attainment would compete successfully for employment and marriage at the younger ages, tending to increase early-nuptiality, but this advantage would wear off with age, as the better educated are selectively removed from the marriage market. Thirdly, attitudes towards marriage formed during adolescence could persuade couples to form prolonged pre-nuptial cohabitational unions, resulting in low initial nuptiality followed by high nuptiality at the later ages; again a constant cohort effect across ages for a cohort is not indicated. These last two points suggest a cohort-experience approach.

3.2.3 Cohort-experience: Hernes' model.

A cohort-experience type of model of nuptiality has been proposed by Hernes (1972), although the cohort-experience element in this model, pressure-to-marry, does not include the cohort-experience mechanisms outlined at the end of section 3.2.2 above.

Hernes (1972) models the proportions ever-married by age in a cohort for each sex, by conceptualising that members of the cohort are subject to two opposing, intuitively reasonable forces: the increasing pressure to marry as the proportion of the cohort already married increases, and the declining marriageability of the single members of the cohort as they grow older. It is assumed that the pressure to marry is directly proportional to the proportion of the cohort ever-married by age a, G(a), and that marriageability declines exponentially with age. The resulting model may then be written:

$$y(a) = \alpha\beta^{a}G(a) \tag{3.13}$$

where, as before, y(a) denotes the nuptiality rate at age a, and where α and β are model parameters. Model (3.13) is applicable to either sex.

G(a) in equation (3.13) can be removed using equation (3.5), but this leads to a rather complicated expression. However, taking logarithms in equation (3.13), reparameterising, adding an error term ε and introducing subscripts gives the simple linear model:

$$\ell n\, y_{ijk} = \mu_k + \theta_k a_i + \ell n\, G_{ijk} + \varepsilon_{ijk} \tag{3.14}$$

where, as in section 3.2.1, G_{ijk} can be estimated from the data using life-table techniques. As in section 3.2.1, model (3.14) possesses the disadvantage that the calculation of G_{ijk} involves the dependent variable itself, although the model now contains only two parameters per cohort.

To accommodate time-period fluctuations, model (3.14) may be re-specified for cross-sections, as follows:

$$\ell n\, y_{ijk} = \mu_j + \theta_j a_i + \ell n\, G_{ijk} + \varepsilon_{ijk} \tag{3.15}$$

in which the first two terms of the model state that the effect of age on marriageability depends on contemporaneous influences, whilst the pressure to marry remains dependent on the proportion already married in the cohort. Thus this cross-sectional version of Hernes' (1972) model is much more readily interpretable than the cross-sectional version of the Coale and McNeil (1972) marriage model (equation (3.8)). From table 3. 1 , model (3.15) fits the male and female nuptiality data about as well as the purely cohort specification of equation (3.14).

The residuals from both models (3.14) and (3.15) show distinct age patterns, suggesting that modifications to the marriageability component are required. Hastings and Robinson (1973) have also noted this lack of fit. The addition of a quadratic age term to the marriageability component of (3.15) gives the model:

$$\ell n\ y_{ijk} = \mu_j + \theta_j a_i + \upsilon_j a_i^2 + \ell n\ G_{ijk} + \varepsilon_{ijk} \qquad (3.16)$$

which, from table 3.1, fits the data extremely well. Alternative modifications to the marriageability component are less successful, and relaxing the functional form of the pressure-to-marry component yields no improvements in fit. The cross-sectional specification (3.16) fits better than the corresponding longitudinal model, as might be expected.

Figure 3.2 illustrates the estimates of marriageability from model (3.16), and exhibits several interesting features. Interpreting these features in terms of the theoretical foundation of the model, it appears that marriageability decreases with age, although the rate of decline slows up slightly at the higher ages. For the earlier time-periods, males are noticeably more marriageable than females of the same age, but male marriageability is steadily decreasing with time, especially for the older men. This could reflect the growing economic independence of women: in the past the financial security offered by older men would have considerably offset the underlying age trend in marriageability for males; consequently increased female earnings would affect male marriageability particularly at the older ages. Interestingly female marriageability has not been affected by changing economic circumstances, which suggests that the attractiveness of females to

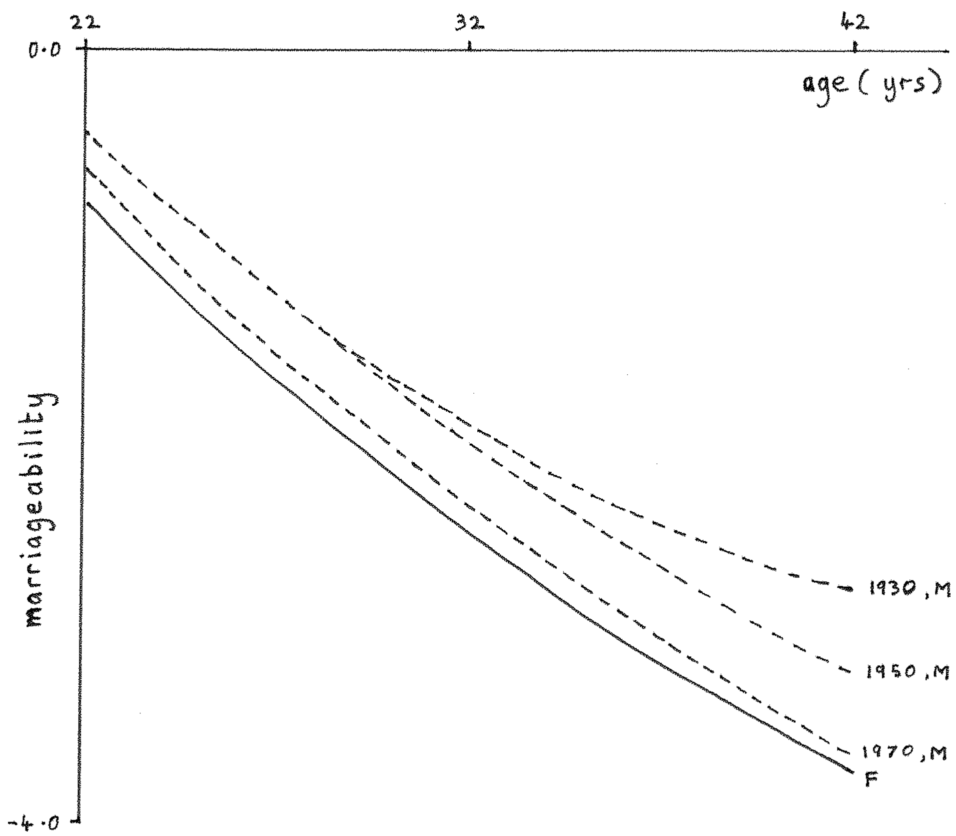


Figure 3.2: The marriageability component (the first three terms) of the modicication of Hernes' (1972) model in equation (3.16), for time periods 1930, 1950 and 1970 for males (broken line) and females (continuous line). The female curves are coincident.

males is not based on economic considerations to any extent.

The time-period parameters in (3.16) soak up the effects of the war. Interpreted in terms of the theoretical assertions of the model, both males and females became less marriageable during the war years.

It is interesting to note that no modifications of the pressure-to-marry component of the model were found necessary. This suggests that pressure-to-marry is not influenced by economic or social conditions, but only by the fear of being 'left on the shelf' which increases only with the proportion already married in the cohort. Neither of the cohort-experience mechanisms referred to at the end of section 3.2.2 above could be accommodated properly by the cohort-experience component of (3.16), since they both suggest that early and late nuptiality levels within a cohort should be inversely related. However, these mechanisms could cause the effects of the pressure-to-marry component to be underestimated, but the experiments with the pressure-to-marry component suggest that a positive relationship between early and late nuptiality within a cohort is dominant.

3.2.4 The bimodel.

Attitudes to marriage may not necessarily be formed in adolescence, as assumed in section 3.2.2 above. Changing macro-economic conditions provide a climate for evolution of attitudes towards marriage and patterns of family living, and it is reasonable to suppose that all cohorts are exposed to new ideas and norms via the mass media. Macro-economic conditions also have a direct affect on economic readiness for marriage, and would consequently affect nuptiality levels particularly at the younger ages, as noted in section 3.2.2 above. These considerations, and also the data themselves, suggest that a model which expresses a gradual change from one cross-sectional age-structure of nuptiality to another would potentially account for much of the data variation. Such a model assumes that cohort-experience and other cohort-related mechanisms are not of great importance.

As noted in section 3.1.2 above, the bimodel represents one way of expressing a transition in age-structures for one sex:

$$\ell n\ y_{ijk} - \tilde{y} = \alpha_i\ \gamma_j + \beta_i \delta_j + \varepsilon_{ijk} \qquad (3.17)$$

where $\{\alpha_i\}$ and $\{\beta_i\}$ are two age-structures and $\{\gamma_j\}$ and $\{\delta_j\}$ are time-period weights, and where $\tilde{y}$ is the mean of the log rates for one sex. The logarithmic scale is chosen here merely to ease comparisons of fit with other models in this section. The removal of the mean $\tilde{y}$ is to prevent the fit reflecting the units of the rates: this effectively adds one degree of freedom to the model. Model (3.17) possesses the same parametric structure as the 2-component biplot, given in equation (2.2) of Chapter 2, with k = 2. This model, applied separately to the male and female nuptiality data, fits very well (table 3.1).

Now from figure 3.1, the period trends, but not the age-structures, are similar between the sexes: therefore it is worth fitting model (3.17) to both sexes simultaneously, providing different age-parameters but the same period-parameters for each sex. Formally this model may be written:

$$\ell n\ y_{ijk} - \tilde{y} = \alpha_{is}\gamma_j + \beta_{is}\delta_j + \varepsilon_{ijks} \qquad (3.18)$$

where subscript s has just two levels denoting sex, and where $\tilde{y}$ is now the mean of both male and female log rates. Model (3.18) may be estimated using the Singular Value Decomposition by arranging the data in the form of a matrix whose rows are time-periods, and whose columns are combinations of age and sex. From table 3.1 it appears that bimodel (3.18) provides a comparable fit to that of the bimodel applied separately to each sex, equation (3.17), but is much more efficient with degrees of freedom.

As discussed in Chapter 2, the use of the Singular Value Decomposition to estimate the bimodel requires that the data matrix should contain no missing values. The present data matrix has missing values corresponding to the pre-1900 birth cohorts, and so the bimodels (3.17) and (3.18) were estimated only for periods after 1934 and ages below 36 years.

Figure 3.3 contains the biplot corresponding to model (3.18). The curvature in the age markers reflects the changing age structure of nuptiality from one in which marriages at the higher ages are usual to one in which marriages are concentrated between ages 20 and 25. The female age markers show even greater curvature, demonstrating that ages- at-marriage for females have become more concentrated than for males. The period markers show a reversal of these trends since 1970.

The impact of the 1939-45 war can be clearly seen in the period markers in figure 3.3, which also display a brief reversal in the trend toward younger marriage in 1946, presumably due to marriages delayed by the war. Glass (1976) has noted that since the war the increasing availability, effectiveness, and use of contraception, and the improving employment prospects for women after marriage, has removed the traditional view of marriage as the start of childbearing and that consequently women have become less reluctant to marry at a young age. The availability of effective contraception together with a recognition of the growing lack of security offered by marriage has recently led to the emergence of pre-nuptial 'trial marriages' (Wilkie, 1981), and this may be partly responsible for the post 1970 reversal in the trend towards younger marriage. Ermisch (1981) finds that the decline in nuptiality at the younger ages in the 1970s is largely due to the increase in women's earnings relative to men's.

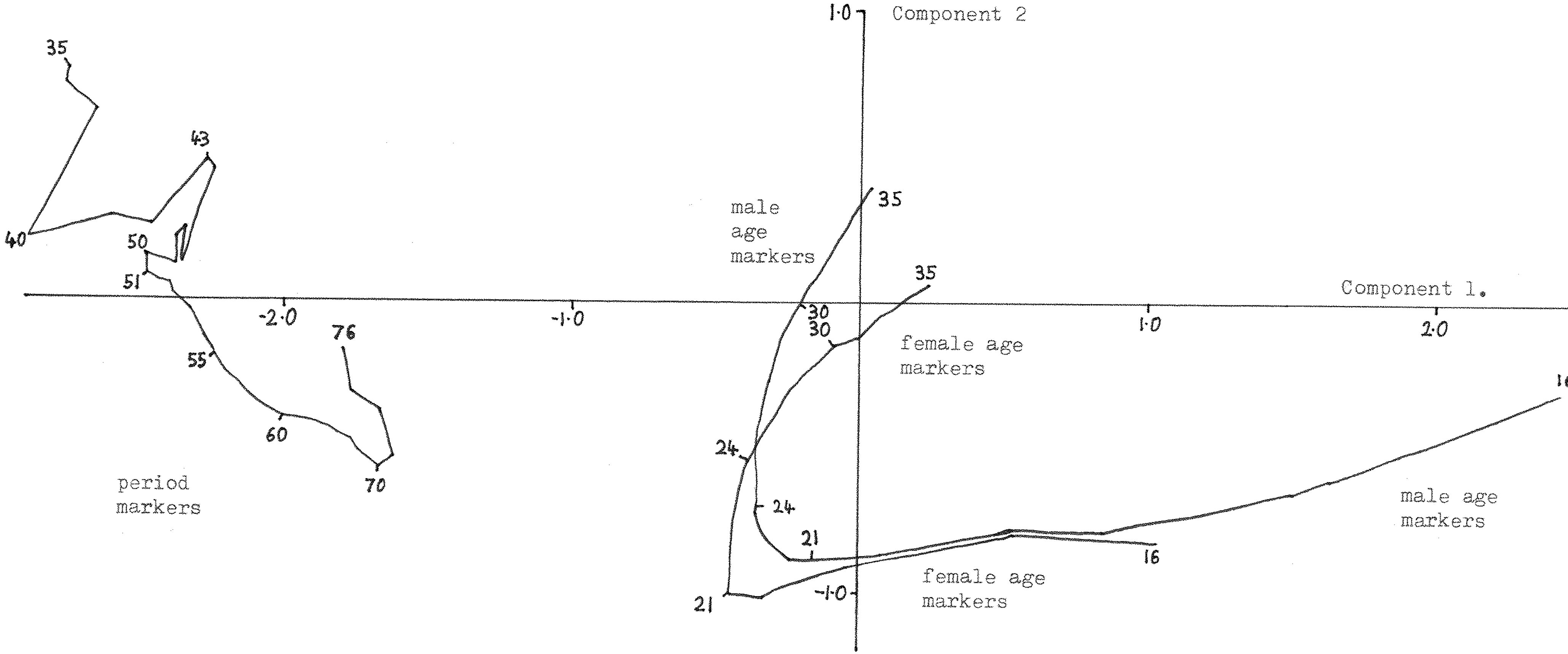


Figure 3.3: Biplot of mean corrected log nuptiality rates for males and females in England and Wales, corresponding to model (3.18)

3.2.5 Discussion of nuptiality models.

In this section a number of causal mechanisms underlying the nuptiality data for England and Wales have been discussed, and it is often the case that there is more than one way of accommodating a given mechanism into a model. None of these mechanisms, however, suggest the formulation of either cohort factor effects which are constant for all ages within each cohort or period factor effects which are constant for all ages within each time-period. The best fitting models represent either cohort-experience type factors (the adaptation of Hernes' (1972) model (3.16)) or period factors which interact with age-structure (the bimodel (3.17)). On grounds of both fit and interpretability, the adaptations of the Coale-McNeil (1972) model (3.8) and (3.9) and the additive age, period, cohort model (3.10) may be dismissed. Possibly there is some prospect of adapting the Coale-McNeil (1972) model to accommodate period factor effects, but the difficulty in doing this is evident from section 3.2.2 above.

The result then is a dilemma; a choice between two extremely well fitting, intuitively reasonable, and highly interpretable models which make somewhat different substantive assertions. In some respects the adaptation of Hernes' (1972) model (3.16) and the bimodel (3.17) are similar: they both contain only period parameters. Potentially the period parameters in (3.16) could accommodate the changing age-structure in nuptiality; however for females the period parameters have changed remarkably little over time: changes in age structure being largely accommodated by the pressure-to-marry term. Consequently it cannot be said that (3.16) is merely an approximate version of (3.17). Moreover the trends in marriageability illustrated in figure 3.2 are tantalisingly interpretable. Both models account for the impact of macro-economic conditions and contraceptive availability, but in rather different ways.

Neither (3.16) nor (3.17) adequately express the 'catching-up' of marriage plans postponed by the war, although both models indirectly accommodate the consequent post-war bulge in late marriages (figure 3.1). However, this is a minor source of data variation. Also neither model explicitly takes into account the effects of imbalances between the sexes, which Ermisch (1981) has found to be of some importance for England and Wales.

So far as projection is concerned, the bimodel (3.18) has the advantage since it contains only two period parameters for both sexes combined, whereas the adaptation of Hernes' (1972) model (3.16) would require six period parameters to be projected for both sexes. Additionally, the use of G_{ijk} in model (3.16) complicates its use for projection since its computation involves the nuptiality rate it is involved in projecting.

3.3 All-women fertility.

The data analysed here are fertility rates for all women in England and Wales, classified by single years of age and time-periods from 1938 to 1979, (O.P.C.S., 1979)[1]. Figures 3.4 and 3.5 display age structures for selected periods and cohorts for these data, from which it is apparent that childbearing has become increasingly concentrated around age 25 years, and that period age-structures are considerably more regular than cohort age-structures.

It might be argued that there is little point in modelling all-women fertility rates unless the proportions married are explicitly taken into account, in which case one may as well model legitimate fertility rates directly. An analysis of legitimate fertility is given in section 3.4 below, but in defense of the present analysis it is noted that not all fertility is legitimate; that some legitimate fertility is pre-maritally conceived; and that the timing of marriage may in many cases be influenced by fertility intentions or expectations. Consequently the case for analysing legitimate fertility alone is not clear cut. Thus this section applies the varyious types of model discussed in section 3.1.2 above to the all-women fertility data for England and Wales. Section 3.4 below contains a similar analysis of legitimate fertility.

3.3.1 The Gompertz function.

Many models of the age-structure of all-women fertility have been

---

1. These data, from an unpublished table, showed marked irregularities between five-year age groups for the earlier time-periods, undoubtedly reflecting a rather crude apportionment of five-year age-specific populations exposed to risk. These irregularities were removed using a bi-cubic spline technique (Hayes and Halliday, 1974) before commencing the analysis above.

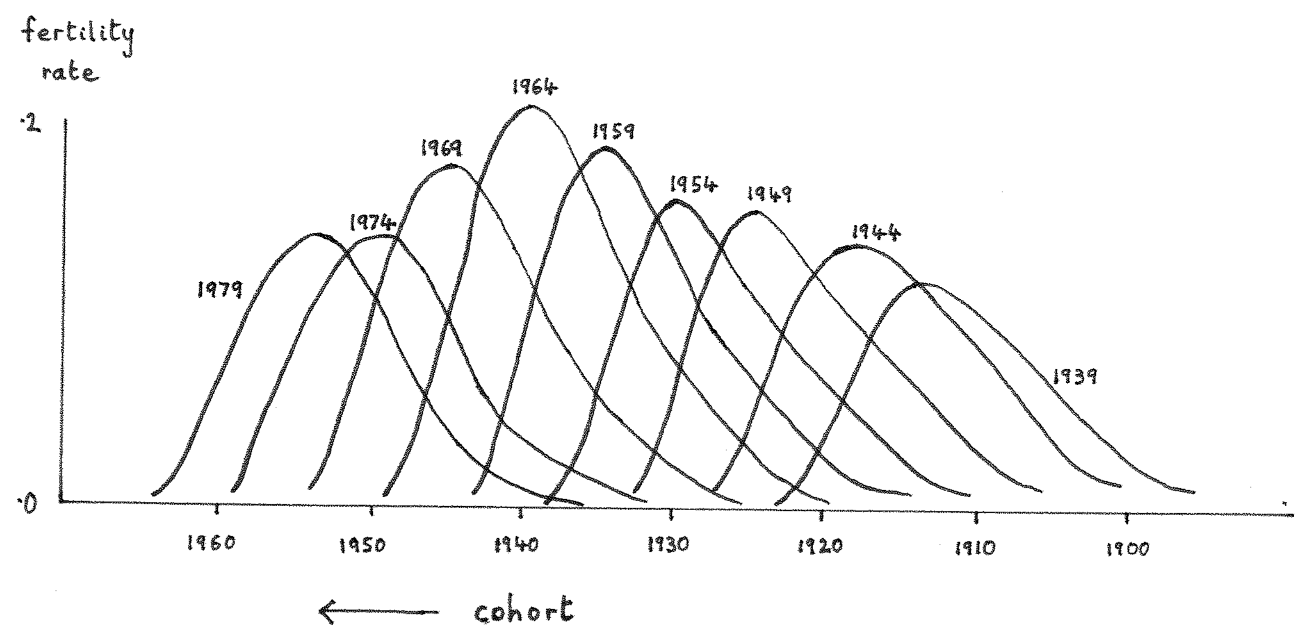


Figure 3.4: Fertility rates for all women in England and Wales for selected time-periods, by birth cohorts. Each curve represents the age-structure for one time-period with age increasing from left to right, from age 15 to age 45.

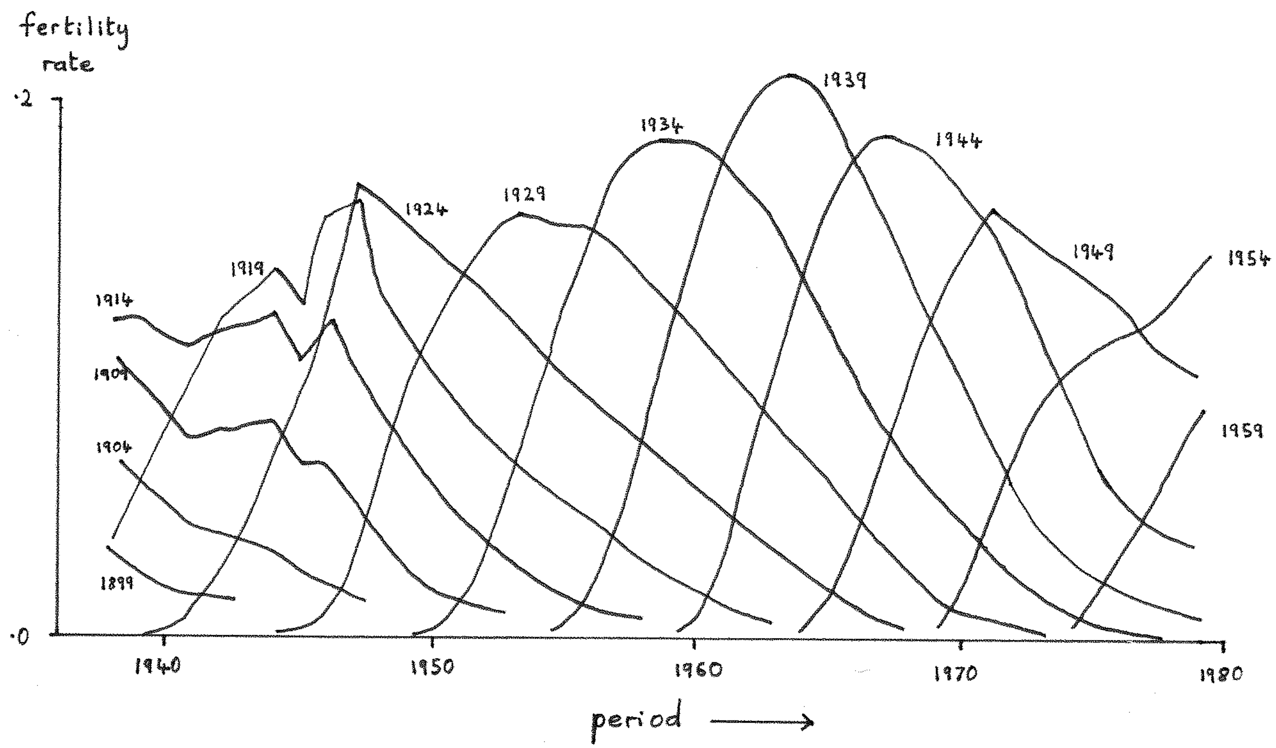


Figure 3.5: Fertility rates for all women in England and Wales, for selected birth-cohorts by time period. Each curve represents the age structure for one birth cohort with age increasing from left to right, from age 15 to age 45.

proposed, including: a polynomial (Brass, 1960); Johnson's (1949) S_B functions (Talwar, 1974); Gompertz and Makeham functions (Wunsch, 1966; Murphy and Nagnur, 1972; Murphy, 1982), Hadwiger's (1940) net maternity function, the Lognormal distribution and the gamma (Pearson III) distribution (Duchêne et al, 1974), the beta (Pearson I) distribution (Mitra, 1967; Brass et al 1968; Romaniuk, 1973; Mitra and Romaniuk, 1973; and Duchêne et al, 1974); and Mazur (1963), Coale and Trussell (1974) and Pittenger (1980) have developed their own specialised forms. The fit of these models to these data ranges from very bad (Hadwiger's 1940 function) to good (Gompertz); however these results are not reported in full here. Hoem et al (1981) have compared some of these models on Danish data, including Brass's (1974a, b) relational models. The Gompertz model, being a well-known and generally well fitting model, is considered here.

Murphy and Nagnur (1972) use the Gompertz function to model cumulative fertility within cohorts, and also within periods. The Gompertz function takes the following form:

$$F(a) = \alpha\beta^{\gamma^a} \tag{3.19}$$

where, in this case, F(a) denotes fertility cumulated (within periods or within cohorts) up to age a. Differentiating (3.19) with respect to age gives an expression for the fertility rate y(a):

$$y(a) = \alpha\beta^{\gamma^a} . \ell n\beta . \ell n\gamma . \gamma^a \tag{3.20}$$

and taking logs, reparameterising, adding an error term and introducing subscript notation gives the cross-sectional model:

$$\ell n\ y_{ijk} = \mu_j + \upsilon_j a_i + \theta_j e^{\upsilon_j a_i} + \varepsilon_{ijk} \tag{3.21}$$

Model (3.21) fits the present data well, as may be seen from the first line of table 3.2. A cohort - oriented version of (3.21) does not fit the data well, as might be expected from figure 3.5.

Murphy and Nagnur (1972) note that the three parameters, α, β and γ , in (3.19) have a straightforward demographic interpretation: α estimates total achieved fertility, β is inversely related to the average age of childbearing; and γ is related to the spread of the age-specific fertility schedule about the average age of childbearing. This 'interpretation' however says nothing about the substantive processes underlying the data, and the model is therefore descriptive rather than theoretical.

The fact that model (3.21) fits better cross-sectionally rather than longitudinally indicates that time-period rather than cohort mechanisms are dominant. The war, macro-economic conditions and contraceptive availability are probably jointly responsible for the time-period trends in the data. The temporary bulge in fertility following the war is not purely a time-period phenomenon since it is confined mainly to the cohorts which were at the prime reproductive ages during the war, (some evidence of this may be seen in figure 3.5); this is clearly a 'making up' of births missed due to the war.

Model (3.21) states that primarily time-period factors are responsible for the 'boom' in fertility between 1951 and 1966, and the subsequent 'bust'. Many researchers have suggested causes for the boom and bust. Some have attempted to find roots of causation in cohorts (notably the 'Easterlin Hypothesis', Easterlin, 1968; applied to fertility movements in England and Wales by Easterlin and Condran, 1976, and Samuelson, 1976), but Lee (1974, 1976, 1978) has developed a version of the Easterlin Hypothesis which states that age-specific labour-force participation rates in a time-period affect fertility levels via the level of income compared to anticipated income ('relative economic status'). (Congdon (1980) has used this hypothesis to forecast births in England and Wales). However, an expanding economy does not necessarily imply increasing levels of fertility: Butz and Ward (1977, 1979) have derived models from the 'New Home Economics' theory of Becker (1960, 1965) and Mincer (1963) which allow for the conflict between the husband's level of income (which is associated positively with fertility) and the wife's earnings potential (which is negatively associated with fertility). Women's earnings potential has been assisted by legislation for equal employment opportunities for men and women (Glass, 1976). Ermisch (1979) compares models derived from the Easterlin Hypothesis and the 'New Home Economics' on data for England and Wales, and finds the latter to be more successful.

Model (3.21) also captures the gradual transition towards younger fertility which continued throughout the 'boom' and up to 1971. Much of this is associated with younger marriage which in turn can be partly explained by macro-economic conditions (see section 3.2 above). Only since 1966 has the fertility of older women declined, which suggests that employment opportunities for older women did not improve until that time. Since 1971 there has been a trend towards delayed fertility which has been accompanied by a rising age-at-marriage. This probably partly reflects the increasing difficulty young couples experience in affording a home and children in the presence of economic recession, and also the increasing employment prospects especially for the more educated women. Both Wilkie (1981) for the U.S.A., and Kiernan and Diamond (1982) for England and Wales, have shown that delayed parenthood is positively associated with women's educational achievement. However the direction of causation between employment and fertility is not one-way: Jones (1982) considers the effects of childbearing on employment.

Equation Number	Model	Degrees of Freedom	R^2	R^2_{adj}
3.21	Gompertz	126	.978	.976
	Additive age, period, cohort			
3.10	Age + period + cohort	140	.986	.984
3.11	Age + period	71	.928	.924
3.12	Age + cohort	100	.928	.922
3.22	Cohort-experience: Achieved fertility	95	.961	.958
3.17	Bimodel, periods	141	.976	.973

Table 3.2: Goodness-of-fit for models fitted to the age by period array of fertility rates for all women in England and Wales. R^2 is the proportion of variance in log fertility rates explained by the model. R^2_{adj} adjusts for model degrees of freedom,(equation (3.1)).

3.3.2 Additive age, period and cohort effects.

The original version of the Easterlin Hypothesis (Easterlin, 1968) states that 'relative economic status' (economic status relative to economic aspirations formed in adolescence) has a positive effect on fertility, and that relative economic status is roughly inversely related to 'relative cohort size' (the size of a cohort relative to the size of the cohort one generation before). Thus relative economic status is constant within cohorts. Ryder (1978) states that couples have intentions concerning both the quantum and tempo of fertility, although the tempo of fertility is often distorted by period phenomena (such as wars, economic conditions, etc.) Ryder's (1978) view is not inconsistent with that of Easterlin (1968), and both suggest an age, period, cohort framework of analysis. Sanderson (1979) has used the additive age-period-cohort model (3.10) to test an aspect of the Easterlin Hypothesis on data for the U.S.A., and Pullum (1980) has also used the model to analyse fertility in the U.S.A.

The results (table 3.2) of fitting model (3.10) (with y_{ijk} denoting the fertility rate for all women) and its submodels (3.11) and (3.12) to the data for England and Wales show that model (3.10) provides a much closer fit than its submodels, the cohort parameters absorbing the changing cross-sectional age structure of fertility to some extent. However the lack of a reliable means to identify the model once again frustrates a more detailed interpretation of parameter estimates.

3.3.3 Cohort-experience: achieved fertility.

Cohorts aiming for a desired quantum of fertility (or 'target' fertility) would be expected to adjust their current fertility to take into account their achieved fertility (Lee, 1977, 1980; Ryder, 1978). The post-war 'catching-up' of births missed during the war is evidence of this phenomenon. The additive age, period, cohort model (3.10) discussed in section 3.3.2 above does not explicitly accommodate such a phenomenon, since it does not posit a dependence on achieved fertility. Lee's (1977, 1980) model of marital fertility accommodates this phenomenon by relating current fertility to the difference between target and achieved fertility. This model is discussed in section 3.3.4 below.

Now the concept of a target fertility is not the only, nor necessarily the best way of bringing achieved fertility into the model: to some extent the target fertility itself is dependent on achieved fertility, since couples are unlikely to consider existing children unwanted. It may be more realistic to assume that each existing child makes demands on a couple's resources of time and money, and as such represents a disincentive to further childbearing. Assuming that in the absence of these disincentives there is a basic age-structure of childbearing determined largely by biological factors, that time-period factors such as wars and macro-economic conditions influence fertility in all age groups proportionately, and that the disincentive which each child represents is constant over time, then a model may be constructed as follows:

$$\ell n\ y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_h + \varepsilon_{ijkh} \qquad (3.22)$$

where subscript h denotes the level of achieved fertility in cell (i, j, k). (To determine integer h for a given cell (i,j,k), achieved fertility for that cell is first calculated by cumulating fertility rates over previous age groups within cohort k. If the achieved fertility falls in the range 0.0 to 0.1 children, then h is assigned equal to 1; if achieved fertility falls in the range of 0.1 to 0.2 children, then h is set equal to 2; and so on. In general if achieved fertility falls in the range (h - 1)/10 to h/10 then h is the level of achieved fertility assigned to the cell). Thus model (3.22) assigns one parameter to each level of achieved fertility, and consequently the functional form of the dependence on achieved fertility is very flexible.

Model (3.22) does not represent a particularly good fit to the data, but it does fit very much better than the simple age-period additive submodel (3.11). The parameter estimates from model (3.22) are interesting and are given in figure 3.6, each panel of which reflects an arbitrary normalisation of the parameters. The age affects at the higher ages resemble Henry's (1961) schedule of natural fertility, and therefore probably reflect mainly biological factors. At the younger ages the age effects are much lower than natural fertility, reflecting later exposure to risk of childbirth relative to the natural fertility population, due to later marriages or unions and use of effective techniques of contraception.

The achieved fertility effects in figure 3.6 show that each child achieved, up to two children, represents roughly a halving of the fertility rate, other factors being held constant. Above two children, however, each additional child achieved represents roughly a ten-fold reduction in the fertility rate, other factors being held constant. This suggests that couples feel a strong disincentive to having more than two children.

The period affects in figure 3.6 initially correspond very closely to the total period fertility rate (T.P.F.R.), however, they increase faster than the T.P.F.R. during the fertility boom, and do not turn down until six years after the turning down in the T.P.F.R. This suggests that the rate of increase in the T.P.F.R. during the boom actually <u>underestimates</u> the extent to which those times were propitious for childbearing, the reason for the underestimation being that couples were having a high pace of fertility despite already larger than average family sizes for their age during the boom years. (The inverse relation-

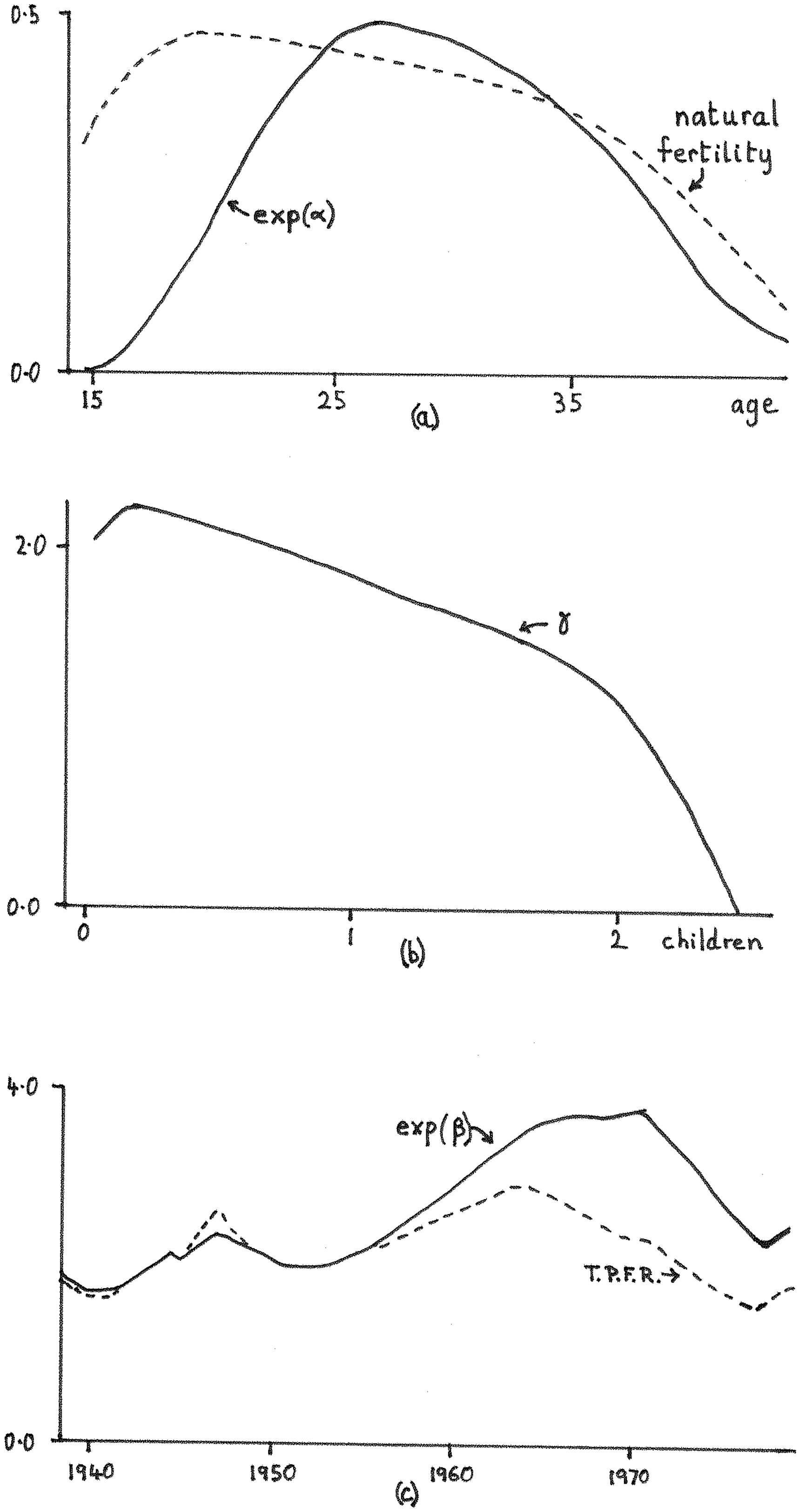


Figure 3.6: Parameter estimates from model (3.22) showing:
(a) exponential age effects and natural fertility.
(b) achived fertility effects; and
(c) exponential period effects and the T.P.F.R.

ship can be seen during the immediately post-war years, in which the T.P.F.R. overestimates the extent to which times were propitious for childbearing because much of this fertility was due to the making-up of births missed during the war.) Now between 1965 and 1971 the period effects are roughly constant, indicating that conditions for child-bearing were not changing; the falling T.P.F.R. during this period reflects the excess of children built up during the previous years. This is a most useful result, since on the assumption of constant period factor effects after 1965 one could have predicted the fall in T.P.F.R. between 1965 and 1971 (however, the age and achieved fertility parameter estimates based on the pre-1965 data are unstable; this is discussed further below). Similarly the increase in T.P.F.R. after 1977 could have been predicted even on the basis of falling period effects after 1977.

The achieved fertility effects in panel (b) of figure 3.6 curiously resemble the natural fertility age schedule in panel (a). It is tempting to try to make some substantive links between the two schedules, and it is not difficult to express the relationship algebraically. However, the relationship is extremely difficult to interpret, and it seems more likely that the similarity is merely coincidental.

The need to determine how much of the variation in levels of fertility is due to time-period factors, and how much is due to achieved or desired family size, has motivated much of the research into target fertility, and recently both Butz and Ward (1979) and Ryder (1980) have proposed essentially the same method of decomposing the T.P.F.R. into two such components. However their methods do not make explicit the underlying model, although their basic assumptions are clear enough. Model (3.22) explicitly decomposes fertility rates into period and achieved fertility components, and leads directly to a method of decomposition of the T.P.F.R. which is very similar to that of Butz and Ward (1979) and Ryder (1980), the main difference being that achieved fertility replaces target fertility in their method.

Model (3.22) makes quite strong substantive assertions: Firstly, the disincentive to further childbearing represented by existing children is assumed to be constant over all ages. This is very different from a model based on the concept of target fertility in which a high parity at one age would tend to have a much greater impact on subsequent fertility than the same parity at an older age. In the target model couples are far-sighted, planning their future fertility in order to avoid missing their target; in the achieved fertility model (3.22) couples are more short-sighted, only reacting to how many children they already have, tending to drastically reduce the risks of childbirth after the second child.

Secondly, model (3.22) asserts that the disincentive due to existing children is constant over all time-periods. This assumption seems unlikely to be true over long periods of time, since family size norms have undoubtedly reduced over the last century. However the assumption may be approximately valid for the period 1938 - 79. Attempts to relax this assumption have unfortunately led to models whose parameter estimates are uninterpretable, although room for further exploration certainly exists.

Thirdly, model (3.22) asserts that time-period factors affect fertility in all age groups and achieved fertility groups proportionately. Now panel (a) of figure 3.6 suggests that mainly biological factors are operating at the older ages, and consequently time-period factors would not be expected to have much influence in this region. At the younger ages time-period factors may have a powerful impact on fertility primarily via delayed exposure to risk of childbirth: it is interesting to note that the time-period parameter estimates in panel (c) of figure 3.6 start to decline after the boom in 1971, which coincides with the initiation of the trend towards delayed marriage discussed in section 3.2 above. Thus it seems that the proportionality in the effects of time-period factors is not realistic. Again there is room to explore methods of adjusting this aspect of the model.

If the data contained no variation across time-periods, model (3.22) would not be able to distinguish between age and achieved fertility effects. This identifies a potential source of instability in the parameter estimates: as noted above, estimating the model on the basis of the pre-1965 data produces irregular and uninterpretable parameter estimates. Consequently, it seems that a fairly long run of data, containing both upward and downward movements in fertility, is necessary in order to obtain stable and interpretable parameter estimates from model (3.22).

In summary, some of the assumptions embedded in model (3.22) are over-restrictive. Nevertheless, the model as it stands is highly interpretable owing mainly to its simplicity: this could easily be lost with more sophisticated versions. Moreover, for projection, model (3.22) requires a long run of data, but only one period parameter need be projected.

3.3.4 The bimodel.

Several time-period factors of fertility are discussed in section 3.3.1 above in connection with the cross-sectional Gompertz model (3.21). This model contains three period parameters. The bimodel (3.17) (with y_{ijk} denoting a fertility rate for all women), which contains only two period parameters, may represent a more efficient way of accommodating time-period factors.

In Chapter 2 the present data are biplotted on the untransformed scale (figure 2.1), and for two components the R^2 value is .997 (table 2.1). However, model (3.17) for the present data obtains an R^2 value of only .976. Examination of the residuals from model (3.17) suggests that its poorer fit is partly due to the dominating effect on the fit of very small fertility rates at the extreme ages, which become large and negative on the log scale. As stated in section 3.2.4 above, the only reason for wishing to use the bimodel on the log scale is to permit comparisons between the fit of different models. It is suggested that for most purposes the bimodel of the natural scale data would be more easily interpreted and better fitting. Nevertheless, the R^2 value of .976 for the log scale biplot (3.17) is only marginally worse than that for the Gompertz model (3.2.1). The biplot corresponding to (3.17)for these data is given in figure 3.7, from which is may be seen that ages 22 to 29 years are not well distinguished, and consequently not well fitted.

Like the cross-sectional Gompertz model (3.21), the bimodel (3.17) does not explicitly take account of the cohort-experience 'catch-up' effect following the war, although the time-period parameters do account for much of the consequent variation in the data.

3.3.5 Discussion of all-women fertility models.

The best fitting of the models discussed above is the additive age-period-cohort model (3.10); yet this model is also the least interpretable, not only because of the lack of identification in the parameters, but also because reasoned discussion of the factors underlying the data does not indicate constant cohort effects within cohorts. Much of this discussion suggests that time-period factors operate in such a way as to change the age-pattern of fertility: the Gompertz model (3.21) and the bimodel (3.17) both capture this phenomenon. Consequently the success of the additive age-period-cohort model seems to be due to the ability of the cohort parameters assigned to the earliest and latest cohorts (which are all incomplete) to

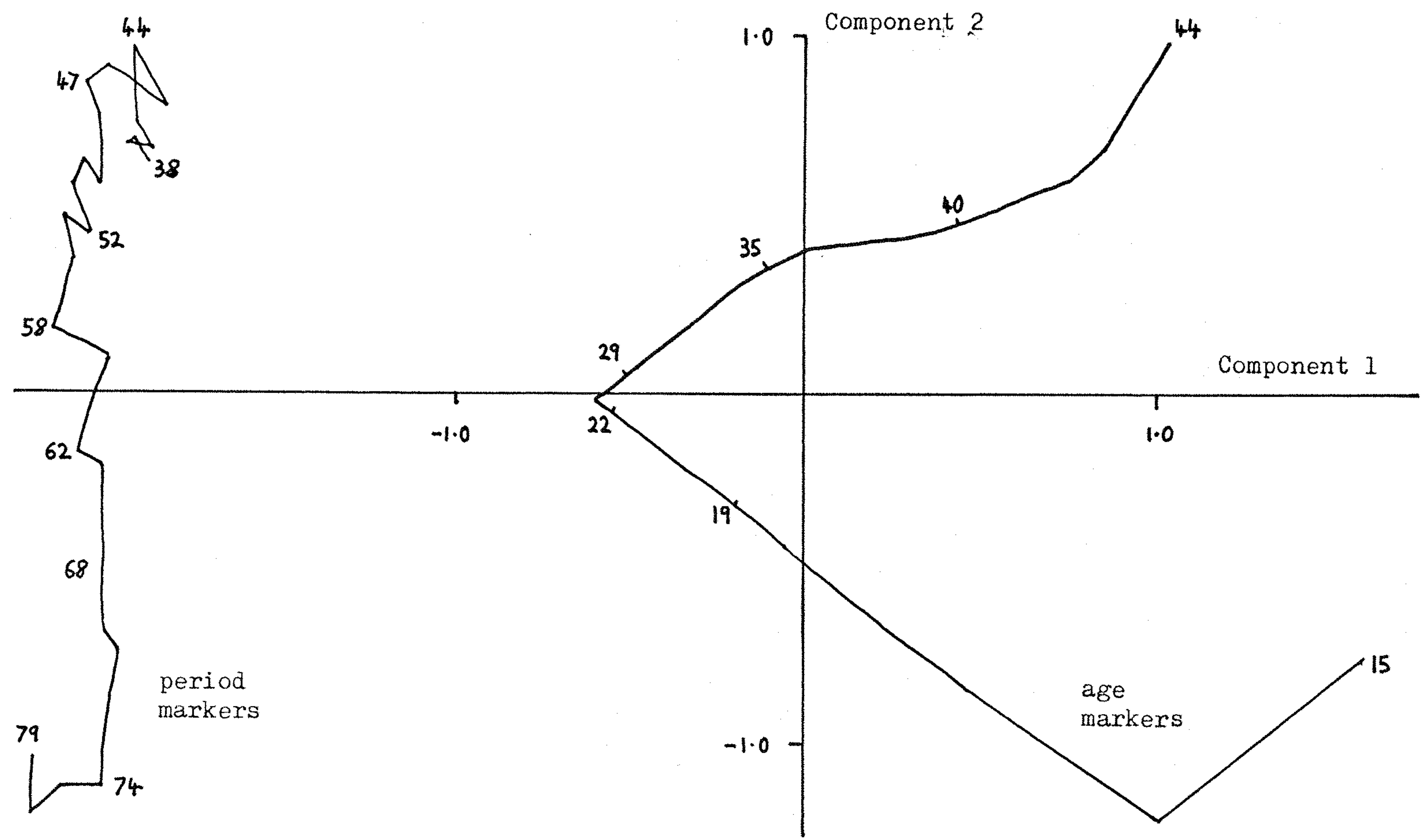


Figure 3.7: Biplot of mean corrected log fertility rates for all women in England and Wales corresponding to model (3.17)

accommodate a trend in cross-sectional age-structures.

By contrast, the most interpretable and interesting model, the cohort-experience achieved-fertility model (3.22), is also the worst fitting model (excepting the submodels of age-period-cohort model (3.10)). Some of the assumptions embedded in model (3.22) are perhaps too rigid, and further research might indicate ways in which to improve the model and its fit without detracting from its essential simplicity, which is the key to its interpretability. Usefully, model (3.22) represents an additive age-period-cohort model of sorts, which does <u>not</u> have a built-in identification problem. Despite its poor fit, the model has considerable potential for projection , since it takes into account previous levels of fertility in a direct and intuitively reasonable way, without recourse to unreliable or internal estimates of 'target' fertility.

3.4 Marital Fertility.

From figures 3.1 and 3.4 it is clear that both nuptiality and fertility for all women in England and Wales have, up to 1971, gradually become more concentrated at the younger ages, and that these trends have reversed since 1971: this may be seen more clearly in the biplots of figures 3.3 and 3.7. This suggests that an examination of marital fertility might reveal simpler patterns. Consequently this section analyses marriage duration specific fertility rates for England and Wales (O.P.C.S., 1977b). These data are classified by five year age-at-marriage categories and single-year duration and marriage-cohort categories.

Figure 3.8 illustrates these data for selected cross-sections, and reveals some very interesting features. The smooth durational profiles within age-at-marriage groups are striking; the decline in fertility with age-at-marriage is also marked. Time-period influences may also be seen, for example between 1941 and 1946, or between 1971 and 1976. The durational profiles have gradually changed throughout the series, initially being monotonically decreasing after duration zero, but later becoming humped, exhibiting a growing trend towards delayed fertility following marriage. This trend towards delayed fertility may be detected as early as 1956 in the youngest age-at-marriage group, but it is not observable in the older age-at-marriage groups until 1971. (The humping in the 1946 profiles is of a different form). Simultaneously fertility at the higher durations has become much reduced. Such a high degree of regularity is not displayed when the data are plotted longitudinally.

The models which follow were fitted for time periods 1950 to 1976, durations 1 to 14 years and five-year age-at-marriage groups from 15 to 45 years. (Duration zero is omitted because of its evident anomalous

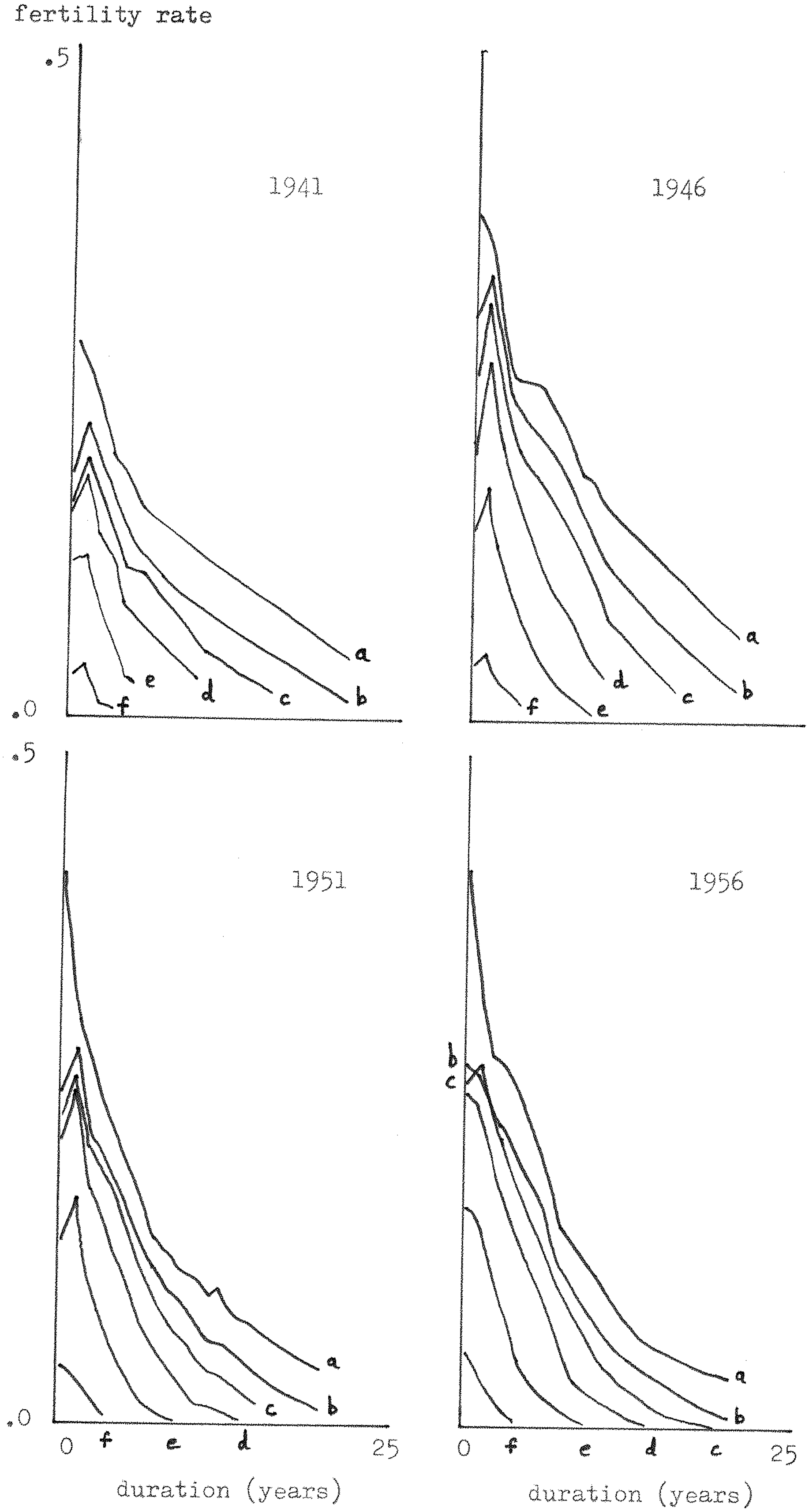


Figure 3.8: Fertility rates in England and Wales for selected time-periods, showing duration of marriage structures for age-at-marriage groups (a) under 20, (b) 20 - 24, (c) 25 - 29, (d) 30 - 34, (e) 35 - 39, and (f) 40 - 44 years.

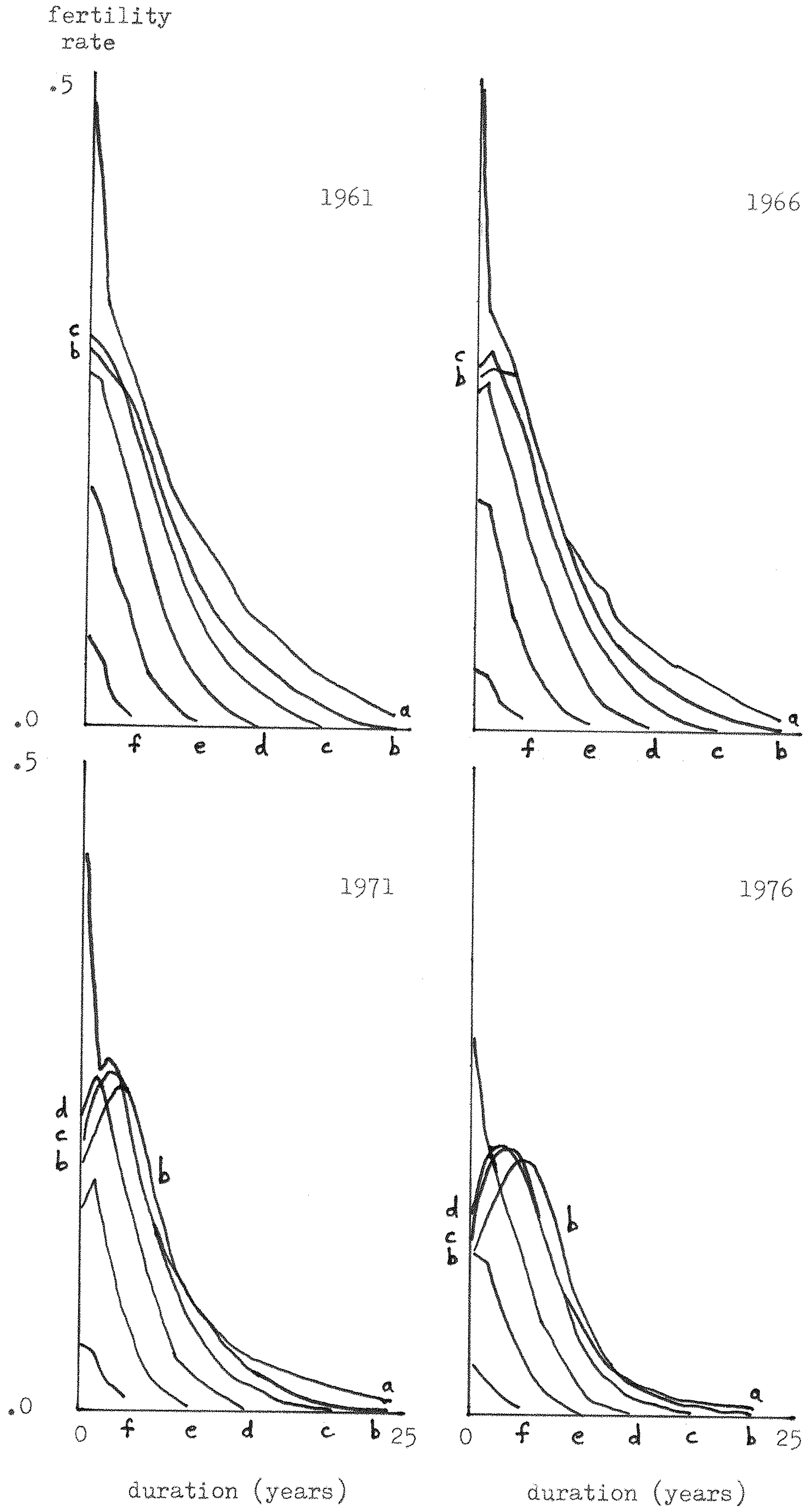


Figure 3.8: continued.

behaviour. Incomplete (up to duration 14 years) cohorts are also omitted).

3.4.1 Additive age, period and cohort effects.

Page (1976) analyses Swedish fertility rates classified by age, duration of marriage and time-period. Her model essentially consists of age and duration effects which are allowed to vary with time-period. These, time-period interactions represent a large number of degrees of freedom, although Page (1976) goes some way to modelling these interactions with fewer degrees of freedom. Gilks (1979) shows that the resultant model does not contain some main-effect terms corresponding to its interaction terms. Moreover, there is some evidence from the Swedish data that age-at-marriage effects are present. Repeating the analysis on similar data for England and Wales, Gilks (1979) finds that there is clearer evidence of age-at-marriage effects.

It is possible that interactions with time-period might be efficiently parameterised with the addition of birth-cohort or marriage-cohort factor effects to a model involving age, period, duration and age-at-marriage factor effects:

$$\ell n\, y_{ijk\ell mn} = \mu + \alpha_i + \beta_j + \gamma_k + \delta_\ell + \zeta_m + \xi_n + \epsilon_{ijk\ell mn} \qquad (3.23)$$

where $y_{ijk\ell mn}$ denotes the fertility rate in age group i, time-period j, birth-cohort k, duration-of-marriage ℓ, marriage-cohort m and age-at-marriage group n. Now there are three logical relationships between these six dimensions, which may be expressed as follows:

$$\begin{aligned} &\text{age} - \text{period} + \text{birth-cohort} = 0 \\ &\text{duration} - \text{period} + \text{marriage-cohort} = 0 \qquad (3.24) \\ &\text{age} - \text{duration} + \text{age-at-marriage} = 0 \end{aligned}$$

and each of these represents a source of confounding of the age, period, cohort type discussed in Chapter 1. Consequently model (3.23) possesses three 'linear' identification problems. Submodels of (3.23) which do not contain any of the sets of three logically related variables represented in (3.24) may still have linear identification problems, since by (3.24) : age, period, age-at-marriage and marriage-cohort are logically related; as are: age, duration, birth-cohort and marriage-cohort. A complete list of these logical relationships is given in Casterline and Hobcraft (1981).

In addition to these, further identification problems exist due to the fact that age-at-marriage categories for the present data are five times longer than the duration and period categories. Fienberg and Mason (1979) demonstrate these.

Since age-at-marriage categories are five years in length for the present data, there seems little point in calculating single-year width categories of age and birth-cohort, since the calculations involve age-at-marriage (equation (3.24)). Single-year width marriage-cohort categories however may be usefully calculated. Using five-year width categories of age and birth-cohort the identification problems associated with unequal width categories disappear, as do some of the 'linear' identification problems. However, linear identification problems which are apparently resolved by grouping categories do not solve the associated interpretational problems: this is explained in section 1.4.2 of Chapter 1.

As may be seen from table 3.3, model (3.23) does not represent a very convincing fit to data. In fact there seems little justification for retaining age-at-marriage, marriage-cohort or birth-cohort effects in the model since they account for little variation. This leaves the submodel:

$$\ell n\ y_{ij\ell} = \mu + \alpha_i + \beta_j + \delta_\ell + \varepsilon_{ij\ell} \qquad (3.25)$$

Equation Number	Model	Degrees of Freedom	R^2	R^2_{adj}
	Additive Effects			
3.23	All six effects	102	.964	.962
3.25	Age, duration, period	45	.954	.953
3.26	Age-at-marriage, duration, period	47	.859	.855
	Cohort experience: Lee's Model.			
3.32	period targets	28	.020	.005
3.33	period, age-at-marriage targets	33	.609	.602
3.34	Bimodel (periods)	189	.994	.993

Table 3.3: Goodness-of-fit for models fitted to the duration-of-marriage by age-at-marriage by period table of fertility rates for England and Wales. R^2 is the proportion of variance explained by the model R^2_{adj} adjust for degrees of freedom in model and data, (equation (3.1)).

which has the merit of not possessing any identification problems. It is perhaps surprising that this model performs better than the model containing the original dimensions of the table:

$$\ell n\, y_{j\ell n} = \mu + \beta_j + \delta_\ell + \xi_n + \varepsilon_{j\ell n} \tag{3.26}$$

since five-year age groups actually encompass a total range of ten years of age as may be seen from figure 3.9. This emphasises the importance of the age variable. Nevertheless, model (3.25) cannot be considered a good model: its structure asserts that age-at-marriage by duration profiles should be proportional between time-periods: an assertion which clearly does not correspond to the reality (figure 3.8). Thus model (3.25) fails to capture the most interesting developments in fertility observable from figure 3.8. Casterline and Hobcraft (1981) have fitted additive effect models of the above type to World Fertility Survey data from a variety of countries, and, as above, find that the best of these models is the age, duration, period model (3.25).

3.4.2 Cohort-experience: Lee's model.

From survey data for the U.S.A., Lee (1977) finds that each married women, at any point in her reproductive history, may be considered to be either a 'terminator' (i.e. she wants no more children) or a 'non-terminator' (i.e. she does want more children), and that her status (terminator or non-terminator) may change back and forth with time, depending on further births and on economic circumstances, etc., which might cause her to alter her desired family size. Lee (1977) further finds that, for non-terminators, both the fertility rate and the average additional number of children desired remain approximately constant over age and time-period. Consequently, assuming no contraception failure amongst women who are terminators:

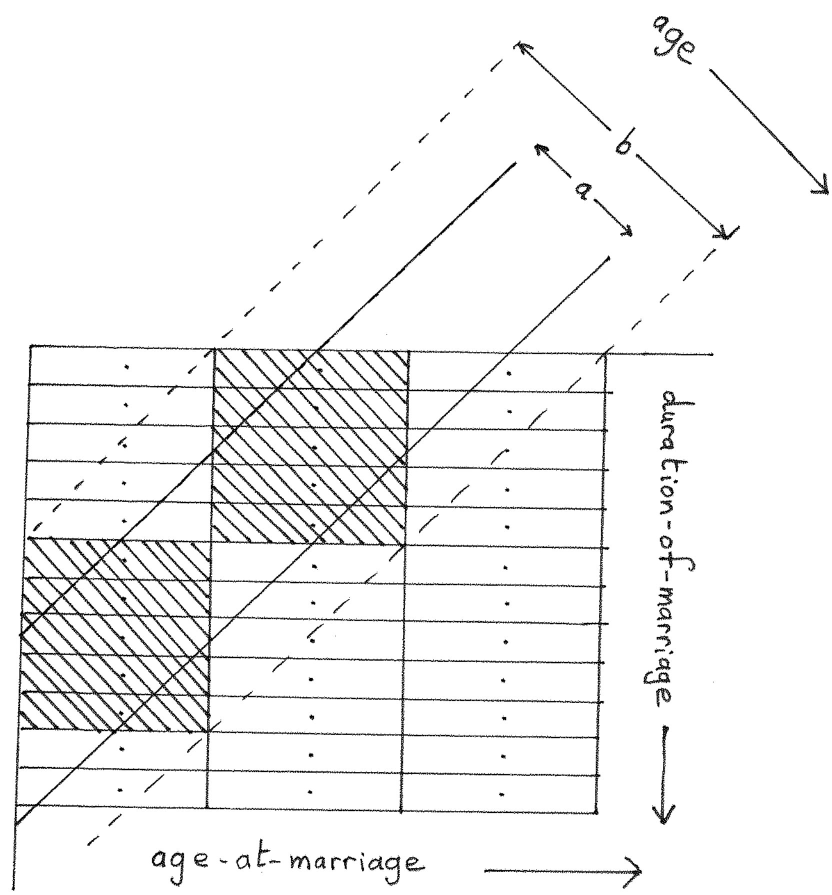


Figure 3.9: Showing five-year age-at-marriage groups by single-year duration-of-marriage groups, and resultant five-year age groups. Cells whose midpoints fall within a five-year age range, denoted 'a' on the figure, are allocated to one age group, and are shown shaded. These shaded cells encompass a total age range of ten years ('b').

$$y(a,t) = p(a,t)y^* \quad (3.27)$$

and

$$A(a,t) = p(a,t)A^* \quad (3.28)$$

where $y(a,t)$ and $A(a,t)$ are the fertility rate and average additional number of children desired for all married women aged a at time t, where y^* and A^* are the corresponding quantities for non-terminators (which do not depend on a or t), and where $p(a,t)$ is the proportion of women at age a and time t who are non-terminators. Equation (3.28) may be substituted into equation (3.27) to eliminate $p(a,t)$ to give:

$$y(a,t) = \lambda . A(a,t) \quad (3.29)$$

where $\lambda = y^*/A^*$. Empirically Lee (1977) finds that $\lambda \simeq 0.2$, but adjusting for contraceptive failure amongst terminators, suggests setting $\lambda = 0.18$.

The average desired family size for a group of married women may be termed its 'target fertility'. Ryder (1978) has proposed a model of fixed targets within each cohort. Lee (1980) considers a 'moving target' model in which targets are determined by time-period rather than cohort factors, and assumes, in order to use equation (3.29), that target fertility for married women is equal to the sum of their average achieved fertility and average additional desired fertility. This assumption implies that no achieved fertility is unwanted. Conceivably individuals may retrospectively consider all their children to be 'wanted': such retrospective rationalisation represents an additional factor. Pure time-period factors may be considered as producing a target fertility which relates to desires in the absence of retrospective rationalisation. Such a target ferility should be permitted to fall below the levels of achieved fertility, and would not in general be equal to the sum of achieved fertility and additional desired fertility.

By equation (3.27), fertility should depend on target fertility via the proportion of married women who are non-terminators. It may be realistic to assume that:

$$\ell n\ p(a,t) = \mu + \beta(T(t) - F(a,t)) \tag{3.30}$$

where T(t) is the target produced by time-period factors at time t, and F(a,t) is the average achieved fertility for married women aged a at time t. (Since p(a,t) is a proportion, it might be more appropriate to use the logit rather than the log transformation on the left of (3.30), but the log transformation simplifies subsequent development considerably). Substituting (3.30) into (3.27) gives:

$$y(a,t) = y^* \exp\{\ \mu + \beta(T(t) - F(a,t))\} \tag{3.31}$$

Now the present data are also duration-specific. Since (3.31) is proposed for all ages it may be reasonable to apply it to all durations also. Taking logs in (3.31), reparameterising, adding an error term and introducing the subscripts defined in section 3.4.1 above, gives:

$$\ell n\ y_{i\ell n} = \Gamma_j - \beta F_{j\ell n} + \varepsilon_{j\ell n} \tag{3.32}$$

where $\Gamma_j = \ell n\ y^* + \mu + \beta T(t_j)$, and $F_{j\ell n}$ denotes achieved fertility, calculated by cumulating fertility over previous duration categories within the marriage-cohort, age-at-marriage group . More realistically, target fertility should vary not only with time-period, but also with age-at-marriage, giving the model:

$$\ell n\ y_{j\ell n} = \Gamma_j + \theta_n - \beta F_{j\ell n} + \varepsilon_{j\ell n} \tag{3.33}$$

Table 3.3 shows that the period target model (3.32) explains almost no more variation than the grand mean model; the addition of the age-at-marriage term in (3.33) is essential but the fit is still very poor. In fact to obtain an adequate fit the six main-effect terms in model (3.23) must all be added to model (3.32). The poor performance of model (3.33) is partly attributable to the recent trend in delayed fertility (figure 3.8) which causes fertility rates to rise with duration within age-at-marriage groups: model (3.33) basically posits decreasing durational profiles.

Despite its poor fit, the parameter estimates for model (3.33) are reasonably interpretable: model (3.33) does not permit individual targets to be estimated but differences between targets are estimable and suggest that target fertility decreases by about 1 child per ten year increase in age-at-marriage, and that between 1966 and 1976 target fertility has dropped by about half a child on average.

3.4.3 The bimodel.

It is apparent from section 3.4.1 above that main effects in the six dimensions underlying the table are not sufficient to capture the intricate cross-sectional patterns observable from figure 3.8. Figure 3.8 essentially demonstrates a gradual transition from one age-at-marriage by duration-of-marriage pattern to another. This may be represented on the log scale by a bimodel:

$$\ell n\, y_{j\ell n} - \tilde{y} = \alpha_{\ell n}\gamma_j + \beta_{\ell n}\delta_j + \varepsilon_{j\ell n} \qquad (3.34)$$

where $\tilde{y}$ denotes the mean of the log fertility rates and where the remaining notation is as defined in section 3.4.1 above. This model fits the data well, as may be seen from table 3.3. The biplot corresponding to model (3.34) is given in figure 3.10. The curvature in the age-at-marriage profiles at the early durations in figure 3.10 represents the introduction of delayed parenthood following marriage; it may be seen that the 20 - 24 years age-at-marriage group has the greatest curvature, suggesting that delayed parenthood is most pronounced for this group. The period markers show that the trend

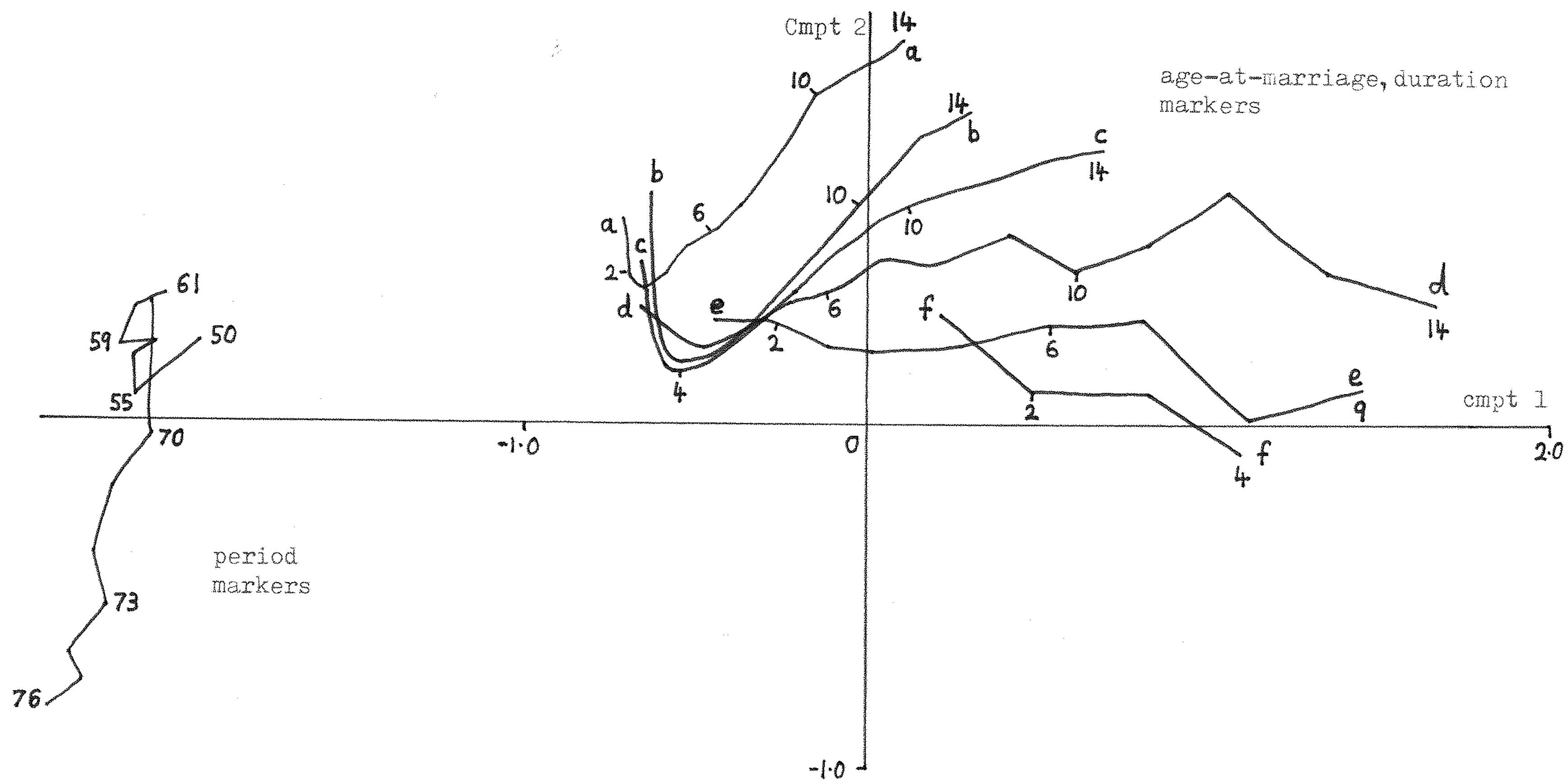


Figure 3.10: Biplot of mean-corrected log legitimate fertility rates in England and Wales, corresponding to model (3.34) showing duration of marriage structures for age-at-marriage groups (a) under 20, (b) 20 - 24, (c) 25 - 29, (d) 30 - 34, (e) 35 - 39, and (f) 40 - 44 years.

towards delayed parenthood within marriage began in earnest in 1961. The youngest age-at-marriage group is separated to some extent from the others in figure 3.10, revealing that for the earlier time-periods fertility was higher for this group, and that the trend has been to reduce this differential.

As mentioned in section 3.2.4 above, Glass (1976) attributes the trend towards delayed fertility within marriage to the increased availability of reliable means of contraception and the increasing opportunities for women to work after marriage, which have combined to remove the traditional view of marriage as the initiation of childbearing.

3.4.4 Discussion of legitimate fertility models.

It is difficult to argue convincingly for the existence of factors which operate uniquely via birth-cohort or marriage-cohort dimensions, and which futhermore have an equal effect over all levels of other factors underlying the data. Consequently it is not surprising that little evidence for such factors is found in section 3.4.1 above. Additive effect models involving the period, duration, age and age-at-marriage dimensions are incapable of expressing the recent transition in cross-sectional patterns of marital fertility. Thus none of the additive models in section 3.4.1 above represent a satisfactory theoretical or descriptive account of the data. Section 3.4.1 does, however, demonstrate that variation is much more effectively summarised by the age dimension rather than the age-at-marriage dimension, and by the period dimension rather than by cohort dimensions.

The adaptation of Lee's (1977, 1980) model (3.33), although theoretically very interesting and perhaps plausible, fits the data badly. Nevertheless, the parameter estimates do seem interpretable, suggesting declines in desired family size with age-at-marriage and time-period. The model is however incapable of accommodating the trend towards delayed parenthood within marriage.

The bimodel (3.34) fits the data extremely well. This model does not attempt to express algebraically the complex patterns within cross-sections, and only concentrates on how these patterns change between cross-sections: this is undoubtedly the reason for its success. However the price is a relatively large number of degrees of freedom. Moreover, although the cross-sectional patterns may be interpreted in substantive terms, these substantive mechanisms are not explicit in the model, and consequently the model is merely descriptive to an extent. (The model is theoretical to an extent also, asserting that time-period influences are primarily responsible for the transition in cross-sectional patterns.) It is interesting to compare the bimodel (3.34) with Page's (1976) model: whilst the former concentrates on changes between cross-section, the latter concentrates on patterns within cross-sections. Consequently bimodel (3.34) should be more more useful for projections, since it has fewer (only two) period parameters.

The trend towards delayed parenthood within marriage, which has affected all age-at-marriage groups since 1971, is coincident with a trend towards delayed marriage (section 3.2 above). It seems likely that both trends reflect the combined effect of contraceptive availability and macro-economic conditions described by Glass (1976). It is possible that fertility rates by age-at-motherhood and duration-of-motherhood would exhibit less complex patterns; many recent researches have indicated the usefulness of motherhood related variables (for example: Bumpass et al , 1978; Finnas and Hoem, 1980; Marini, 1981; Casterline and Hobcraft, 1981). Although such data for England and Wales are not published by O.P.C.S., the O.P.C.S. Longitudinal Study might prove a suitable source for such data.

3.5 Mortality.

The data analysed here are age-specific mortality rates in five-year age groups above age 30 years for five-year time-periods between 1801 and 1930 for Swedish males and females, obtained from Cramer and Wold (1935). Cramer and Wold (1935) found these data to show greater regularity within periods than within cohorts. Curiously, Kermack et al (1934) found greater regularity within cohorts when using a very similar data set for Sweden, and so it seems worth reanalysing these data here.

Figure 3.11 contains cross-sectional age-structures for these data, and figure 3.12 longitudinal age-structures. It is immediately apparent that there is a high degree of regularity cross-sectionally which does not exist longitudinally, supporting the findings of Cramer and Wold (1935). Because of its anomalous age-structure, the first time-period is omitted in the following analysis.

3.5.1 The Gompertz function.

Cramer and Wold (1935) fitted Makeham functions to the cross-sectional and longitudinal age-structures in the Swedish mortality data. The Makeham function for the proportion surviving, S(a), to age a in a cohort is:

$$S(a) = \alpha\beta^{\gamma^a}\theta^a \qquad (3.35)$$

Setting $\theta = 1$ gives the Gompertz function defined in equation (3.19). Both the Makeham and Gompertz functions are commonly used for graduation of mortality data (Miller, 1949; Wolfenden, 1954).

Now the mortality rate y(a) at age a is related to S(a) by:

$$y(a) = \frac{-S'(a)}{S(a)}$$

$$= \ell n\ \beta .\ \ell n\gamma .\ \gamma^a \qquad (3.36)$$

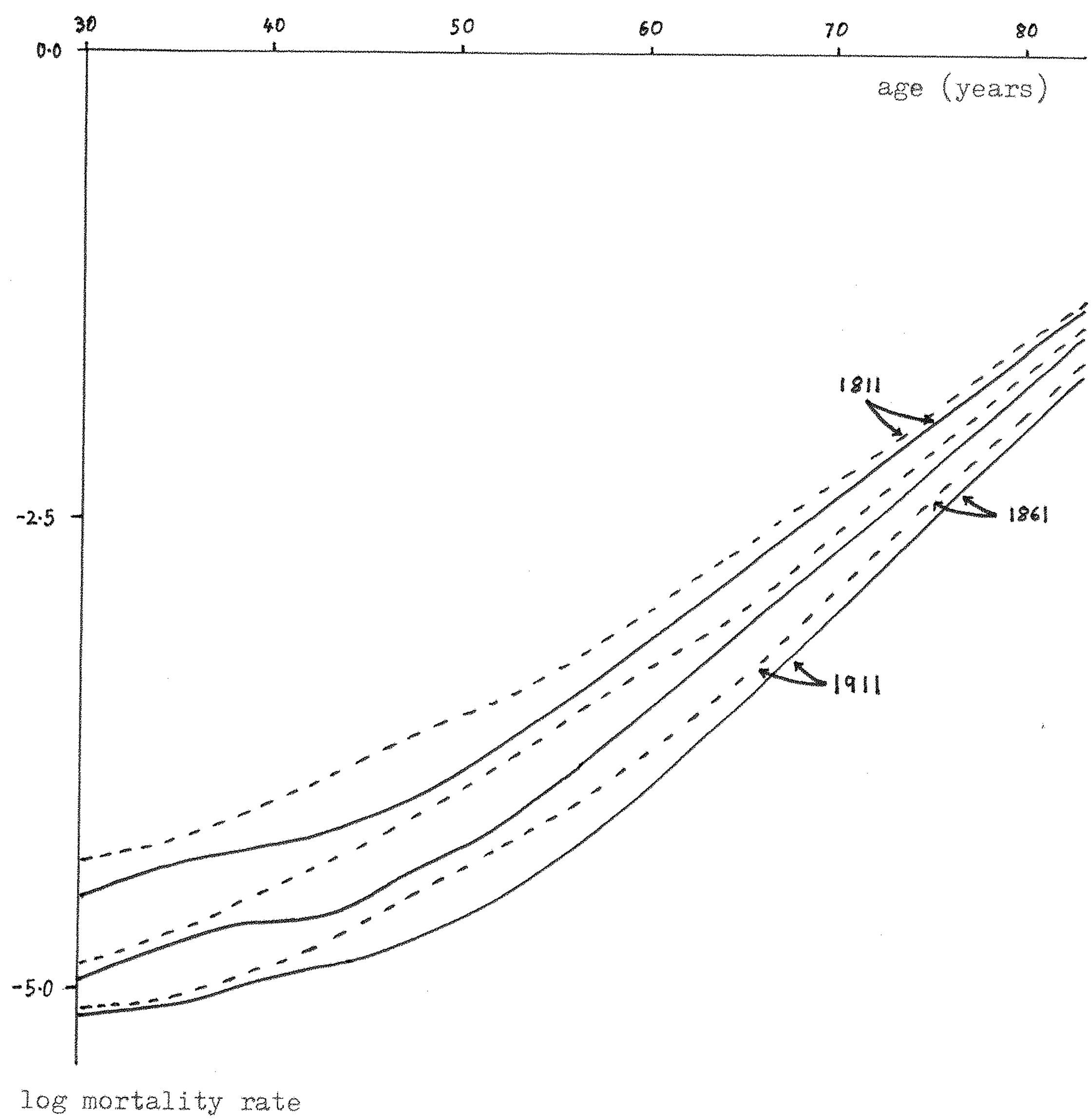


Figure 3.11: Age-specific log mortality rates for selected time-periods for Swedish males (broken line) and females (continuous line).

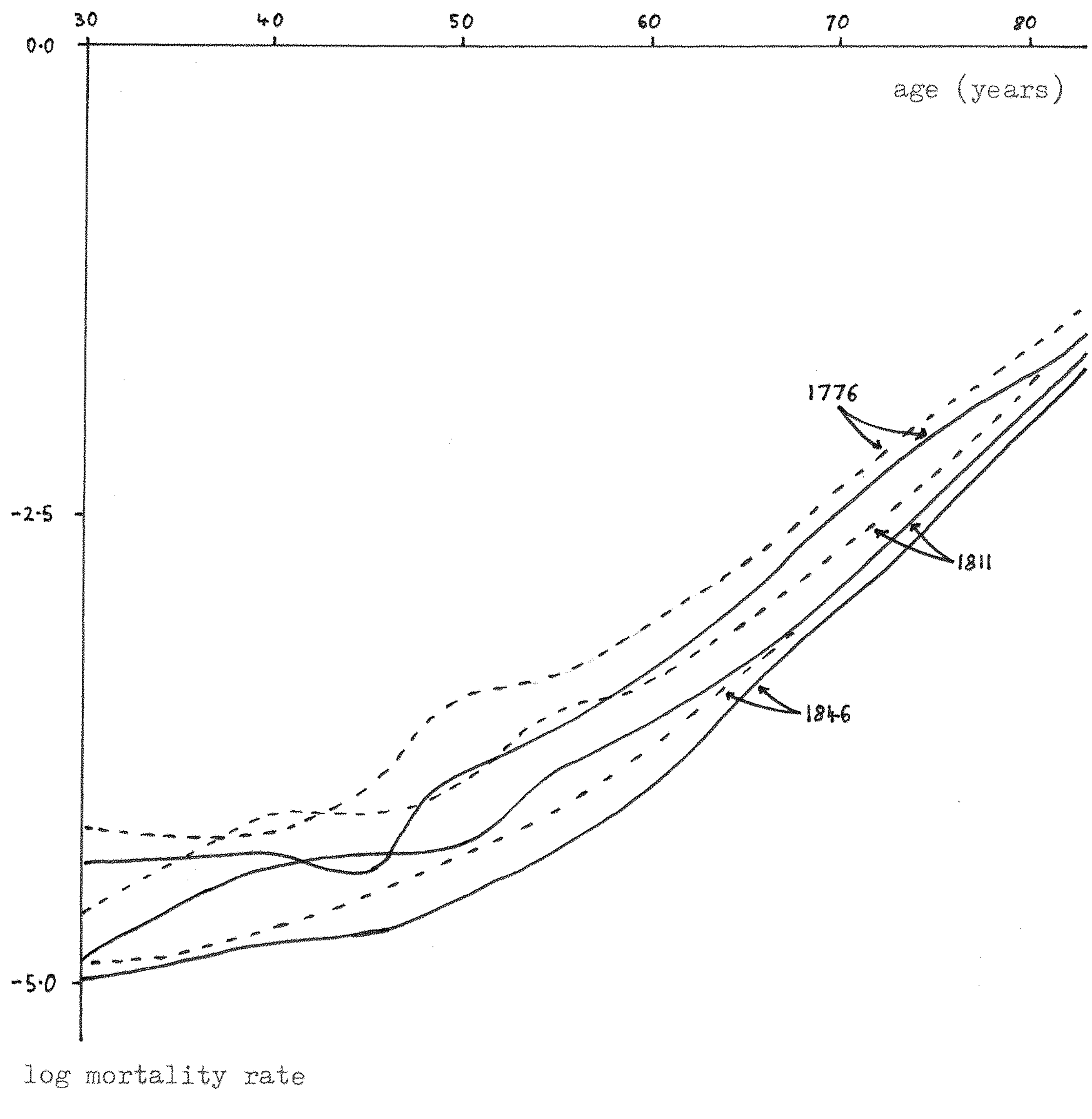


Figure 3.12: Age-specific log mortality rates for selected <u>cohorts</u> for Swedish males (broken line) and females (continuous line).

for the Gompertz function, using (3.10). The corresponding expression for the Makeham function is much more complicated, and consequently will not be used here. Taking logs, reparameterising an error term and introducing the usual subscripts gives the cross-sectional model:

$$\ln y_{ijk} = \mu_j + \alpha_j . a_i + \varepsilon_{ijk} \qquad (3.37)$$

A cohort oriented version of (3.37) is not considered here because of the evident lack of longitudinal regularity (figure 3.12).

Model (3.37) produces a rather poor fit to the Swedish data, for both males and females, as may be seen from table 3.4. Generalising the Gompertz to include a quadratic dependence on age:

$$\ln y_{ijk} = \mu_j + \alpha_j . a_i + \phi_j . a_i^2 + \varepsilon_{ijk} \qquad (3.38)$$

improves the fit considerably. Model (3.38) asserts that time-period rather than cohort-specific factors are in operation: these time-period factors could include wars, famines, epidemics, sanitary conditions, etc. The following section suggests some possible cohort factors which might be operating in addition to period factors.

3.5.2 Additive age, period and cohort effects.

Many researchers have analysed aggregate level mortality data using age, period, cohort model (3.10) (with y_{ijk} denoting a general or cause specific mortality rate); for example : Greenberg et al (1950); Sacher (1957, 1960, 1977); Beard (1963); Barrett (1973, 1978a, b); Osmond and Gardner (1982). In each case the intention is to gain insight into disease processes. For the present data cohort factors may

Equation Number	Model	Sex	Degrees of Freedom Model	Data	R^2	R^2_{adj}
	<u>Gompertz</u>					
3.37	linear age	F	50	300	.961	.953
		M	50	300	.974	.969
3.38	Quadratic age	F	75	300	.997	.996
		M	75	300	.999	.999
	<u>Additive age, period, cohort.</u>					
3.10	Age, period cohort	F	70	300	.998	.997
		M	70	300	.998	.997
3.11	Age,period	F	36	300	.997	.997
		M	36	300	.994	.993
3.12	Age, cohort	F	47	300	.995	.994
		M	47	300	.994	.993
	<u>Cohort-experience: Selection and debilitation.</u>					
3.39	Proportion surviving	F	37	234	.998	.998
		M	37	234	.997	.996
3.40	Log proportion dead	F	37	234	.999	.999
		M	37	234	.999	.999
	<u>Bimodel</u>					
3.17	periods	F	71	300	.999	.999
		M	71	300	.999	.999
3.18	periods	F + M	95	600	.999	.999

Table 3.4: Goodness of fit for models fitted to the Swedish mortality data. R^2 is the proportion of variance in log mortality rates explained by the model. R^2_{adj} adjusts for the degrees of freedom in model and data (equation (3.1)).

include lifetime experiences up to the age of 30 years, since this is where the present data commence. Nutritional standards and exposure to disease may have a lasting effect on an individual, particularly those experienced during childhood. This appears to be particularly true for mortality from tuberculosis (Sacher, 1957, 1960, 1977; Mason and Smith, 1979; see also Chapter 1 above). Period factors may include famines and epidemics as well as medical innovation, and the biological process of ageing clearly represents an age factor.

Fitting model (3.10) to the present data for males and females separately produces a good fit (table 3.4). The age, period submodel (3.11), however, also produces a high R^2, but figure 3.11 clearly does not suggest an additive age, period model on the log scale since for each sex the age structures are not parallel. This demonstrates that even very high R^2 values do not necessarily represent a very good fit when the data are very regular.

3.5.3 Cohort experience: selection and debilitation.

Model (3.9) assumes that only experiences below age 30 years have a lasting effect on individual healthiness. It may be safer to assume that experiences after age 30 years also have lasting effects. The direction of this effect is uncertain since some experiences could leave the individual debilitated whilst others could improve resistance to disease. Selectivity represents another type of cohort-experience factor: those individuals who are more frail are more likely to die young leaving a relatively less frail population surviving to the later ages. Vaupel et al (1979) have examined the impact of selectivity mechanisms on mortality.

To build these cohort-experience factors into a model the proportion surviving from age 30 to age a, $S^*(a)$, may be used as a measure of the amount of selectivity that has taken place and simultaneously as a measure of the exposure to debilitating conditions. Replacing the fixed cohort effect in (3.10) with a dependence on $S^*(a)$ gives:

$$\ell n\ y_{ijk} = \mu + \alpha_i + \beta_j + \theta S^*_{ijk} + \varepsilon_{ijk} \tag{3.39}$$

where S^*_{ijk} is the proportion surviving to age a_i in cohort k. The earliest, incomplete, cohorts were omitted when fitting model (3.39) since S^*_{ijk} cannot be calculated for these cohorts using the present data. It may be seen from table 3.4 that model (3.39) fits about as well as the additive age, period, cohort model (3.10), but uses only about half the degrees of freedom. The estimates of θ are hard to interpret, being positive for males and negative for females.

Further experimentation with models of the form of (3.39) reveals that using log proportion dead in place of proportion surviving:

$$\ell n\ y_{ijk} = \mu + \alpha_i + \beta_j + \theta \ell n\ (1 - S^*_{ijk}) + \varepsilon_{ijk} \tag{3.40}$$

produces an even better fit. The θ estimates are more easily interpreted with this model being .64 for males and .78 for females, suggesting that high previous mortality predisposes towards high subsequent mortality in a cohort, through the debilitating effect of poor nutrition and disease.

3.5.4 The bimodel

As for the other data analysed in this chapter, these data exhibit a gradual transformation from one cross-sectional pattern to another (figure 3.11). The bimodel (3.17), with y_{ijk} denoting a mortality rate, may be applied to these data, for each sex separately. Although for each time-period the mortality age structures for males and females differ, the general trend towards reduced mortality especially at the younger ages is common to both sexes, suggesting bimodel (3.18) for both sexes simultaneously. All these bimodels fit the data well (table 3.4), the both-sex bimodel (3.18) being more efficient with degress of freedom.

The biplot corresponding to bimodel (3.18) is given in figure 3.13. The male and female age markers are very close above age 55 years, reflecting a slight excess in male mortality which diminishes with time. Below age 55 the differentials between the sexes are greater, but these also diminish with time. The curvature in the age markers reflects the greater fall in mortality for the younger age groups. The period markers show that these trends had begun by 1806 but halted in 1821 and did not continue their path until the 1850s. The direction of the period markers indicates that for all age-sex groups, mortality was declining up to 1916, when the first World War and its aftermath appear to have reversed this trend.

Bimodel (3.18) asserts that primarily period-specific causes were responsible for the above trends. Declining mortality at the younger ages suggests that improving standards of hygiene and nutrition were largely responsible.

3.5.5 Summary of models of mortality.

The Gompertz model (3.38) and the bimodel (3.17) make the same substantive assertions: that time-period influences are responsible for the changing level and structure of mortality in the Swedish data; and there is little to choose between the two in terms of fit. The bimodel has the advantage that for both sexes combined, model (3.18) has only two parameters for projection whereas the Gompertz model (3.38) has six.

The additive age-period-cohort model (3.10) and the cohort-

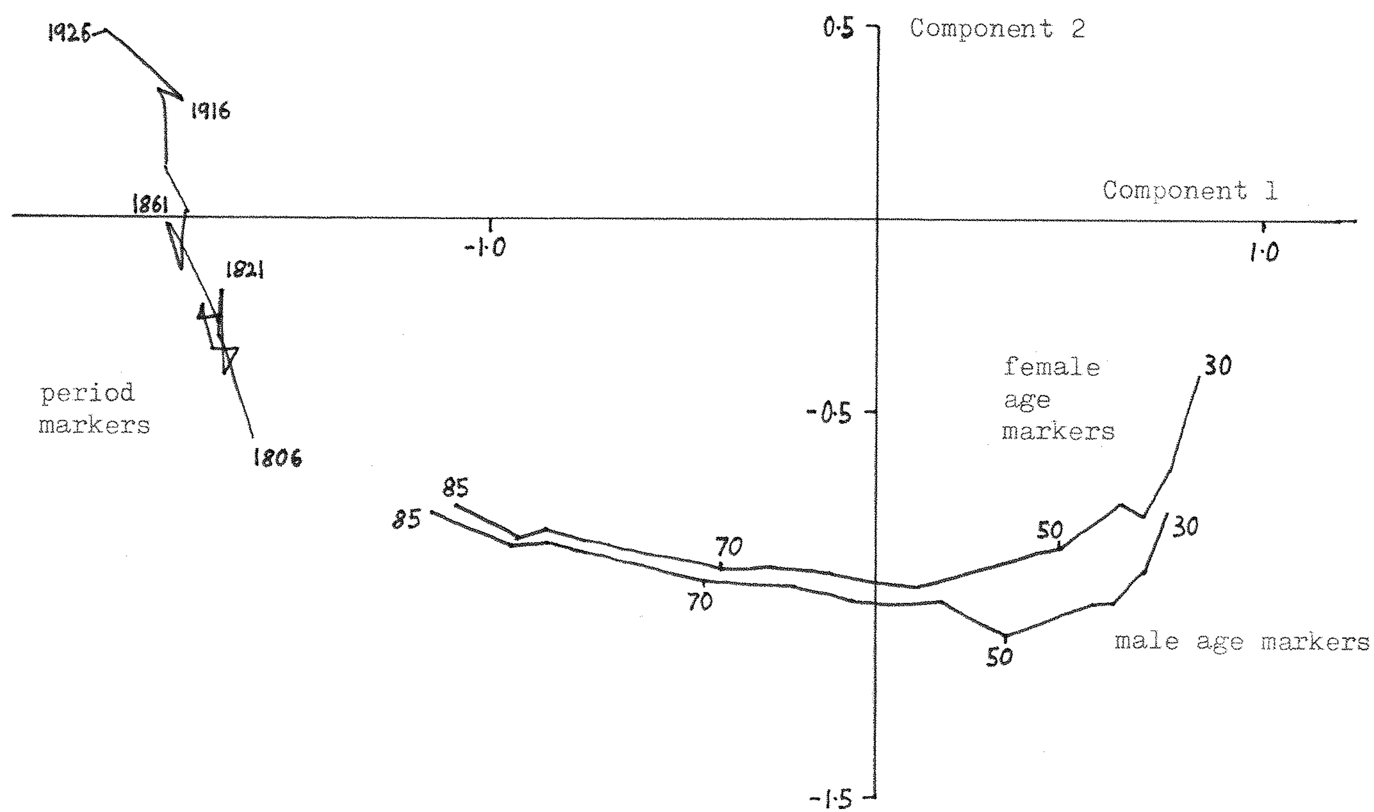


Figure 3.13: Biplot of mean corrected log mortality rates for males and females in Sweden, corresponding to model (3.18)

experience models (3.39) and (3.40) produce fits which are as close as those of the Gompertz (3.38) and bimodel (3.17), but make quite different substantive assertions. The interpretation of models (3.10), (3.39) and (3.40) is not easy: the effects of age, period and cohort factors in model (3.10) are inestimable, and the sensitivity of the cohort-experience parameter in (3.39) and (3.40) to changes in the functional form of the cohort-experience covariate is not at all reassuring. Model (3.40), however, does make reasonable substantive assertions, and the parameter estimates are in accord with intuition: that high mortality up to any age leads to high subsequent mortality. Model (3.40) also has the property that only one period parameter need be projected for each sex: given the similarity in period trends between the sexes, it looks as if a model which contains only one period parameter for both sexes combined might be developed from model (3.40).

That so many models fit this data extremely well is probably due to the small number of cells associated with each time-period: better discrimination between models might be attainable on the basis of single-year age-specific mortality data. It would also be interesting to compare these models for cause-specific mortality data: in particular different disease processes could indicate different formulations of cohort-experience type models.

3.6 Conclusions.

Several models or types of model have been fitted to a variety of aggregate level data above. The purpose in doing this was not simply to find explanations for these particular data sets, but to draw some wider conclusions concerning approaches to constructing and evaluating models of aggregate level demographic data in general. This section aims to draw together the experiences and results of this chapter under three headings: strategies for comparing models; evaluation of modelling strategies; and summary remarks.

3.6.1 Strategies for comparing models.

As stated in section 3.13 above, models should be compared on the basis of two criteria: goodness-of-fit and interpretability. The results of this chapter throw considerable light on how these two criteria should be evaluated and compared.

The very high degree of regularity in each of the data sets above, particularly when viewed from a cross-sectional perspective, suggests that very close fits should be attainable. In most cases very high R^2 values have been obtained, but these can be deceptive: this is well illustrated by the additive age-period model (3.11) applied to log mortality rates in section 3.5.2, which gives an R^2 value of .997 for females despite the evidently non-parallel age-profiles of the log rates in figure 3.11. Furthermore, all of the models produce clumping in the residuals: that is, areas of the data table which are consistently underestimated, or consistently over-estimated; although the clumping is generally less severe for the better fitting models. Thus even the best fitting models fail to capture all the systematic variation in the data.

Closeness of fit is influenced by model degrees of freedom: as a general rule, any model can be made to fit well by incorporation of sufficient period, cohort or age parameters. In order to take degrees of freedom into account, the adjusted R^2 statistic defined in equation (3.1) is presented in tables 3.1 to 3.4. However, as may be seen from these tables, this statistic does not seem repsonsive enought to degrees of freedom and rarely alters the ranking of models implied by the unadjusted R^2 statistic. Other methods of taking degrees of freedom into account also failed to produce satisfactory results.

Discrimination between models on the basis of fit is further impeded by the fact that often several equally well fitting models exist for one data set. This is particularly evident for the male mortality data of section 3.5, for which no fewer than three models score an R^2 value of .999 in table 3.4.

Thus it is seen that goodness-of-fit cannot be relied on to isolate one good model for a table of vital rates. The R^2 value may however be useful in eliminating models which fit badly, and models with large numbers of degrees of freedom may also be discounted.

Interpretability is a much less tangible quality, and may be thought of as comprising two separate components: the intuitive reasonableness of the theoretical foundation of the model, and the extent to which the parameter estimates may be reasonably interpreted in terms of that theoretical foundation. A model which does well in terms of the first of these components should make substantive assertions which strike a delicate balance between being too vague to be of any practical use, and being so specific that no account is taken of other processes which might well be of importance. The second component, concerning parameter estimates, can provide an acid test of the viability of a model: for example a model of fertility which produces estimates of target fertility which are either negative or greater than a certain amount would be immediately dismissed.

Interestingly in practice there seems to be a trade-off between goodness-of-fit and interpretability; for example, one of the most interesting and interpretable models discussed in this chapter is the achieved fertility model (3.22), yet its fit is not good. This trade-off is probably mainly due to the greater flexibility available when formulating models whose substantive assertions are very vague.

In concluding this subsection, it may be said that models of aggregate level demographic data may be compared only by an informal and intuitive balancing between closeness-of-fit, model degrees of freedom, reasonableness in theoretical assertions, and consistency of parameter estimates with those assertions.

3.6.2 Evaluation of modelling strategies.

The experiences of this chapter may be drawn together in order to evaluate various aspects of modelling strategy. Firstly the case for period factor effects versus cohort factor effects may be examined. The raw data themselves, in each of the data sets analysed, exhibit much greater regularity within periods than within cohorts. This may be seen in figures 3.4, 3.5, 3.11 and 3.12. It is hardly surprising therefore that cross-sectional specifications fit better than longitundinal specifications. In all of the data sets considered, the most recent cohorts are incomplete - as is generally the case for census and registration based data. All models fitted which contain cohort specific parameters produce very unstable and quite uninterpretable parameter estimates corresponding to these incomplete cohorts, and this also means that the closeness-of-fit of these models is artificially high. The instability of longitudinal specifications leads Brass (1974b) to advocate the use of cross-sectional specifications for projections. Thus, if a choice must be made between longitudinal and cross-sectional specifications of a model, the latter should be preferred.

Secondly, the question of whether age, period and cohort factor effects should be simultaneously built into a model may be addressed. The experience here seems to suggest that such factors can be usefully considered simultaneously, although automatically including all such dimensions is of little value, and this is well illustrated by the six-factor model (3.23) of legitimate fertility. The cohort-experience models demonstrate the variety of ways in which age, period and cohort can be combined. However, the period oriented bimodel generally fits as well as any age, period, cohort formulation, and the structure of the bimodel is in accord with many substantive mechanisms which would be expected to produce different age-structures at different times. Consequently, a clear case for the existence of any substantial cohort phenomena in addition to period phenomena cannot be made on the basis of the analysis above.

Thirdly, the way in which factors should be represented in a model may be considered. It may be said that none of the data sets examined above provide any real evidence of the existence of additive period or cohort factor effects on the scale of log rates. Furthermore, none of the substantive arguments lend much support to the notion either, although as a first approximation an additive period effect can be useful, as in the achieved fertility model (3.22). The bimodel, and mathematical trend curves such as the Gompertz, when applied cross-sectionally, do accommodate time-period factors which do not have an equal effect over all age groups within a time period; equally the cohort-experience models allow for cohort oriented mechanisms which do not produce equal effects over all age-groups within a cohort. These non-constant effect models are more reasonable intuitively, and also often fit better than models containing simple additive period or cohort factor effects.

If theoretical arguments suggest that age, period and cohort factors should be simultaneously represented in the model, then the way in which they are combined should be as simple as possible without being grossly inconsistent with the theoretical considerations: a certain amount of inconsistency may be tolerated in the interests of simplicity and interpretability, as in the case of the achieved fertility model (3.22) in which the possibility of a time-varying existing-children-effect is not accommodated. As intimated above, simply adding age, period and cohort factor effects together is perhaps too inconsistent with theory to be useful, and the attendent identification problem prohibits any compensation in terms of interpretability. Moreover, simply assigning some parameters to periods and others to cohorts in a theoretically based model age-schedule can produce serious interpretational difficulties, as in the treatment of the Coale-McNeil (1972) model in section 3.2.1 above.

Fourthly, the relative merits of the theoretical and the descriptive approaches to modelling may be assessed. On the one hand, a purely theoretical model (such as the versions of Lee's (1977, 1980) model (3.32) and (3.33)) may well fail to correspond to the data; and on the other, a model such as the bimodel (which might be suggested purely on the basis of the observed trends in cross-sectional age-structures) perhaps says too little about what types of period factors are involved, although the parameter estimates might lead to hypotheses. Interestingly, with highly aggregated data, it seems unlikely that an age, period, cohort formulation would be adopted through a purely descriptive approach, since clear differentials attributable to each dimension are generally difficult to pick out by eye: indeed any cohort-specific component to a model could be re-expressed as a form of age- period interaction. By definition, therefore, age, period, cohort formulations are theoretical. Conversely, theoretically it is hard to ignore altogether cohort-related phenomenon. Consequently it seems as though the two approaches cannot lead to the same model.

3.6.3 Summary remarks.

Aggregate level demographic data must surely conceal a wealth of important and potentially illuminating differentials on unmeasured variables, produced by a great complexity of causal processes. In fact the observed high levels of regularity across age (or duration) dimensions is itself evidence of a smoothing effect obtained from using an imprecise control for underlying variables. In these circumstances it is reasonable to ask whether there is any point in constructing inevitably vastly over-simplified theoretical explanations of the data. The answer lies in the hope that 'aggregate level substantive phenomena' exist; that is, causal mechanisms which dominate the variation in aggregated data. In this chapter the search for aggregate level phenomena has led to some interesting yet equivocal hypotheses, and perhaps more importantly, some sound perspectives on modelling highly aggregated demographic data.

The regularity across age or duration dimensions provides an impetus to discover an extremely well fitting model. In fact, it is not difficult to find several models of the same data having very high R^2 values, but it seems impossible to capture all of this regularity since residual patterns persist. This represents a double blow to the hopes of discovering a unique theoretical explanation of the data. Consequently purely objective criteria cannot be relied on to distinguish between models, and subjective assessments involving interpretability and substantive reasonableness must be heavily relied upon. Even so, for a given set of data, there is generally no one model indicated.

Accepting then, that there may be several equally good explanations of the data, it is still possible to use a theoretical model to examine the implications of theoretical assertions rather than to validate them. For example, Lee's (1977, 1980) model (3.33) may be used to estimate trends in period target fertility, on the assumption that such a concept has validity. Such an approach can lead to much deeper interpretations and more useful insights than would a more descriptive approach, as is well illustrated by the cohort-experience models of marriage and fertility. These interpretations however can only be held tentatively until suitable individual level data can be found to test the model.

The theoretical arguments in the preceding sections suggest that additive period or cohort factor effects are often somewhat implausible and that period factors are more easily defined and probably more powerful than cohort mechanisms, although it is generally more reasonable to accommodate both period and cohort related factors. The cohort-experience models possess all of these desired properties, and usually produce highly interpretable estimates (unlike the additive age, period, cohort model (3.10) whose parameters are not even estimable), even in the case of Lee's (1977, 1980) model (3.33) which does not fit at all well. However, this is not always the case: mortality cohort-experience model (3.39) is evidence of this. Moreover the slight change in specification of this model (3.39) to produce model (3.40) gives quite different results for males; also the achieved fertility model (3.22) gives unstable parameter estimates when estimated on the basis of a subset of the data. Thus it is suggested that steps are taken when usuing a cohort-experience formulation to investigate the robustness of the model to minor modifications in its algebraic structure and in its data base.

There are also attractions in using a less theoretical, more descriptive model such as the bimodel. To begin with, the cohort-experience term in the cohort-experience models approximates an age-structure, and consequently it could be held that the cohort-experience models sometimes accommodate the data simply because they approximate the two age-structure formulation of the bimodel. The bimodel, by making only weak substantive assertions, can in principle reflect a large number of causal processes, although the depth of interpretation may be limited. The bimodel is therefore less pretentious in a sense than the more theoretical models. The graphical capability of the bimodel, and its ability to easily incorporate both sexes, represent additional advantages. The Gompertz model essentially fulfills the same role as the bimodel, and has the advantage that it does not require age parameters, and the disadvantage that it generally requires three period parameters to achieve the same closeness-of-fit as the two period parameters of the bimodel, thus making projections less reliable. Interestingly the algebraic structure of the Gompertz model is not dissimilar to that of the cohort-experience models above, and as such these cohort-experience models represent generalisations of the Gompertz. Murphy (1982) has developed other generalisations of the Gompertz.

Both the bimodel and the cohort-experience models must sometimes be fitted to a subset of the available data. When the earliest cohorts in an age by period array are missing, the use of the Singular Value Decomposition in fitting the bimodel prohibits the use of the earliest incomplete periods. When the younger ages of the earliest cohorts are missing, the cohort-experience component in the cohort-experience models cannot be calculated, and so those early incomplete cohorts must be omitted. However, provided a long run of data is available, neither of these limitations should be of great consequence.

It would be interesting to apply the modification of Hernes (1972) model (3.16) and the achieved fertility model (3.22) to data from other countries. It would also be interesting to compare the cohort-experience models and the bimodel in relation to projection. To do this properly it would be necessary to try to link trends in period parameters with external data including macro-economic indicators. The two-sex bimodel would have an advantage over the cohort-experience models applied separately to each sex, through having only two period parameters to link externally. However, the achieved fertility model (3.22) has only one period parameter, and this could prove to be a powerful tool for projections. Projections are inevitably subject to wide margins of error, and therefore it is perhaps not important that model (3.22) does not fit as well as other models of the same data.

Chapter 4 - Proportional hazards models of World Fertility Survey data using birth-hirstory covariates.

4.1 Introduction

As seen in Chapter 3 above, there are severe limitations on the depth of reliable interpretation obtainable from highly aggregated demographic data. This is principally for two reason: firstly, realistic hypotheses usually involve unmeasured factors (such as the marriage model of Hernes (1972), which is based on 'marriageability' and 'pressure to marry' factors); and secondly, several quite different hypotheses can usually be found which fit the data equally well. In this chapter, individual level fertility data from a wide variety of World Fertility Survey (WFS) countries are analysed. The larger number of variables available with the data permit a much greater depth of reliable interpretation.

The approach in this chapter is to attempt to find a 'universal' model of fertility in developing countries. Such a model (if it exists) could provide valuable insights into mechanisms affecting fertility, and provide a firm basis for projection and for comparisons between countries. Casterline and Hobcraft (1981) have also set out with this aim, using variables derived from the birth-history including: 'age' variables (age, duration of marriage, duration of motherhood); 'age-at-entry' variables (age-at-marriage, age-at- motherhood); 'cohort' variables (birth-cohort, marriage-cohort, motherhood-cohort); and time-period. These and other variables calculated from dates in the birth-history may be referred to collectively as 'birth-history covariates'. The attractions of using birth-history covariates are that they are well-defined, generally well-measured, and can reflect underlying factors of fertility which have been poorly recorded or or which have not been recorded at all. The disadvantages are that they cannot themselves be considered to be direct causal factors, and there can be no guarantee that their relationships with fertility should have the same causal interpretation for different countries.

The present analysis, like that of Casterline and Hobcraft (1981), concentrates mainly on fertility differentials with respect fo birth-history covariates, although in the present analysis a more extensive set of birth-history covariates and a different technique of analysis are used. The data used here are, however, the same as those used by Casterline and Hobcraft (1981), being WFS data from Bangladesh, Columbia, Indonesia, Jamaica, Jordon, Kenya, Korea, Mexico and Sri Lanka. Amongst the developing countries these nine are believed to be geographically, economically and demographically diverse, and their data are of reasonable quality and sample size. Any model which holds across such a diversity of populations might reasonably be expected to hold for a large number of other populations.

Some of the birth-history covariates which previous research has shown to be important in fertility analysis are now discussed.

Numerous researches have demonstrated and modelled the relationships between age and fertility, (for example: Coale and Trussell, 1974; Page, 1976; Hoem et al, 1981: Casterline and Hobcraft, 1981). The popularity of the age variable derives partly from its availability in registration, census and survey data; but its association with fecundity (Henry, 1961), the biological ability to reproduce, is undoubtedly the principal causal factor underlying the relationship (although Rindfuss and Bumpass (1978) discuss age effects which are not simply physiological in origin).

Recent work has shown the importance of age-at-marriage (Ruzicka, 1976; McDonald et al, 1980; Finnas and Hoem, 1980) and age-at-motherhood (Bumpass et al , 1978) in determining subsequent fertility. One explanation (Bumpass et al , 1978) is that early fertility limits the pursuit of other life options which might later compete with child-bearing; another is the so-called 'catch-up' effect (Kendall, 1979; McDonald et al, 1980; Freedman and Casterline, 1981) of more rapid childbearing among women with later ages-at-marriage or motherhood; educational attainment may also be reflected in fertility differentials by age-at-marriage or motherhood.

Duration of marriage (Page, 1976) and of motherhood (Casterline and Hobcraft, 1981) have also been shown to be associated with fertility. Coital frequency could be partly responsible, but it is possible that these duration variables are largely a substitute for birth-order. Braun (1980) demonstrates a dominant relationship between the average length of previous birth-intervals and current fertility: this variable is a function of duration and birth-order, and measures previous fertility levels.

Birth-, marriage-, and motherhood-cohort may also be useful variables, although it is difficult to argue convincingly for the presence of causal mechanisms which would be clearly linked to them. Hobcraft et al (1979) and Casterline and Hobcraft (1981) suggest possible mechanisms, but it has not been demonstrated empirically that they are efficiently represented by the cohort dimensions.

Time-period may act as a surrogate for a whole set of contemporaneous influences, including economic circumstances and availability of contraception. Page (1976) and Casterline and Hobcraft (1981) have clearly demonstrated the importance of this variable.

Lastly, the time since the previous birth is also responsible for large differentials in fertility, since within this time fertility may be affected by post-partum abstinence and amenorrhoea; also gestation time effectively prevents births occurring within eight months of the previous birth.

The analysis of Casterline and Hobcraft (1981) does not take explicit account of birth-order or time since the previous birth: these are potentially series omissions. Their methodology is essentially multiple classification analysis, in which a model parameter is assigned to each category of each control (as in the additive age-period-cohort model (1.3) of chapter 1): this is a rather extravagant use of degrees of freedom in the present circumstances, where a large number of effects and interactions are to be modelled. The present analysis concentrates on efficiently parameterising effects and interactions, and pays particular attention to the effects of, and interactions with, birth-order and time-since-previous-birth.

The use of 'time-since-previous-birth' indicates a life-table approach to the analysis of fertility within birth-intervals. Hobcraft and Rodriguez (1980) and Rodriguez and Hobcraft (1980) have performed non-parametric life-table analyses of fertility within birth-intervals for World Fertility Survey data, controlling for various birth-history covariables. Finnas and Hoem (1980) and Bumpass et al (1978) have treated other birth-history data in a similar way. The difficulty with these non-parametric analyses is that the numbers of women within subgroups become too small for useful analysis unless the number of control variables is strictly limited (to about two variables) and the number of categories for each control variable is also strictly limited (to at most four categories).

Because of these limitations, several researches have attempted to model birth-interval life-tables, so that only two or three parameters are needed for each subgroup. Hobcraft and Rodriguez (1980) make a start at this, suggesting several possible models. Their attention, however, is focussed on capturing the general shape of the hazard function (see definition below) of birth-intervals rather than on comparisons between subgroups. Stoto and Menken (1977) attempt to build a model of the birth-interval hazard function from substantive considerations; again, their interest is with what happens within the birth-interval rather than between subgroups. The present analysis requires a model which parameterises efficiently not only the hazard function, but also the impact of birth-history covariates on the shape and level of the hazard function.

Braun (1980) represents the birth-interval density function with a Gamma density; this analysis incorporates birth-history covariates in the manner of a regression, but is applicable only to closed birth-intervals - that is, birth-intervals which are terminated by a subsequent birth. This approach is inappropriate to the present data in which many birth-intervals are censored (i.e. terminated by interview): omitting the censored birth-intervals can cause series biases. Nevertheless this approach is much closer to the present requirements. Braun and Hoem (1979) model the birth-interval hazard function using a Gamma distribution, where birth-intervals may be either closed or censored, but the algebraic structure of their model does not easily generalise to accommodate covariables.

Cox (1972) sets out a much more general class of life-table models, called 'proportional hazards' models, incorporating regression on any choice of covariables and applicable even in the presence of censoring. These models are therefore ideally suited to the present requirements. Menken et al (1981) have used proportional hazards models to investigate socio-demographic influences on marriage dissolution. A general class of proportional hazards models may be written:

$$\ln \ \lambda(t) = \underline{\beta}' \ \underline{X}(t) \tag{4.1}$$

where $\lambda(t)$ is the hazard at survival-time t (that is, the instantaneous rate of decrement from the life-table population, due to the event of interest, t units of time after the start of the life-table), $\underline{\beta}$ is a vector of parameters and $\underline{X}(t)$ is a vector of covariate values at t. (Expression (4.1) implies that the survival-time main-effect term, usually referred to as the 'baseline hazard', is parameterised as a linear combination of known functions of survival-time. Cox (1972) does not attempt to parameterise the baseline hazard, treating it as a nuisance function. For present purposes, however, the shape of the baseline hazard is of interest).

In the present context, the life-table population comprises a set of individuals about to commence a birth-interval, the event of interest is a subsequent birth, censoring occurs when the interview occurs before a subsequent birth, and the hazard at t is the instantaneous fertility (or 'force' of fertility) t units of time after commencing the birth-interval amongst those individuals who have survived to time t (i.e., who have not been removed from the life-table population before time t by interview or by a subsequent birth). Note that, for present purposes, a woman may represent more than one 'individual' if she has more than one birth-interval; for each birth-interval her covariates $\underline{X}$ may take different values (for example, one covariate could be birth-order).

Now the birth-history covariates discussed above may all be considered as fixed for an individual within the birth-interval (even age and duration of marriage or motherhood, which actually vary with survival-time, may be accommodated within this framework, as will be shown below), and it is both realistic and efficient to assume that the effects of these covariates on the hazard function vary smoothly with survivial-time. Consequently the model required here should be in the form:

$$\ln \lambda(t) = \mu(t) + \underline{\beta}'(t)\ \underline{X} \qquad (4.2)$$

where $\mu(t)$ is the baseline hazard, $\underline{\beta}(t)$ is a vector of the effects of $\underline{X}$, the fixed covariates, and $\mu(t)$ and all the $\underline{\beta}(t)$ are smooth functions of t. Now model (4.2) looks rather different to model (4.1), but it is demonstrated in appendix 4.B that provided that each of the functions $\mu(t)$, $\underline{\beta}(t)$, is a linear combination of known functions of survival-time, then (4.2) is a special case of (4.1). Choosing an appropriate functional form for $\mu(t)$ and $\underline{\beta}(t)$, for example a cubic polynomial, will ensure their required smoothness. Polynomials, however, are likely to produce undesirable characteristics at the extremes of survival-time; consequently the functional form of a cubic spline with 'knots' at 10, 20, 40, and 80 months survival-time is chosen. Further details of this are given in appendix 4.B but for the present all that need be understood is that each of the functions $\mu(t)$, $\underline{\beta}(t)$ is completely determined by specifying the values of four model parameters which represent the function values at 10, 20, 40 and 80 months survival-time; the function values at other survival-times being generated by drawing a special type of smooth curve (called a cubic spline) through these four known points. This parameterisation permits efficiency and flexibility in accommodating the data.

Cox (1972) maximises a 'partial' likelihood function to estimate the parameters of model (4.1). This is not convenient for the present data and models, and instead an approximate maximum likelihood solution is obtained, as described in appendix 4.A below.

Model (4.2) provides a framework for a wide variety of specifications. For example, age-at-marriage may be represented by a set of dummy covariates, one for each five year category of age-at-marriage; alternatively and more efficiently, a quadratic or even simply linear dependence on age-at-marriage may be utilised. Variables such as age, or duration of marriage or motherhood, which vary with survival-time, may be accommodated within the model by replacing them with their values at the start of the birth-interval. For example, for a linear dependence on age, this substitution does not affect the estimated age effects, although it does affect the baseline hazard.

Sets of variables such as age-at-marriage, marriage duration and age; or birth-cohort, age and time-period; are logically related since in each case the third variable is equal to the sum of the other two. If all three variables in a logical relationship are included in the model then estimation problems arise (see Chapter 1). In the present analysis this problem is avoided by including at most two of a set of three logically related variables. Since, for the greater part of the analysis, only linear dependences on birth-history variables are utilised, and since a linear dependence on any two of the variables in a logical relationship automatically embodies a linear dependence on the remaining variable, then this procedure does not represent any disadvantage to the omitted variable.

In the sequel a 'current' birth-interval is a birth-interval contributing to the life-table population currently being investigated; the 'previous' birth-interval is the birth-interval immediately preceding the current birth-interval; a 'prior' birth-interval is any birth-interval which precedes the current birth-interval; and the k^{th} birth-interval is the birth-interval following the $(k - 1)^{th}$ birth (or marriage or first union if $k = 1$).

In this section a class of proportional hazards model (4.2) with survival-time varying covariate effects has been developed, and the use of covariates constructable from the birth-history has been discussed. The following section, 4.2, outlines the stages gone through in selecting from these covariates to derive a final model of the form of (4.2) applicable to all nine countries analysed here. The emphasis in this derivation is on the extraction of patterns, and detailed interpretation of the intermediate and final models is postponed until section 4.3.

4.2 Deriving the Model.

There are a large number of effects and interactions which could be present in the data, but it is not practical to look for them all simultaneously as this would involve excessive computation. Therefore, the initial stages in the analysis reflect some prior judgements concerning the likely importance of effects and interactions, although later stages include checks on some of these assumptions. Specifically, it was thought that fertility control mechanisms might produce interesting interactions with survival-time and with birth-order, and so these were given full expression where possible. In general only linear or quadratic relationships with other variables were considered.

The first stage of the analysis involved the use of a forward selection procedure to indicate terms which help to provide a good fit to the data. The procedure begins with the baseline hazard, and at each subsequent step examines each of the remaining terms before adding to the model that term which gives the greatest improvement in fit, as measured by the χ^2 statistic. However, it does not examine terms already in the model for possible exclusion, and it is possible that this deficiency could lead to terms of importance being overlooked.

To begin with, for countries and birth-orders separately, the forwards selection procedure was used on the linear main-effect variables listed in table 4.1. Some separate checks indicated that omission of interactions with survival-time at this stage would not cause any variable of importance to be overlooked. The results for birth-order 3 (that is, the birth-interval following the third birth) are shown in table 4.1. Results for birth-orders 1, 3 and 5 are given in appendix tables 4.E1 to 4.E3. The results indicate, for birth-orders 2 to 6,

	Step 1		Step 2		Step 3	
Bangladesh	ABI (PBI)	215.	FBA,PBA	12.	PBI	6.
Columbia	ABI (PBD,PBA,PBI)	129.	FBD,PBD (FMD)	96.	EDU	41.
Indonesia	ABI	251.	FBD,PBD	62.	FBI,FMD (FBA,PBA,OBD)	16.
Jamaica	ABI	42.	FMD	9.	EDU (FBA,PBA,FMA,OBD)	4.
Jordan	ABI (PBI,EDU)	55.	EDU	47.	OBD	13.
Kenya	PBI (ABI)	109.	ABI (EDU,FBD)	14.	EDU	11.
Korea	PBA (EDU)	187.	EDU	117.	PBI (ABI,FBA)	27.
Mexico	PBA (ABI)	197.	EDU	63.	ABI,FBA	64.
Sri Lanka	PBA	281.	ABI,FBA (PBI)	65.	EDU (FBD,PBD,OBD,FMD)	31.

FBI	length of first birth-interval (from first marriage or first union to first birth)
PBI	length of previous birth-interval
ABI	average length of birth-intervals between first and previous births
FBA	age at first birth
PBA	age at previous birth
FMA	age at first marriage or first union
FBD	date of first birth
PBD	date of previous birth
FMD	date of first marriage or first union
OBD	date of mother's own birth
EDU	length of full-time education

Table 4.1. The first three steps of a forwards selection amongst linear main-effect terms for birth-order 3, with χ^2 values corresponding to the selected terms. Terms enclosed in parentheses have χ^2 values within 75% of that for the selected term, and are listed in order of decreasing χ^2 . When more than one term is selected in a single step, this is due to logical relationships between the selected terms (see section 4.1) Terms with $\chi^2 < 3.0$ are not shown. (Each main-effect term represents one degree of freedom.)

that ABI (average length of birth-intervals between first and previous births), PBA (age at previous birth), EDU (length of full-time education) and PBD (date of previous birth) are, in that order, the most important variables to include in the model. For birth-order 1 the pattern is different, not surprisingly since ABI is not defined for this interval, and other variables are identical to each other; here FMA (age at first marriage or first union) and FBI (length of first birth-interval, from first union to first birth) are the most important.

Further experiments using the forwards selection procedure revealed that duration of marriage is unimportant, as are quadratic components in these variables with the possible exceptions of ABI and PBD.

In view of these results, for the next stage, the model:

$$\ell n\ \lambda(t) = \mu(t) + \alpha(t).ABI + \beta(t).PBA + \gamma(t).PBD + \delta(t)EDU \qquad (4.3)$$

was fitted for countries and birth-orders separately. (The same model, but without the ABI interaction, was fitted for birth-order 1, despite the anomalous behaviour of this birth-order noted above, in the interests of finding a model consistent over all birth-orders and countries.) It was found that the ABI and PBA interactions with survival-time are fairly regular and consistent across birth-orders and countries, but the PBD and EDU interactions are quite erratic.

Replacing the PBD and EDU terms with their main-effect counterparts yields the model:

$$\ell n\ \lambda(t) = \mu(t) + \alpha(t).ABI + \beta(t).PBA + \gamma.PDB + \delta.EDU \qquad (4.4)$$

which was fitted for countries and birth-orders separately, and the results are given in figure 4.1. All parameter structures are broadly similar across countries, excepting Korea. Across birth-orders the

$\mu(t)$ and $\beta(t)$ estimates are remarkably consistent, although the smaller sample sizes for the higher birth-orders introduce some instability. The most striking feature of these results, however, is the roughly linear trend in the $\alpha(t)$, γ and δ estimates across birth-orders. In particular $\alpha(t)$ is roughly proportional to (birth-order - 1). This is illustrated more clearly for $\alpha(20)$ and $\alpha(40)$ in figure 4.2. It was found that the addition of quadratic components to the ABI and PBD terms in model (4.4) does little to improve the fit, although the quadratic ABI parameter estimates do tend to moderate the effects of the larger ABI.

The results above suggest that a model for all birth orders including the first may be constructed as follows:

$$\ell n\ \lambda(t) = \mu(t) + \varepsilon_1.\text{BOR} + \alpha(t).(\text{ABI} + \varepsilon_2).(\text{BOR} - 1) + \beta(t).\text{PBA} + \gamma.\text{PBD} + \theta.\text{BOR.PBD} + \delta.\text{EDU} + \phi.\text{BOR.EDU} \qquad (4.5)$$

where BOR denotes birth-order. Note that ε_1 is introduced to prevent the model reflecting the arbitrary normalisations of PBD and EDU given in table 4.2, and ε_2 is similarly introduced in respect of ABI. Model (4.5) may be reparameterised and written:

$$\ell n\ \lambda(t) = \mu(t) + \eta(t).\text{BOR} + \alpha(t).\text{MOD} + \beta(t).\text{PBA} + \gamma.\text{PBD} + \theta.\text{BOR.PBD} + \delta.\text{EDU} + \phi.\text{BOR.EDU} \qquad (4.6)$$

where MOD is duration of motherhood (which appears because MOD is equivalent to $\{\text{ABI}.\ (\text{BOR} - 1)\}$ ignoring multiple-births), and where the following parameter constraint holds:

$$\eta(t) = \varepsilon_1 + \varepsilon_2.\alpha(t) \qquad (4.7)$$

Figure 4.1. Parameter estimates from model (4.4). In panels (a), (b) and (c), for each country and birth-order, parameter estimates at 10, 20, 40 and 80 months of survival-time are connected with straight lines(although strictly they should be connected with cubic spline curves). In panels (d) and (e) straight lines connect parameter estimates across birth-orders. Vertical lines of length one standard error are drawn to each side of each parameter estimate, (standard errors for the baseline hazard are too small to be shown). The scale for each panel is indicated in its upper right corner. A broken line indicates that the line should continue. The normalisation of the variables is given in table 4.2.

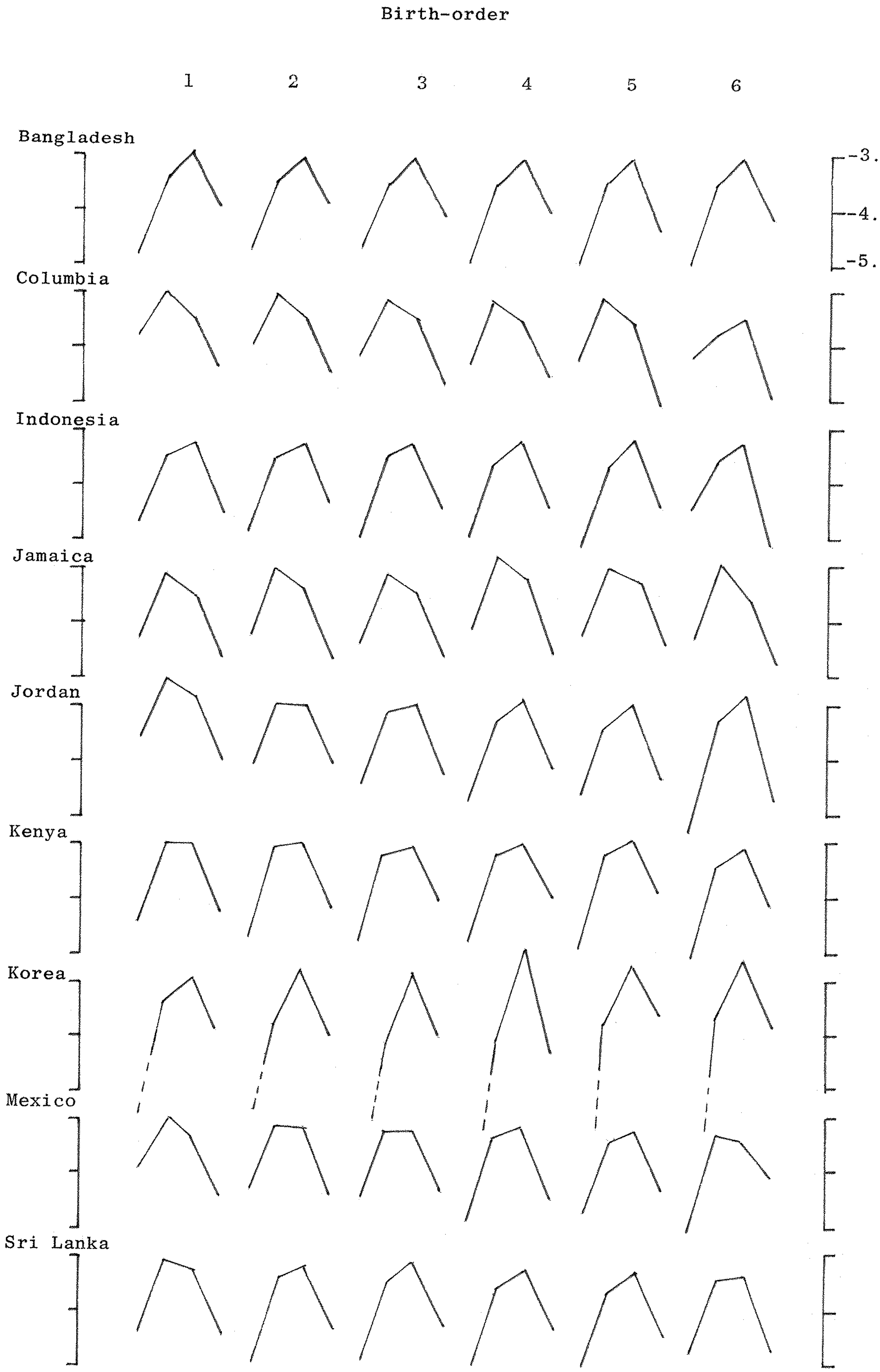


Figure 4.1(a): Baseline hazard, $\mu(t)$, from model (4.4)

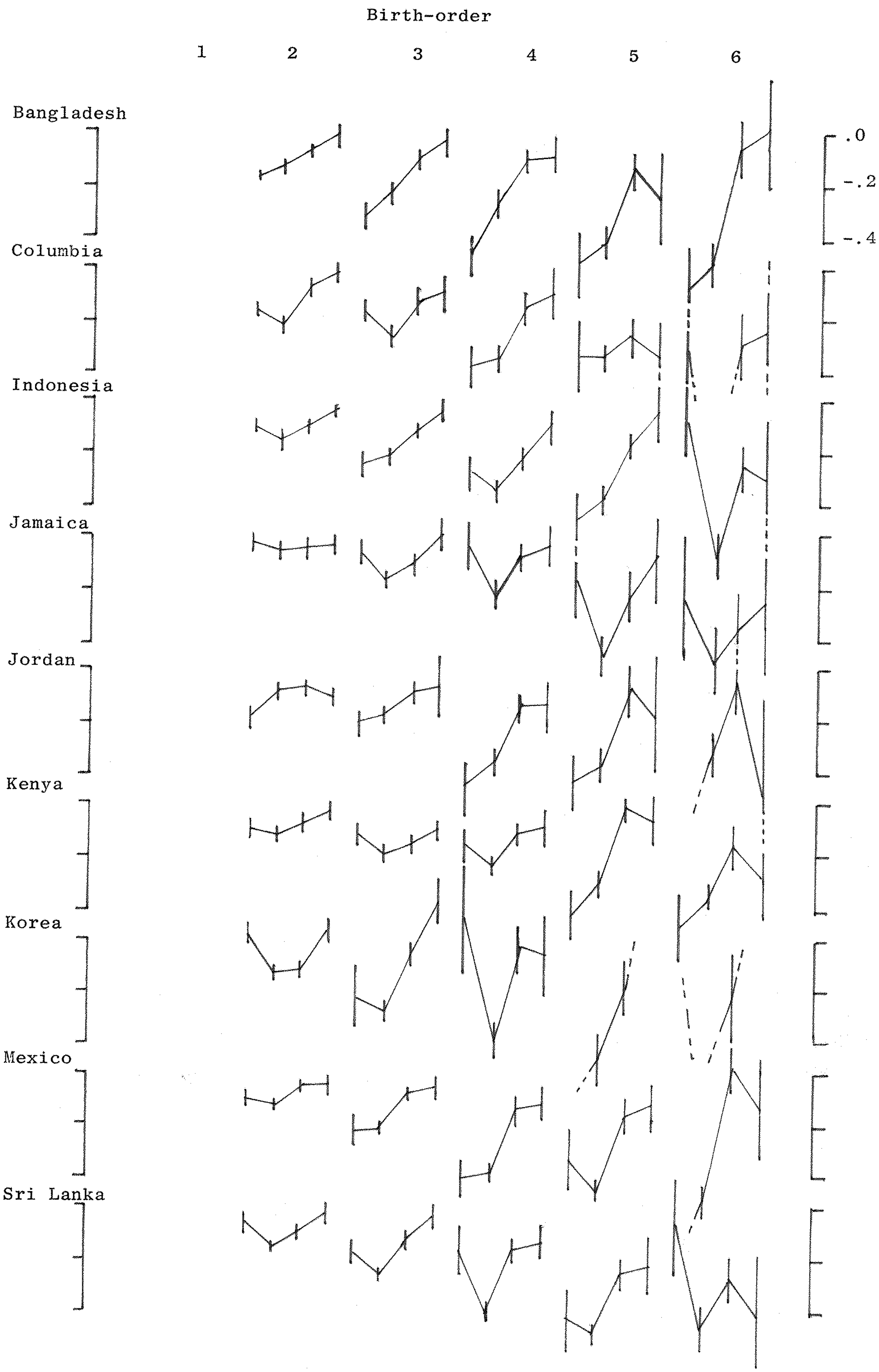


Figure 4.1(b): ABI effects, $\alpha(t)$, from model (4.4)

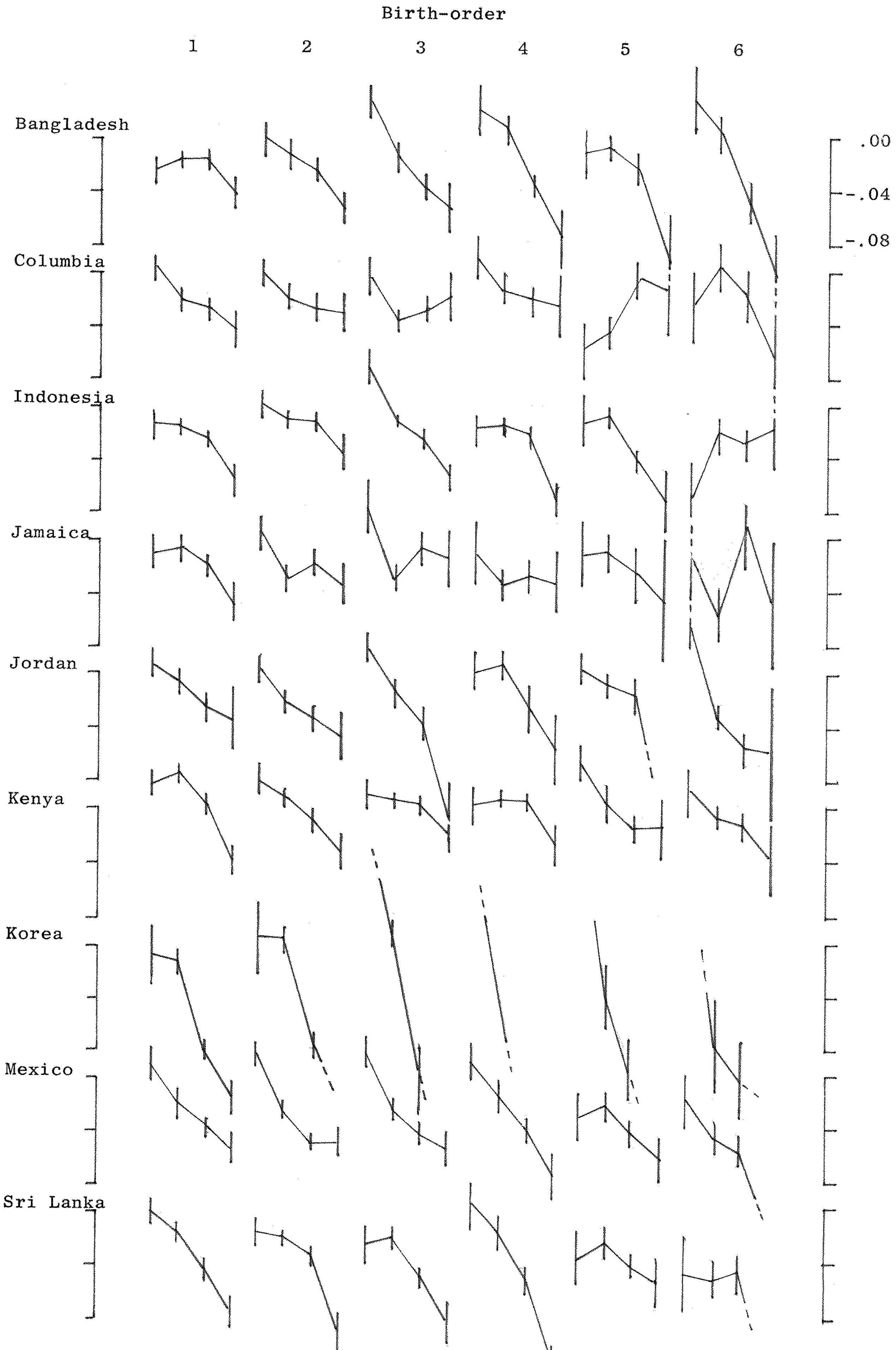


Figure 4.1(c): PBA effects, β(t), from model (4.4)

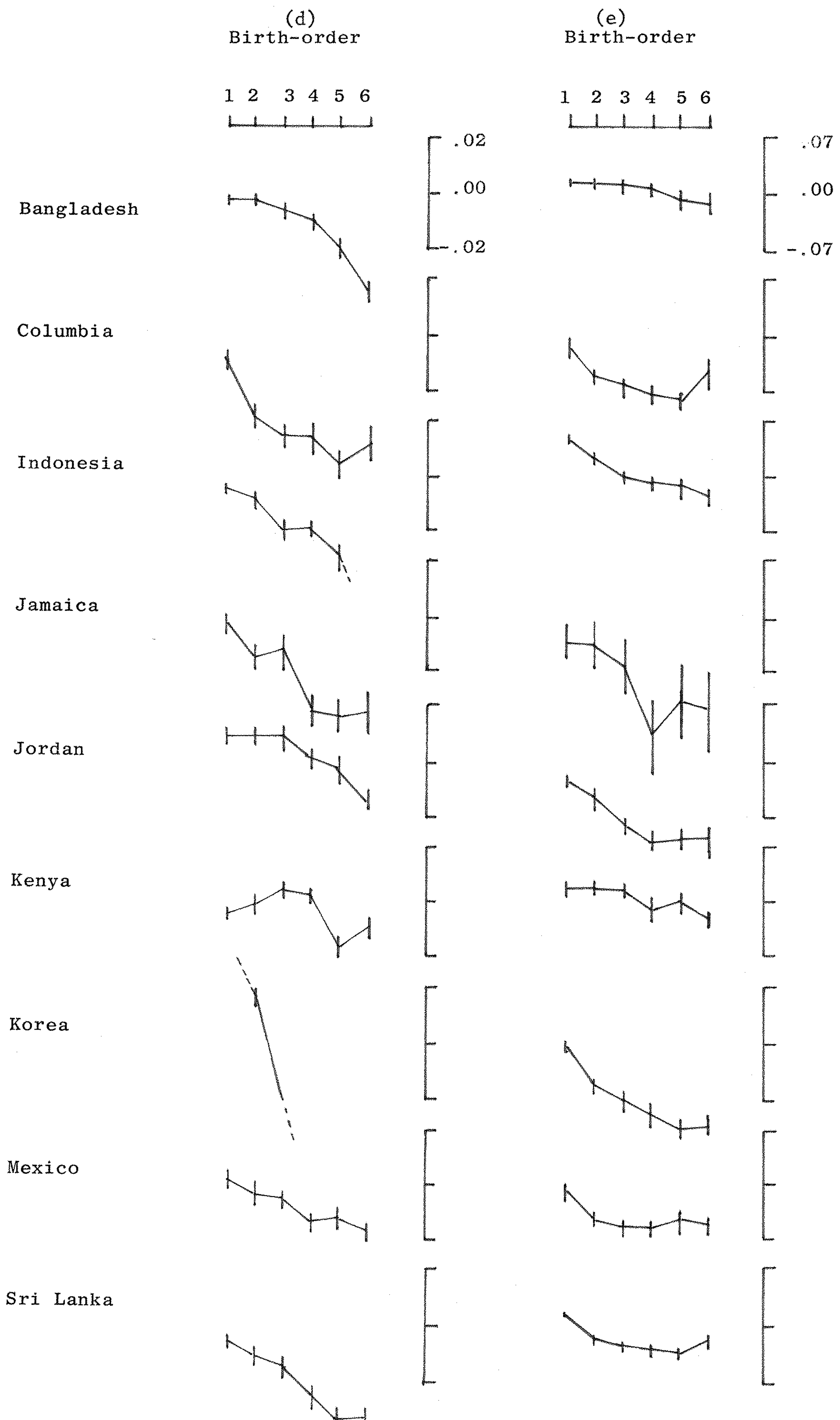


Figure 4.1: d) PBD effecfs,γ, and e) EDU effects,δ, from model (4.4)

Table 4.2 Normalisation of Variables[a]

Variate	Location	Scale
ABI	3 years	1 year
PBA	25 years	1 year
PBD	1960 years	1 year
EDU	3 years	1 year
BOR	3	1
MOD	10 years	1 year
PBI	2.5 years	1 year
LNPBI	ℓn (2 years)	1
CEP	0	1
BFD	0	1

a. For example, EDU is the number of years of full-time education, less 3.

Figure 4.2. Estimates of ABI effects in model (4.4) at a) t = 20 months and b) t = 40 months, across birth-orders. Vertical lines of length one standard error are drawn to each side of each parameter estimate. The scale for each panel is indicated in its upper right corner. Normalisation is given in table 4.2.

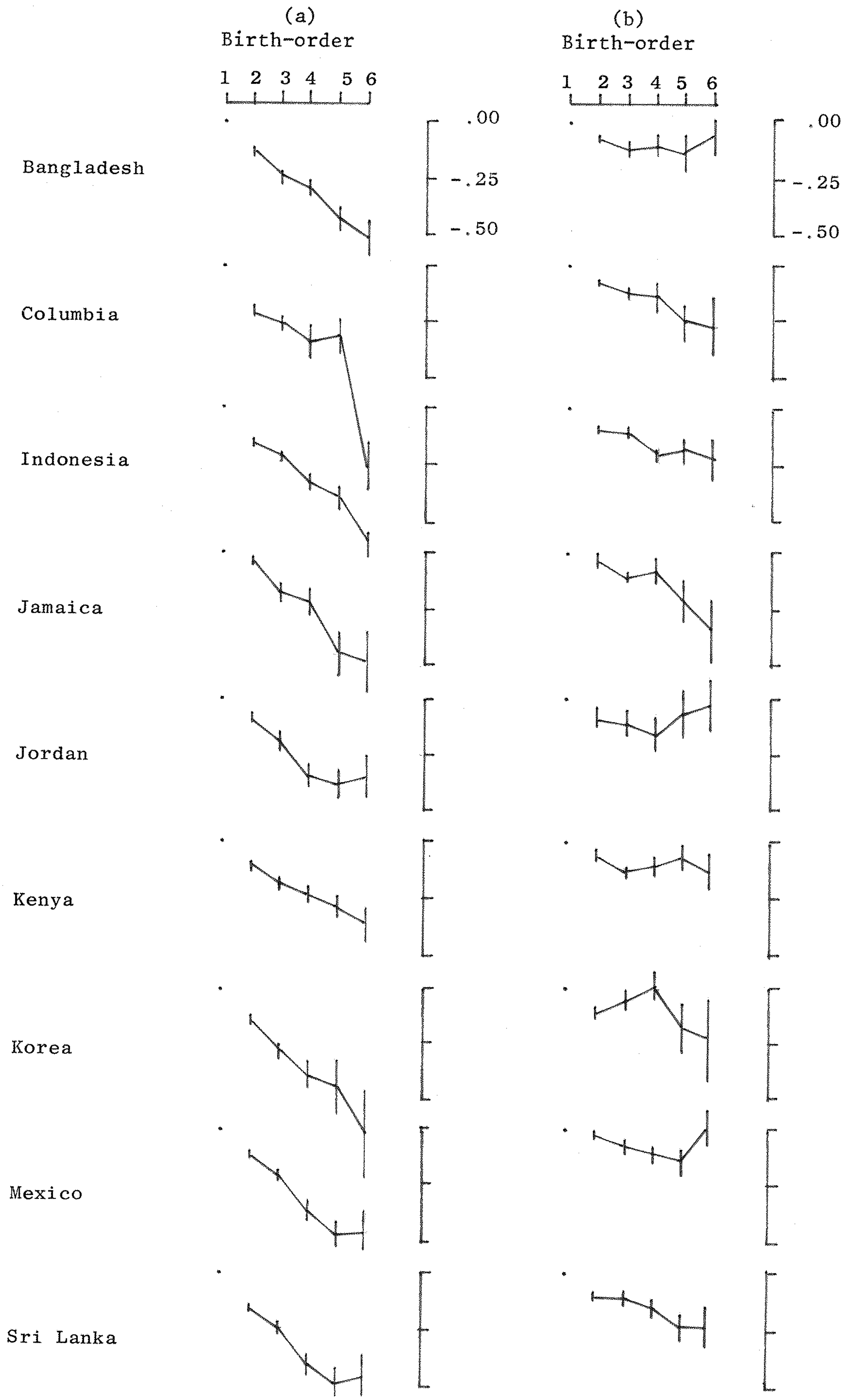


Figure 4.2: ABI effects a) $\alpha(20)$, and b) $\alpha(40)$ from model (4.4)

Figure 4.3. Parameter estimates from model (4.6). For $\mu(t)$, $\eta(t)$, $\alpha(t)$ and $\beta(t)$, estimates at $t = 10$, 20, 40 and 80 months are connected with straight lines. Vertical lines of length one standard error are drawn to each side of each parameter estimate. The scale for each effect is the same for all countries and is indicated for Bangladesh. A broken line indicates that the line should continue. Normalisation is given in table 4.2.

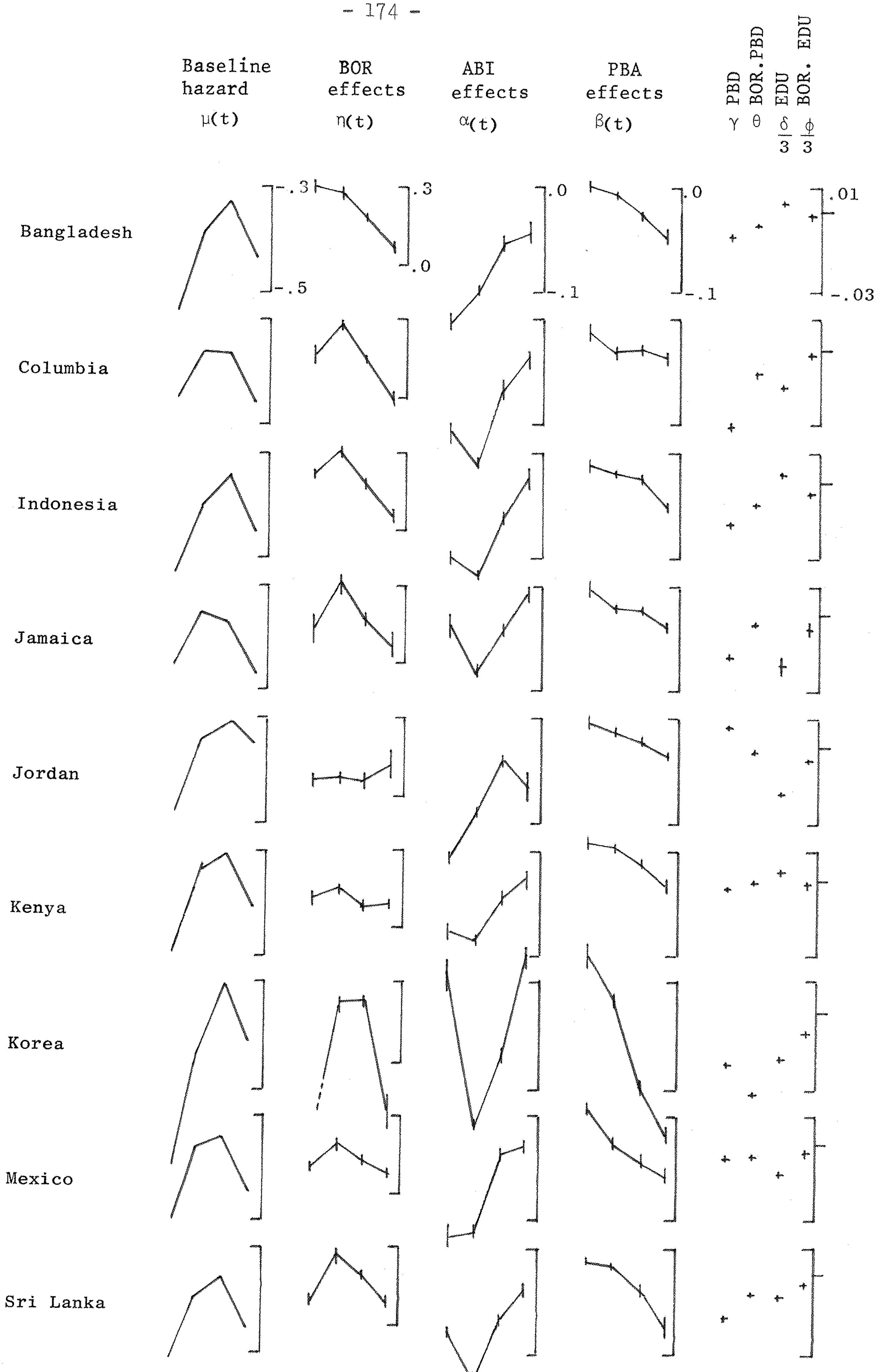


Figure 4.3: Parameter estaimtes from model (4.6)

Table 4.3 Difference in χ^2 between model(4.4)fitted to birth-orders 1 to 6 separately, representing a total of 80 degrees of freedom, and model(4.6)fitted to birth-orders 1 to 6 simultaneously, representing 20 degrees of freedom.

Country	χ^2
Bangladesh	93.
Columbia	153.
Indonesia	151.
Jamaica	68.
Jordan	99.
Kenya	120.
Korea	200.
Mexico	103.
Sri Lanka	141.
d.f.	60

Fitting model (4.6) with constraint (4.7) is inconvenient because of the non-linearity of (4.7). Alternatively, if ε_1 and ε_2 are roughly estimated for each country from previous results, then constraint (4.7) is linear in the free parameters and consequently less inconvenient. The approach adopted, however, was to ignore constraint (4.7) temporarily, and the resulting parameter estimates are given in figure 4.3. As anticipated, the $\eta(t)$ and $\alpha(t)$ parameter estimates in general reflect relationship (4.7), the only exception being Jordan where evidently the $\eta(t)$ parameters are picking up the slight trend across birth-orders in $\mu(t)$ in figure 4.1(a). Table 4.3 contains the χ^2 differences between model (4.4) applied to birth-orders separately, and model (4.6) applied to birth-orders simultaneously. These χ^2 values are generally in the upper 1% tail of the distribution and so, on the basis of a strict statistical test, model (4.4) would be preferred. However, model (4.6) does conveniently summarise the results in figure 4.1, and the χ^2 values in table 4.3 are in general much less significant than the corrected χ^2 values for the remaining terms in model (4.6), given in appendix table 4.E4. In balance, therefore, it seems sensible to proceed with model (4.6).

The fact that relationship (4.7) holds suggests that interpretations should be in terms of model (4.5) rather than unconstrained model (4.6), i.e. in terms of factors underlying ABI rather than MOD. The interpretation is then as follows: ABI is acting as a proxy for underlying factors such as length of breastfeeding, use of contraception and fecundity; the accuracy with which ABI captures the effects of these factors increases with the additional information contributed by each successive birth-interval, giving rise to the observed trend of increasing effects of ABI with increase in BOR. This is discussed in more detail in the next section, and demonstrated algebraically in appendix 4.D.

If the above interpretation is correct then PBI (the length of the previous birth-interval) should act as an alternative to ABI, since it should be able to capture the effects of the factors underlying ABI. However, the effects of PBI would not be expected to increase with BOR since PBI only ever contains information from one birth-interval. It would be expected, then, that replacing ABI in model (4.4) with PBI to give:

$$\ell n\ \lambda(t) = \mu(t) + \alpha(t).\text{PBI} + \beta(t).\text{PBA} + \gamma.\text{PBD} + \delta.\text{EDU} \qquad (4.8)$$

and fitting to birth-orders separately would produce parameter estimates similar to those for model (4.4) except that $\alpha(t)$ would be unaffected by changing BOR. This would permit construction of a simultaneous birth-orders model analogous to model (4.6):

$$\ell n\ \lambda(t) = \mu(t) + \eta(t).\text{BOR} + \alpha(t).\text{PBI} + \beta(t).\text{PBA} + \gamma.\text{PBD} + \theta.\text{BOR.PBD} + \delta.\text{EDU} + \phi.\text{BOR.EDU} \qquad (4.9)$$

where, in place of constraint (4.7):

$$\eta(t) = \eta \qquad (4.10)$$

because the second term in (4.9) is introduced only to prevent the model reflecting the arbitrary normalisations of PBD and EDU. It would be expected that parameter estimates from model (4.9) would be similar to those from model (4.6) except that the $\eta(t)$ estimates should correspond to constraint(4.10), but the fit of model (4.9) would be expected to be worse than that for model (4.6) owing to the omission of information from earlier birth-intervals.

All of these expectations are fulfilled. The constancy of PBI effects in model (4.8) across BOR may be seen from figure 4.4, and the profiles of the unconstrained $\eta(t)$ estimates from model (4.9) shown in figure 4.5 are considerably attenuated in comparison with those in figure 4.3 from model (4.6). Table 4.4 shows that with one exception model (4.9) does not fit as well as model (4.6).

Model (4.9) possesses a distinct advantage over model (4.6) despite the fact that it does not fit as well: it is not encumbered with an awkward non-linear constraint. Moreover, the loss in precision through using PBI rather than MOD is not great in comparison to the total explanatory power accountable to MOD, as may be judged by comparing table 4.1 and appendix table 4.E4. For these reasons the ensuing analysis develops model (4.9) rather than model (4.6).

Intuitively a proportionate increase in PBI might be expected to correspond to a proportionate decrease in current fertility, indicating a linear relationship between log PBI and log hazard. This suggests that large PBI might be detrimentally affecting the fit of model (4.9), and a better formulation would be as follows:

$$\ell n\ \lambda(t) = \mu(t) + \eta(t).\mathrm{BOR} + \alpha(t).\mathrm{LNPBI} + \beta(t).\mathrm{PBA} + \gamma.\mathrm{PBD} + \theta.\mathrm{BOR.PBD} + \delta.\mathrm{EDU} + \phi.\mathrm{BOR.EDU} \qquad (4.11)$$

where LNPBI is ℓn(PBI - 6 months), (removing from PBI a conservative estimate of gestation time). From table 4.4 model (4.11) tends to fit better than either (4.9) or (4.6). All parameter estimates in model (4.11) are similar to those in model (4.9).

Constraint (4.10) is applicable to model (4.11) and table 4.5 shows that the χ^2 differences produced by this constraint, and by the additional constraint $\phi = 0$, are small in comparison to the corrected χ^2 values for other terms in the model (see table 4.6). There is less justification for additionally setting $\Theta = 0$, although if this is done then one could reasonably omit BOR from the model altogether, by placing $\eta = 0$. These results suggest the model:

$$\ell n\ \lambda(t) = \mu(t) + \alpha(t).\mathrm{LNPBI} + \beta(t).\mathrm{PBA} + \eta.\mathrm{BOR} + \gamma.\mathrm{PBD} + \theta.\mathrm{BOR.PBD} + \delta.\mathrm{EDU} \qquad (4.12)$$

and the corrected χ^2 values for terms in this model are given in table 4.6, from which it may be seen that PBA and especially LNPBI are much more important than any other terms in the model. Figure 4.6 and appendix table 4.E5 contain the parameter estimates for model (4.12).

Figure 4.4. Estimates of PBI effects in model (4.8) at 10, 20, 40 and 80 months survival-time, for birth-orders 2 to 6 separately. Vertical lines of length one standard error are drawn to each side of each parameter estimate. The scale for all countries is the same and is indicated for Bangladesh. Normalisation is given in table 4.2.

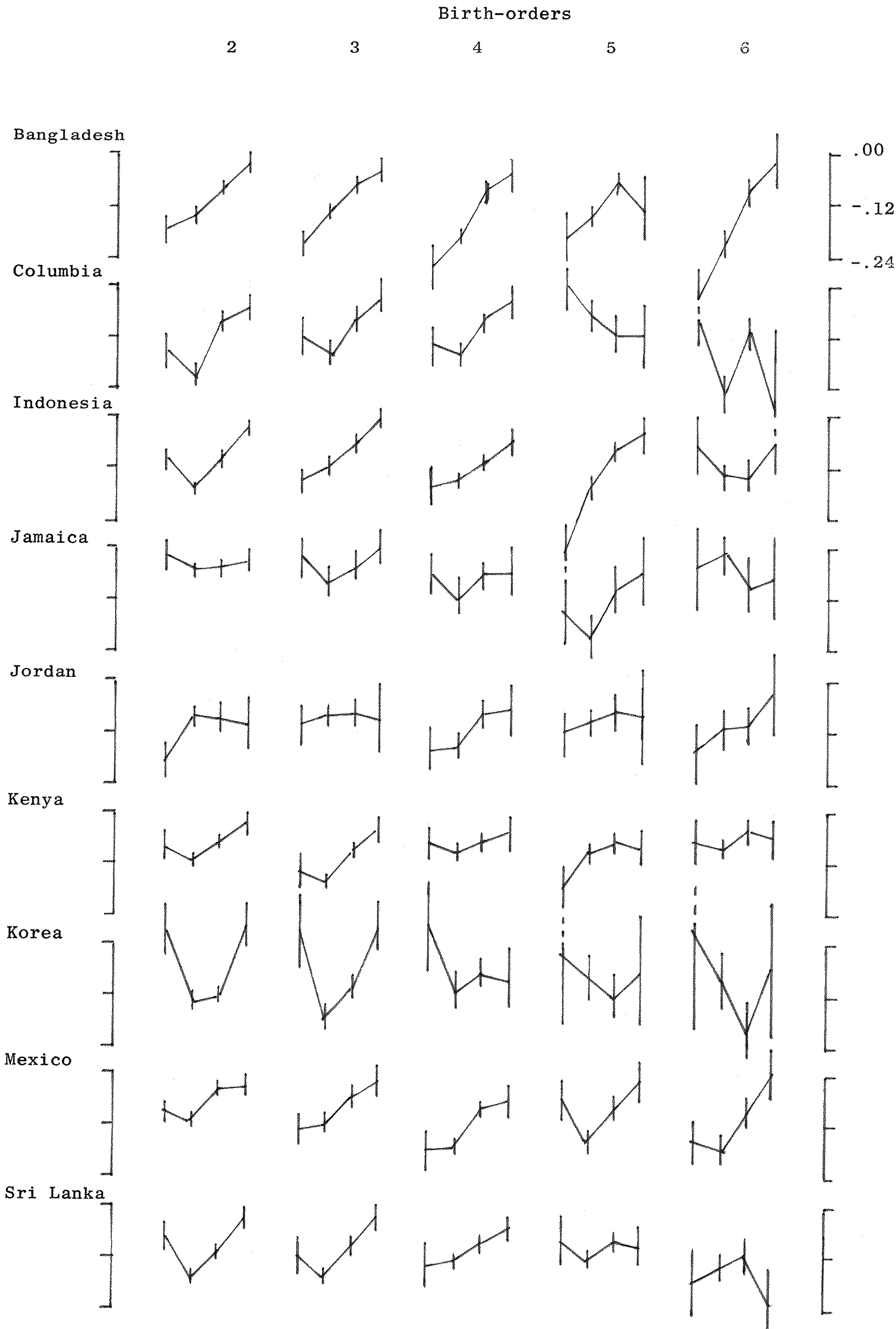


Figure 4.4: PBI effects, α(t), from model (4.8)

Figure 4.5. Estimates of BOR effects in model (4.9) at 10, 20, 40 and 80 months survival-time, for birth-orders 2 to 6 simultaneously. Vertical lines of length one standard error are drawn to each side of each parameter estimate. The scale is the same for all countries and is indicated for Bangladesh. Normalisation is given in table 4.2.

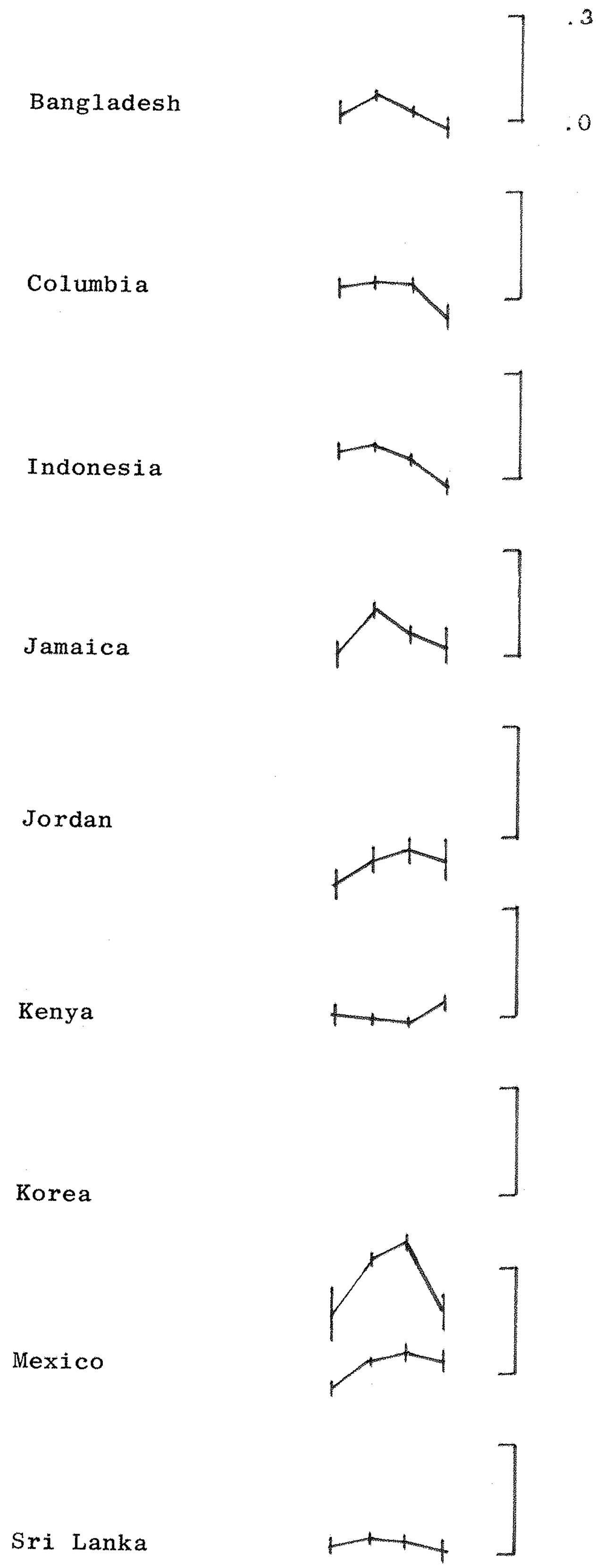


Figure 4.5: BOR effects, $\eta(t)$, from model (4.9)

Table 4.4 Differences in χ^2 between models (4.6), (4.9) and (4.11) fitted to birth-orders 2 to 6 simultaneously[a].

	Model (4.6)-Model(4.9)	Model(4.6)- Model(4.11)
Bangladesh	+33.	+269.[c]
Columbia	-46.	+42.[c]
Indonesia	-99.	+232.[c]
Jamaica	-29.[b]	-70.[b,c]
Jordan	-66.	-28.
Kenya	-40.	+41.
Korea	-85.	-39.
Mexico	-98.	-42.
Sri Lanka	-125.	+39.

a. Birth-intervals commencing with a multiple birth omitted.

b. Adjusted due to one case for which date of first birth is unknown.

c. Adjusted due to birth-intervals where length of previous birth-interval $\leqslant$ 6 months.

Table 4.5 Successive χ^2 differences produced by cumulatively placing constraints on model (4.11)

	Constraints			
	$\eta(t) = \eta$	$\phi = 0$	$\theta = 0$	$\eta = 0$
Bangladesh	18.	1.	51.	6.
Columbia	11.	0.	4.	6.
Indonesia	17.	10.	100.	17.
Jamaica	17.	2.	20.	17.
Jordan	17.	5.	23.	3.
Kenya	5.	5.	24.	7.
Korea	63.	32.	305.	11.
Mexico	27.	0.	20.	8.
Sri Lanka	5.	2.	52.	11.
d.f.	3	1	1	1

Table 4.6 χ^2 contributions of terms in model (4.12) corrected for remaining terms.

	LNPBI	PBA	BOR.PBD	BOR.PBD and PBD	BOR.PBD and BOR	EDU
Bangladesh	632.	277.	51.	77.	57.	2.
Columbia	322.	86.	4.	243.	10.	122.
Indonesia	712.	301.	100.	226.	117.	5.
Jamaica	61.	76.	20.	78.	37.	17.
Jordan	171.	124.	23.	40.	26.	149.
Kenya	322.	56.	24.	24.	31.	7.
Korea	222.	438.	305.	318.	316.	314.
Mexico	427.	345.	20.	169.	63.	23.
Sri Lanka	473.	476	52.	36.	28.	212.
d.f.	4	4	1	2	2	1

Figure 4.6. Parameter estimates for model (4.12). For $\mu(t)$, $\alpha(t)$ and $\beta(t)$, estimates at $t = 10$, 20, 40 and 80 months are connected with straight lines. Vertical lines of length one standard error are drawn to each side of each parameter estimate. Normalisation is given in table 4.2. The scale for each effect is the same for all countries and is indicated for Bangladesh.

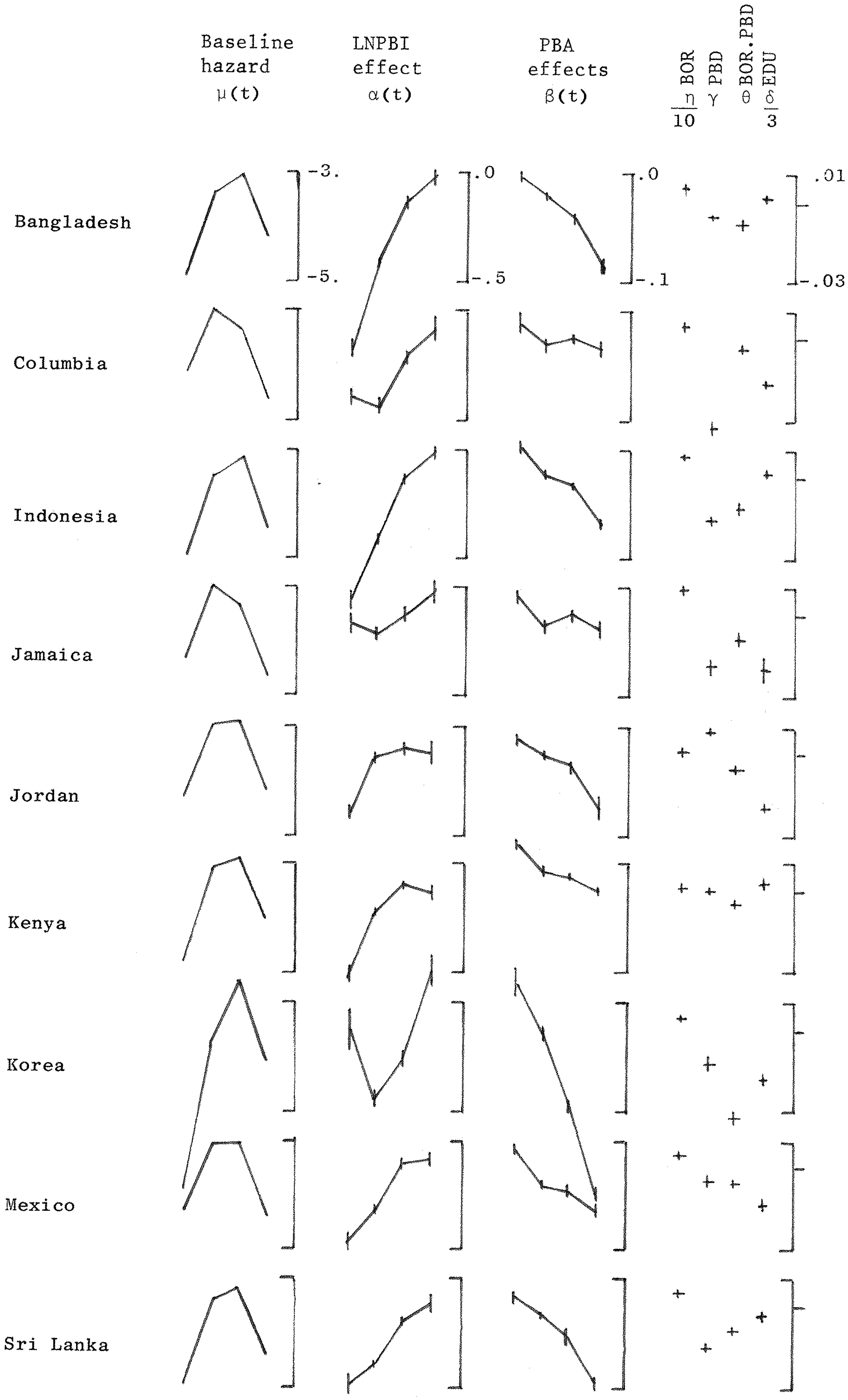


Figure 4.6: Parameter estimates from model (4.12)

4.3 Discussion

4.3.1 Adequacy of the model.

To some extent the motivation to discover a universal model has led to a rather forced application of model to data, notably in the case of Korea which resists most attempts at model simplification (see tables 4.3 to 4.5). Considering however the geographic, ethnic and socio-economic diversity between the nine countries involved in the analysis, the degree to which they support the same model is encouraging; moreover, the similarity between all countries except Korea in survival-time profiles of parameter estimates lends further support to the notion of a common model.

So far the analysis has not allowed for interactions other than with survival-time or birth-order. Interactions with current time-period, or PBD, might be anticipated however, reflecting changes in family building patterns brought about by various aspects of modernisation. To indicate the presence of such interactions, model (4.12) without its PBD terms was fitted to that part of each birth-interval for each woman which falls within the five years before interview. The resulting parameter estimates were virtually no different than those based on the full data, apart from some slight instability in the LNPBI effects which may be seen by comparing panels (a) and (b) of figure 4.9. This indicates that model (4.12) adequately incorporates time-period influences.

Some loss in precision results from using PBI in place of ABI in model (4.6), although this is regained by using LNPBI instead (table 4.4). This suggests that birth-intervals prior to the previous one contain useful information about the current birth-interval, but that ABI is not the best way of combining this information. Intuitively it might be expected that of the prior birth-intervals, the previous one would be most relevant to current fertility. To explore this possibility, correlation coefficients between birth-interval lengths have been constructed and are given in table 4.7. (To avoid selectivity bias, due to the fact that only the shortest of the more recently commenced birth-intervals are completed before interview, all birth-intervals commenced up to 100 months before interview have been omitted from this calculation, as have been all those of length greater than 100 months). Generally the greater the differences in birth-orders, the lower is the correlation between corresponding birth-interval lengths. All this suggests that perhaps the logarithm of a weighted geometric mean of prior birth-interval lengths would be an efficient way of summarising their information about current fertility (i.e. a weighted average of log birth-interval lengths, with higher weights assigned to the more recent prior birth-intervals to reflect their greater relevance to current fertility). However these refinements must await further research.

The correlations between birth-interval lengths may seem surprisingly small considering the large corrected χ^2 values in table 4.6 for LNPBI in model (4.12). Appendix 4C shows theoretically that the correlation between two adjacent birth-interval lengths should be approximately $\{-\alpha(1 \text{ year})/3\}$, and the empirical results in table 4.7 and figure 4.6 are consistent with this relationship.

4.3.2 Biasing Mechanisms

Age and the length of the previous birth-interval appear to be the key determinants of current fertility. Variables such as fecundity (the biological capacity to reproduce), coital frequency, and ability to breastfeed probably depend on age, and so it is possible that age

Table 4.7 Correlations between birth-interval lengths for birth-orders (BOR) 1 to 4[a].

Bangladesh

BOR	2	3	4
1	.20	.14	.04
2		.26	.13
3			.21

Columbia

BOR	2	3	4
1	.18	.13	.04
2		.17	.11
3			.12

Indonesia

BOR	2	3	4
1	.18	.13	.11
2		.18	.14
3			.19

Jamaica

BOR	2	3	4
1	.03	-.02	-.01
2		.01	.07
3			.04

Jordan

BOR	2	3	4
1	.07	.05	.07
2		.01	.02
3			.08

Kenya

BOR	2	3	4
1	.09	.08	.05
2		.15	.05
3			.04

Korea

BOR	2	3	4
1	.10	.12	.10
2		.08	.05
3			.10

Mexico

BOR	2	3	4
1	.09	.07	.08
2		.06	.09
3			.17

Sri Lanka

BOR	2	3	4
1	.12	.09	.07
2		.14	.14
3			.14

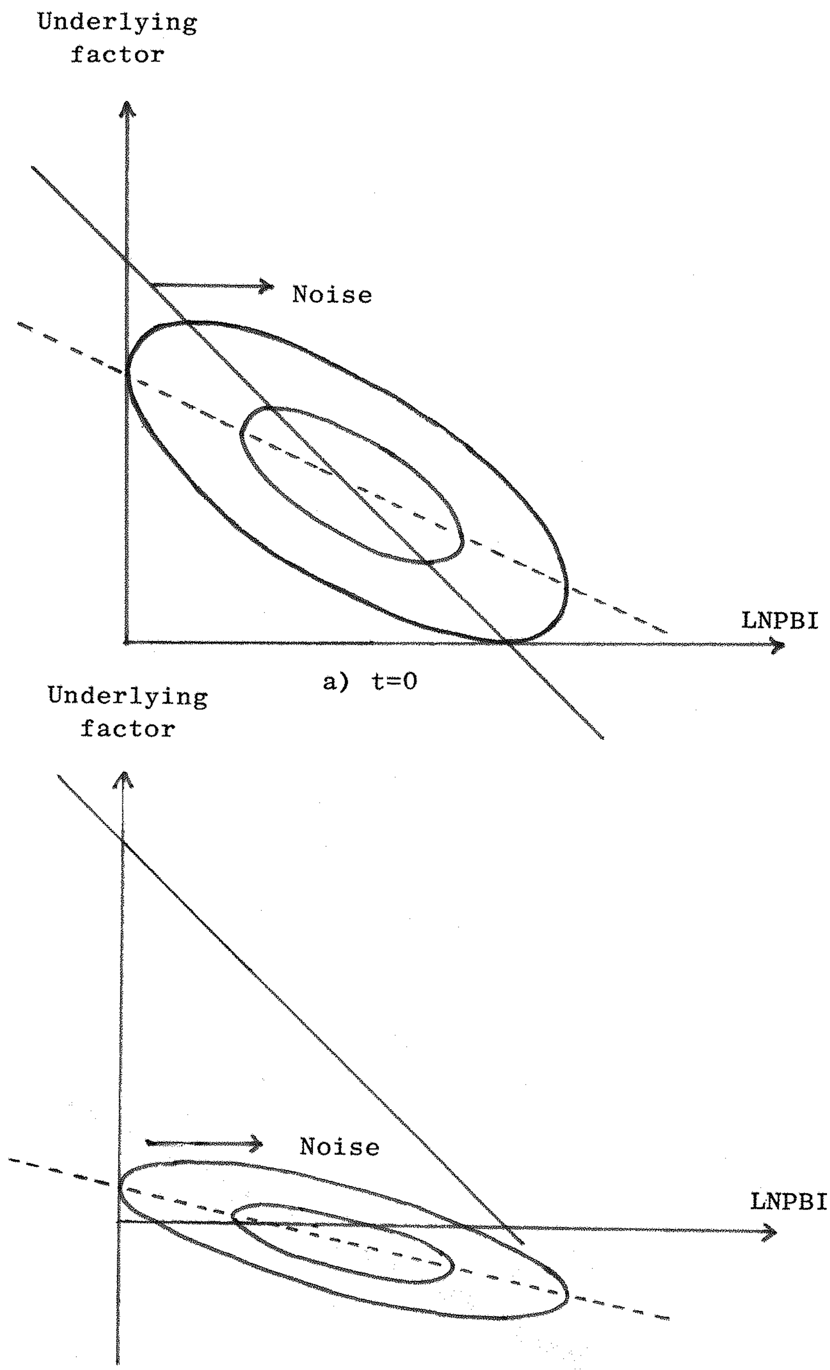
a. For women reaching 5^{th} birth whose birth-intervals for birth-orders 1 to 4 are all $\leqslant$ 100 months and whose 4^{th} birth was $\geqslant$ 100 months before interview (to avoid selectivity bias due to censoring).

effects reflect all of these factors. The role of the other variable, the length of the previous birth-interval, however, is quite different: it cannot be considered to be a cause of factors underlying current fertility, but it is a consequence of them.

Before attempting to associate possible causal factors with the observed survival-time profiles in the parameter estimates, it is important to gain understanding concerning the potential direction and size of various sources of bias.

The age-related factors listed above would all be expected ultimately to affect the length of the current birth-interval. Since age would not normally be much different for the previous birth-interval, it is reasonable to suppose that the length of the previous birth-interval reflects some of the current age-related factors, perhaps to a large extent. The length of the previous birth-interval may also reflect other factors of current fertility: consequently it is likely that age effects would tend to be diminished when controlling for the length of the previous birth-interval.

Now the length of the previous birth-interval undoubtedly also contains a component of 'noise' due to factors which have no bearing on current fertility. The consequent correlation between the length of the previous birth-interval and its noise component will produce biased parameter estimates: this is analogous to the classical 'errors-in-variables' situation (Kendall and Stuart, 1951). As a result of its noise component, the length of the previous birth-interval does not control precisely for its underlying factors, and this will tend to moderate its ameliorating action on the age effects, described above. The 'errors-in-variables' biases are demonstrated in panel (a) of figure 4.7. The solid diagonal line represents the expected value of the length of the previous birth-interval variable



b) t >> 0

Figure 4.7 Errors-in-variables and selectivity biases controlling for age, in LNPBI effects a) initially - b) late in birth-interval.

(LNPBI) given the value of one factor which underlies it, at the start of the birth-interval, controlling for age. The presence of noise, which is assumed at the start of the birth-interval to be distributed symmetrically and independently of the factor, produces the initial joint probability density of the factor and LNPBI, which is indicated on the figure by ellipses representing its contours; note that for each value of the factor, the contour lines are symmetrically placed about the solid diagonal line, reflecting the assumed symmetry and independence of the noise. Now for each value of LNPBI, the contour lines are symmetrically placed about the broken diagonal line, not the solid one. The expected value of the factor given LNPBI is therefore biased in comparison to the factor value which would be associated with LNPBI in the absence of noise. The presence of noise therefore causes factor differences to be underestimated (because the solid diagonal line is steeper than the broken line) and consequently the impact on current fertility of factors underlying LNPBI will be underestimated.

As survival-time increases, controlling for age and LNPBI, the level of the factor underlying LNPBI is reduced through the tendency for individuals with high factor scores to be removed from the birth-interval at the shorter survival-times. Simultaneously, and for the same reason, the variability in factor levels within these controls is reduced, and ultimately vanishes. However, no such selectivity mechanism operates on the noise component, since it is independent of current fertility. As survival-time increases, the noise component maintains its independence of the factor, and its variability is unaffected. Ultimately, therefore, LNPBI will reflect purely noise, and at intermediate survival-times the joint probability density of LNPBI and the factor will take the form indicated in panel (b) of figure 4.7. The solid diagonal line and the broken line represent the same relationships as in panel (a), for a survival-time greater than zero. The broken line is now even less steep than before, and consequently the impact on current fertility of factors underlying LNPBI will be even more severely underestimated. Thus, as survival-time increases, the baseline hazard should become increasingly negative, and the effects of LNPBI should disappear. This is an agreement with the results in figure 4.6.

From figure 4.7, the size of biases are determined by the relative size of the variance in the noise component in comparison to that of the underlying factor. The low correlation between birth-interval lengths (table 4.7) suggests that the noise component is relatively large, and this indicates the LNPBI severely underestimates the impact of factors which it reflects. By using information from several prior birth-intervals when measuring previous fertility, the noise variance is reduced and the effects of previous fertility are correspondingly increased, providing less biased estimates of the impact of factors underlying previous fertility. This is in agreement with the results of figure 4.1(b).

The large noise component of LNPBI means that LNPBI exerts little control over age-related factors initially, and this control becomes increasingly ineffective as the noise content of LNPBI increases with survival time. However, the trend of <u>increasing</u> PBA effects with survival-time cannot be attributed to the increasing inability of LNPBI to control for age-related factors, since the age-only model:

$$\ell n\ \lambda(t) = \mu(t) + \beta(t).\mathrm{PBA} \qquad (4.13)$$

produces age-effects which are similar to those obtained from model (4.12), as may be seen from figure 4.8. Now the rapidity with which

Figure 4.8. Effects of PBA from a) model (4.12) and b) model (4.13). Parameter estimates for $\beta(t)$ at $t = 10$, 20, 40 and 80 months are connected with straight lines. Vertical lines of length one standard error are drawn to each side of each parameter estimate. The scale is indicated in the upper right corner of the figure.

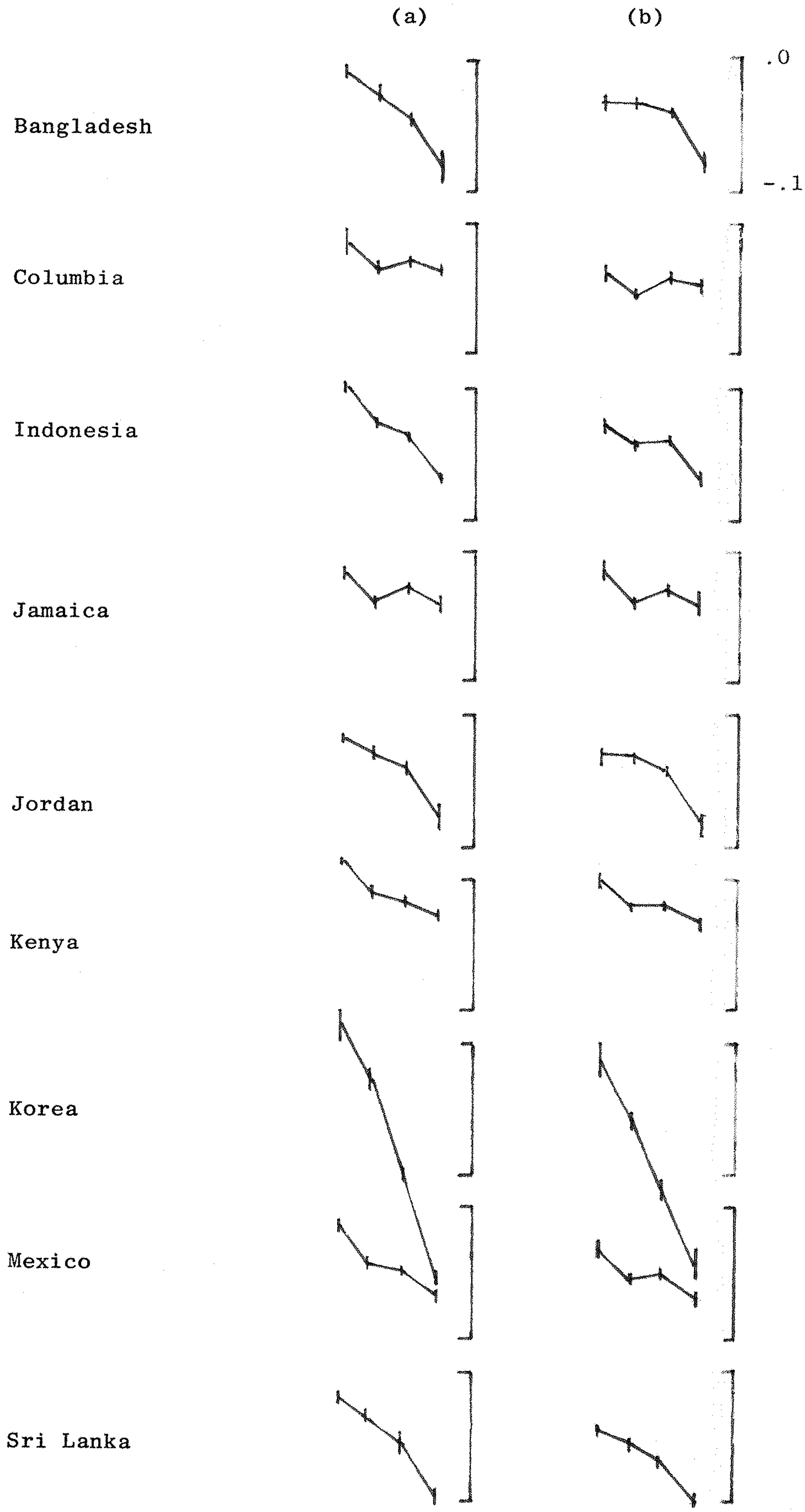


Figure 4.8: PBA effects, $\beta(t)$, from models a) (4.12) and b) (4.13)

the age-effects increase with survival-time seems to rule out a purely behavioural explanation; this could, however, be explained by age-related infecundity. Infecund individuals can only be removed from the birth-interval by interview, and so they tend to accumulate at the higher survival-times. If infecundity following the previous birth is related to age, then initially, and increasingly with survival-time, the age-effects will be dominated by the proportions fecund at each age, and all other effects will tend to disappear. This provides an additional explanation for the diminishing effects of LNPBI with survival-time. The anomalous behaviour of the Korean data could be explained by high levels of infecundity.

It is possible that a major source of age-related infecundity is the menopause. To check this, individuals were artificially censored upon reaching 40 years, and model (4.12) was refitted; but no change in the parameter estimates was produced. Menopause, therefore, does not contribute substantially to selection biases due to age-related infecundity; it seems likely that infecundity occurs at all ages, although more often at the higher ages. It is possible that infecundity occurs as a side-effect of the physiological processes of pregnancy, childbirth and breastfeeding.

It is unlikely that the age effects substantially reflect age related factors other than infecundity, since initially these are partly controlled for by LNPBI and as survival-time increases the age effects are increasingly dominated by infecundity.

All of the above assertions concerning biasing mechanisms are supported theoretically in appendix 4D. In particular, appendix 4D shows that the impact of factors underlying LNPBI could be about ten times those of LNPBI itself.

4.3.3 Breastfeeding and Contraception.

To gain further insight into the behavioural factors underlying the age and the length of the previous birth-interval variables, the impact of breastfeeding and contraception on the hazard function may be examined.

Breastfeeding and contraception data were collected only for the last closed birth-interval and the following open (censored) birth-interval, for each woman interviewed. Serious selectivity problems, additional to those described above, can result from the use of data for the last two birth-intervals (Hobcraft and Rodriguez, 1980) because the closed birth-intervals of the more fertile women (whose last two birth-intervals tend to be short) are under-represented. To control for this, only the experience within the five years prior to interview was used when analysing data for the last two birth-intervals: this avoids these selectivity biases to a large extent because the last two birth-intervals together generally exceed five years, and consequently selection is approximately solely on the basis of time-period rather than fertility. To restrict the period before interview still further would reduce further these selectivity biases, but stability in parameter estimates would be lost due to the contraction of the sample base.

Panel (c) of figure 4.9 displays the estimates of $\alpha(t)$ in model (4.12) estimated on the basis of the last two birth-intervals, within the five years before interview. Panel (b) contains the corresponding estimates for all experience within the five years before interview. Columbia and Sri Lanka show some signs of bias in $\alpha(t)$ due to the restriction to the last two birth-intervals. The remaining effects show no signs of such bias, for any of the countries analysed. Korea

Figure 4.9. Effects of LNPBI in model (4.12)

a) on full data

b) within the five years before interview

c) within the five years before interview, for the last two birth-intervals.

d) as for (c) when breastfeeding and contraception variables are included (model (4.14)).

Parameter estimates for $\alpha(t)$ at t = 10, 20, 40 and 80 months are connected with straight lines. Vertical lines of length one standard error are drawn to each side of each parameter estimate. Panels (c) and (d) are omitted for Korea through lack of convergence, and for Mexico because data on contraception in the last closed birth-interval was not collected. The scale is indicated in the upper right corner of the figure.

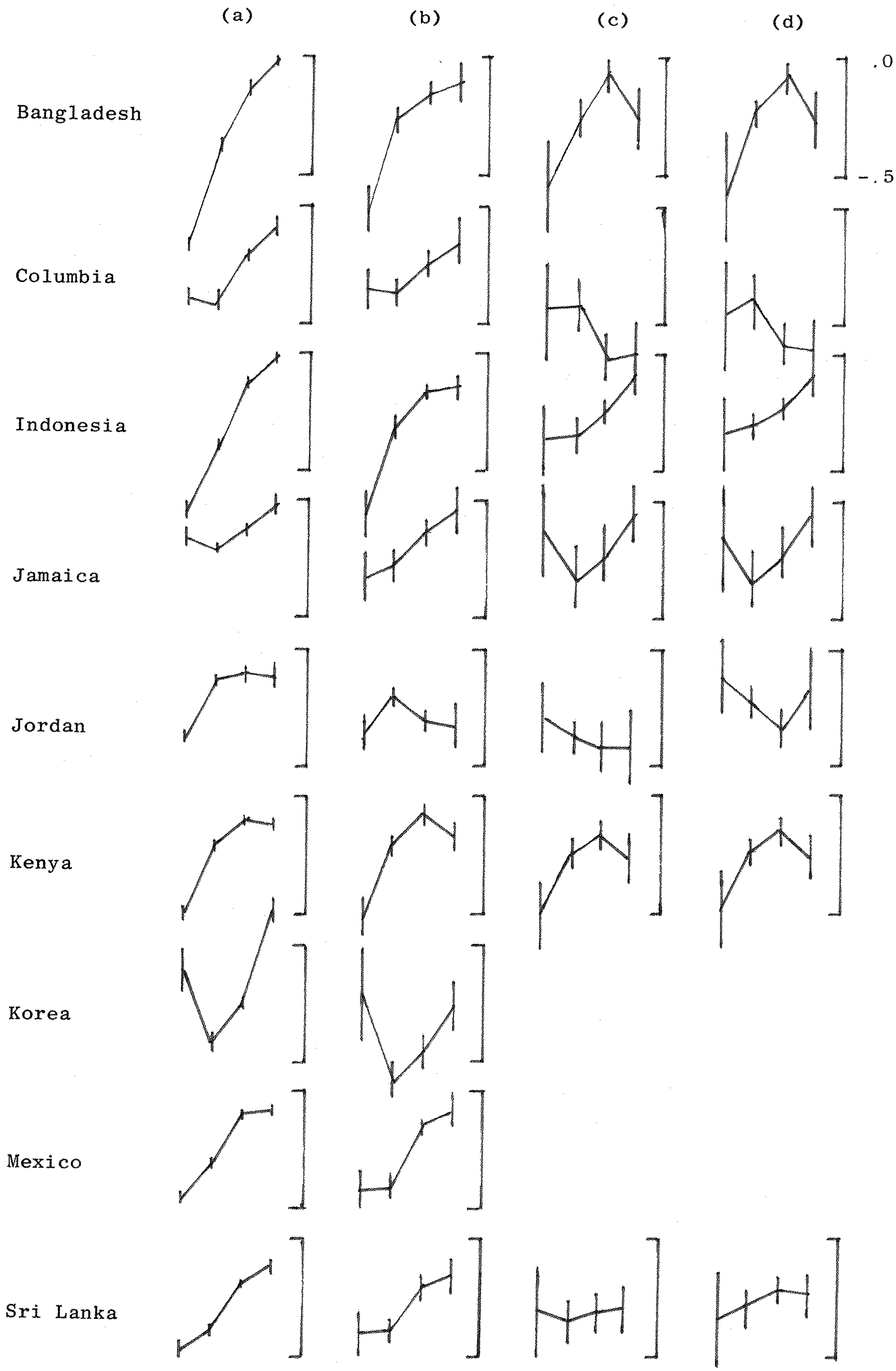


Figure 4.9: LNPBI effects, α(t) under various conditions

and Mexico are not analysed in this section owing to instability in parameter estimates when breastfeeding and contraception variables are included for the former, and to the fact that information on contraception in the last closed birth-interval is not available for the latter.

So far in this analysis all the variables utilised have been such that, for each woman, they do not vary within the birth-interval. Neither breastfeeding nor contraception can be considered to be constant within the birth-interval. This is not a problem provided that states of breastfeeding and contraception are known for each woman at each point in the birth-interval. This information is known for breastfeeding, although the distributions of reported durations of breastfeeding exhibit considerable clumping, indicating rather unreliable data. For contraception, however, it is only known whether it was ever used within the birth-interval (and if so, what type). It may not be unreasonable, however, to assume that use of contraception is uniform throughout the birth-interval, although this will probably lead to slight underestimates of the impact of contraception. However, this assumption can lead to yet another source of bias: for several countries in the analysis rapid changes in contraceptive practices are taking place, with the result that substantial numbers of women have changed their practice of contraception within the last two birth-intervals; for such women it is unreasonable to assume uniform use of contraception within the birth-interval, and the remedy is simply to omit these cases from the analysis.

In the interests of simplicity it was decided to ignore information on method of contraception. Thus, model (4.12) was augmented by full survival-time interactions with current contraceptive status, CEP (0 = not using; 1 = using), and current breastfeeding status,[1] BFD (0 = not breastfeeding; 1 = breastfeeding), but the dependence on date of previous birth was simultaneously dropped since this is effectively controlled by the restriction to the five years before interview, giving the model:

$$\ell n \ \lambda(t) = \mu(t) + \alpha(t).\text{LNPBI} + \beta(t).\text{PBA} + \eta.\text{BOR} + \delta.\text{EDU} + \rho(t).\text{CEP} + \tau(t).\text{BFD} \qquad (4.14)$$

The parameter estimates for contraceptive and breastfeeding status are given in figure 4.10. Contraception tends to have an initial negative impact on fertility, becoming positive after about 40 months survival-time (with the exception of Bangladesh). The negative impact of contraception on fertility is obviously due to the decreased risk of conception for contracepting women. The subsequent positive effect is possibly due to a tendency for non-contraception amongst infecund women, and is in the wrong direction to support the idea that contraceptive sterilisation is a primary source of infecundity.

Figure 4.10 does not show breastfeeding effects above 40 months survival-time due to their extreme instability, resulting from the small numbers still breastfeeding at these survival-times. The initial negative impact of breastfeeding is expected (see, for example, Jain and Bongaarts, 1980), and its diminishing effect with increasing survival-time is probably due partly to the decreasing intensity of breastfeeding with survival-time amongst women who are still breast-

---

1. It would have been better to use breastfeeding status nine months before the current survival-time. This oversight has probably led to slight underestimates of the impact of breastfeeding on fertility.

feeding, partly to measurement errors in the breastfeeding variable resulting in the errors-in-variance biases described above, and partly to the biasing effects of age-related infecundity, also described above.

With perhaps the exception of the education effects, the remaining effects in model (4.14) are virtually identical to those estimated from model (4.12). However, this does not imply that the LNPBI variable does not contain any variation originating from contraceptive or breastfeeding practices. To check this, corrected χ^2 -statistics for each term in model (4.14) are given in table 4.8, and these may be compared with those in table 4.9 for the model which omits the LNPBI term:

$$\ln \lambda(t) = \mu(t) + \beta(t).PBA + \eta.BOR + \delta.EDU + \rho(t).CEP + \tau(t).BFD \qquad (4.15)$$

Comparing these two tables reveals that age is the only variable which gains any explanatory power through the omission of the LNPBI variable, proving that the LNPBI effects do not reflect contraception, breastfeeding, birth-order or education to any real extent.

It may also been seen from table 4.8 that breastfeeding has a powerful impact on fertility, but that the impact of contraception is much more variable between countries.

4.3.4 The remaining effects.

To this point the birth-order, time-period and education effects have not been discussed. From figure 4.6 it is evident that fertility at the higher birth-orders is decreasing with time-period, indicating increasing use of fertility control, and also that education plays a part in reducing fertility in some countries, although further

Figure 4.10. Contraception effects, $\rho(t)$, and breastfeeding effects, $\tau(t)$ from model (4.14). Estimates of $\rho(t)$ at 10, 20, 40 and 80 months of survival-time, and of $\tau(t)$ at 10, 20 and 40 months of survival-time are connected with straight lines (estimates of τ(80 months) have very large standard errors, and are therefore omitted). Vertical lines of length one standard error are drawn to each side of each parameter estimate. The scale for each effect is the same for all countries and is indicated for Bangladesh. The normalisation of the variables is given in table 4.2.

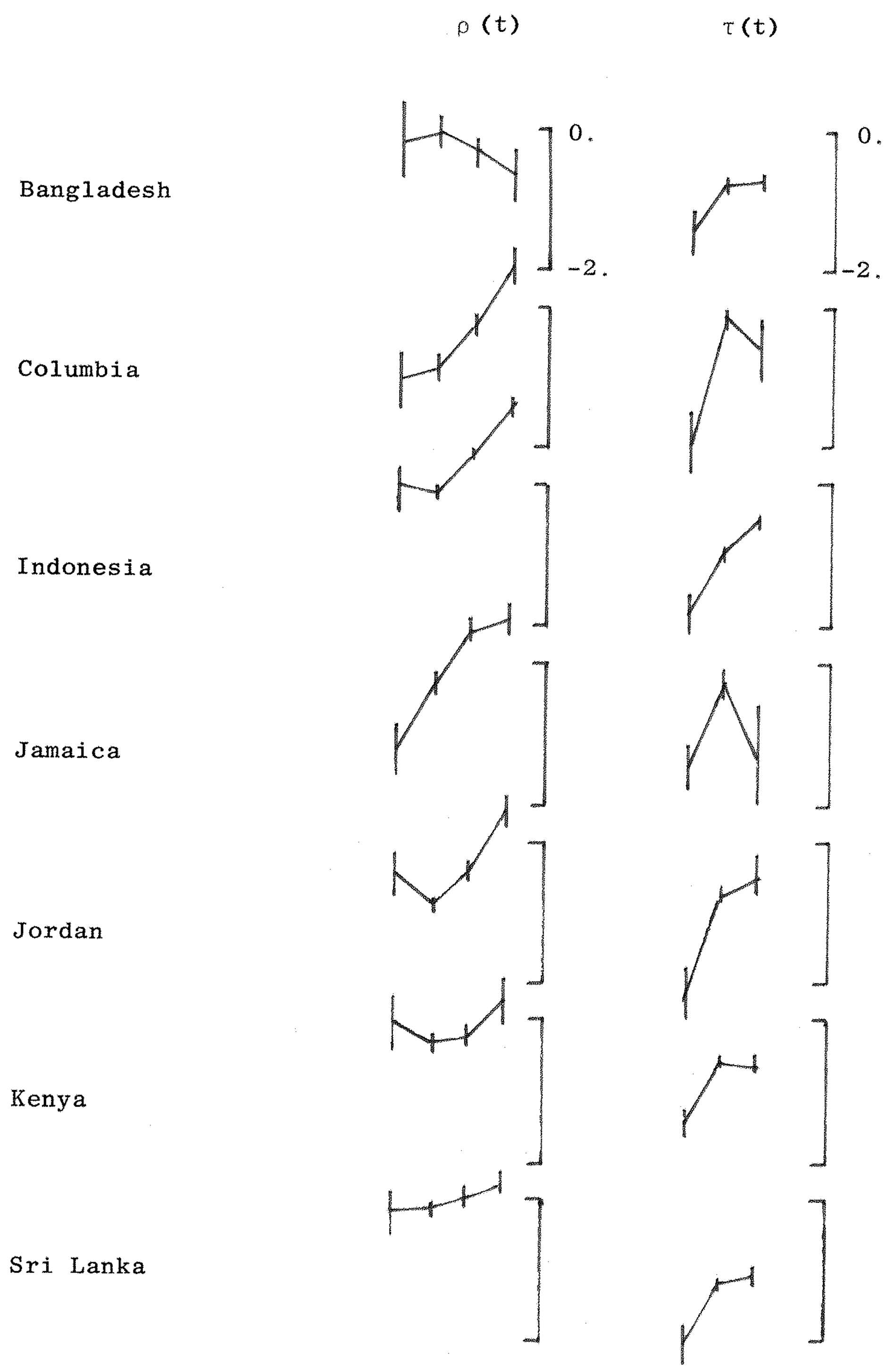


Figure 4.10: Contraception effects ρ(t), and breastfeeding effects τ(t), from model (4.14)

Table 4.8 χ^2 contributions of terms in model (4.14) corrected for remaining terms, based on the last two birth-intervals, within the five years before interview.

	LNPBI	PBA	BOR	EDU	CEP	BFD
Bangladesh	20	84	1	2	6	125
Columbia	53	35	2	23	48	33
Indonesia	35	131	2	3	85	113
Jamaica	10	26	1	6	20	21
Jordan	26	45	0	8	66	87
Kenya	32	34	1	15	10	101
Sri Lanka	25	136	0	15	4	231
d.f.	4	4	1	1	4	4

Table 4.9 χ^2 contributions of terms in model (4.15) corrected for remaining terms, based on the last two birth-intervals, within the five years before interview.

	PBA	BOR	EDU	CEP	BFD
Bangladesh	127	0	1	5	131
Columbia	87	0	24	54	33
Indonesia	199	8	1	83	123
Jamaica	51	2	4	19	21
Jordan	71	0	6	80	99
Kenya	59	0	16	11	103
Sri Lanka	217	0	11	4	246
d.f.	4	1	1	4	4

experimentation shows this to be partly through its association with contraception. However it is evident from table 4.6 that these terms are less important than age and the length of the previous birth-interval.

Figure 4.6 also shows substantial differences in baseline hazard functions between countries. These persist when controlling for breastfeeding and constraception, as may be seen from figure 4.11. The present research can throw very little light on the causes of these differentials, except to rule out those variables which have been investigated above.

4.4 Conclusions

A woman's age and her previous level of fertility are the components of her birth-history which are most strongly associated with her current fertility. Age acts principally to control for infecundity following the previous birth, and previous fertility simply predicts current fertility amongst those individuals who are still fecund. Survival-time trends in the effects of age and previous fertility are the results of selection associated with 'errors-in-variables' and infecundity biases.

Factors underlying the effects of previous fertility are unknown, but could include sub-fecundity, coital frequency and proneness to miscarry; they do not include breastfeeding or contraception. The impact on current fertility of factors underlying previous fertility are underestimated because of 'errors-in-variables' bias, but could be about ten times the impact of the length of the previous birth-interval. Using several prior birth-intervals to measure previous fertility substantially reduces this bias. The age effects do not reflect infecundity due to the menopause.

Contraception and especially breastfeeding have a negative impact on fertility at the shorter survival-times. Later in the birth-interval breastfeeding is less effective partly through less intensive usage amongst those still breastfeeding, and contraception effects become positive, possibly through the tendency for non-contraception amongst fecund women. Fertility at the higher birth-orders is decreasing with time-period indicating increasing usage of fertility control. Education also has a negative effect on fertility partly through its association with contraception. Substantial differences in the baseline hazard persist despite these controls, as illustrated in figure 4.11; the present analysis provides little insight into the causes of these differentials. Figure 4.12 illustrates the effects on the hazard function of the more important variables.

The aim of this research was to develop a model of fertility applicable to a wide variety of countries, using birth-history variables. This has been achieved, and the most important result is the strong suggestion of unknown factors associated with previous fertility which have a profound effect upon current fertility. This result is not entirely new: Braun (1980) in his analysis of closed birth-intervals for three historical populations also found that the level of previous fertility, as measured by the average length of prior birth-intervals was the most, indeed the only, important determinant of current fertility. The present research extends the result of Braun (1980) through its application to modern developing populations; through its assessment of the roles of breastfeeding and contraception; through its suggestion of the magnitude of the true impact of the underlying factors; and through its demonstration of the importance of age when analysis is not confined to closed birth-intervals. The present results are also consistent with those of Bumpass et al (1978) who find that women who commence childbearing at a young age generally continue with a rapid pace of childbearing.

An important aspect of the birth-history model (4.12) is its lack of dependence on marriage variables (above the second birth-interval), cohort variables, duration of motherhood and quadratic age and time-period variables. The absence of duration effects and quadratic age and time-period effects is particularly interesting since Casterline and Hobcraft (1981) on the same data have demonstrated the existence of such effects. This apparent contradiction is probably mainly due to the omission of birth-order effects in the Casterline and Hobcraft (1981)

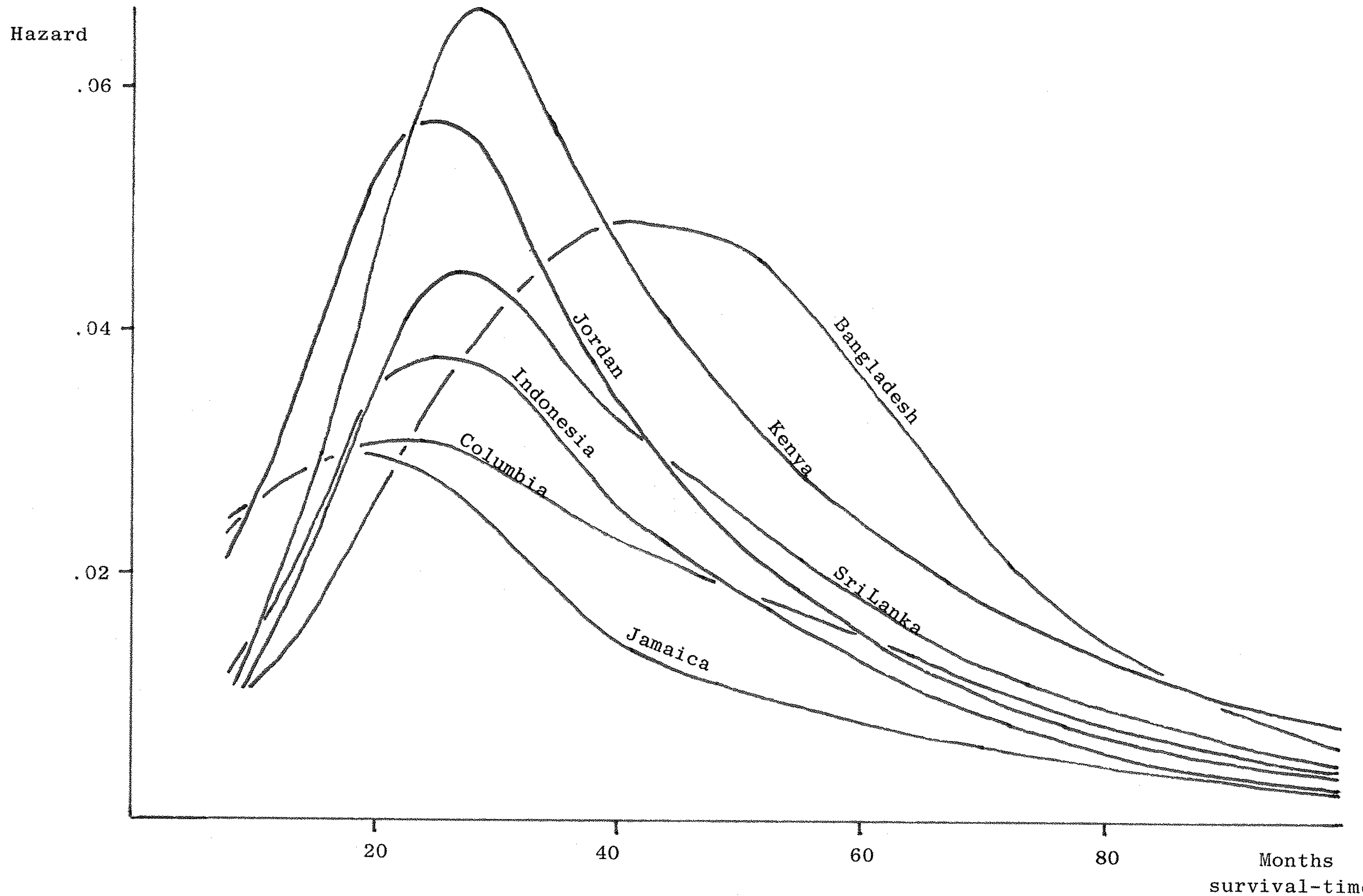


Figure 4.11 Baseline hazards from model(4.14) (Normalisation as in table 4.2).

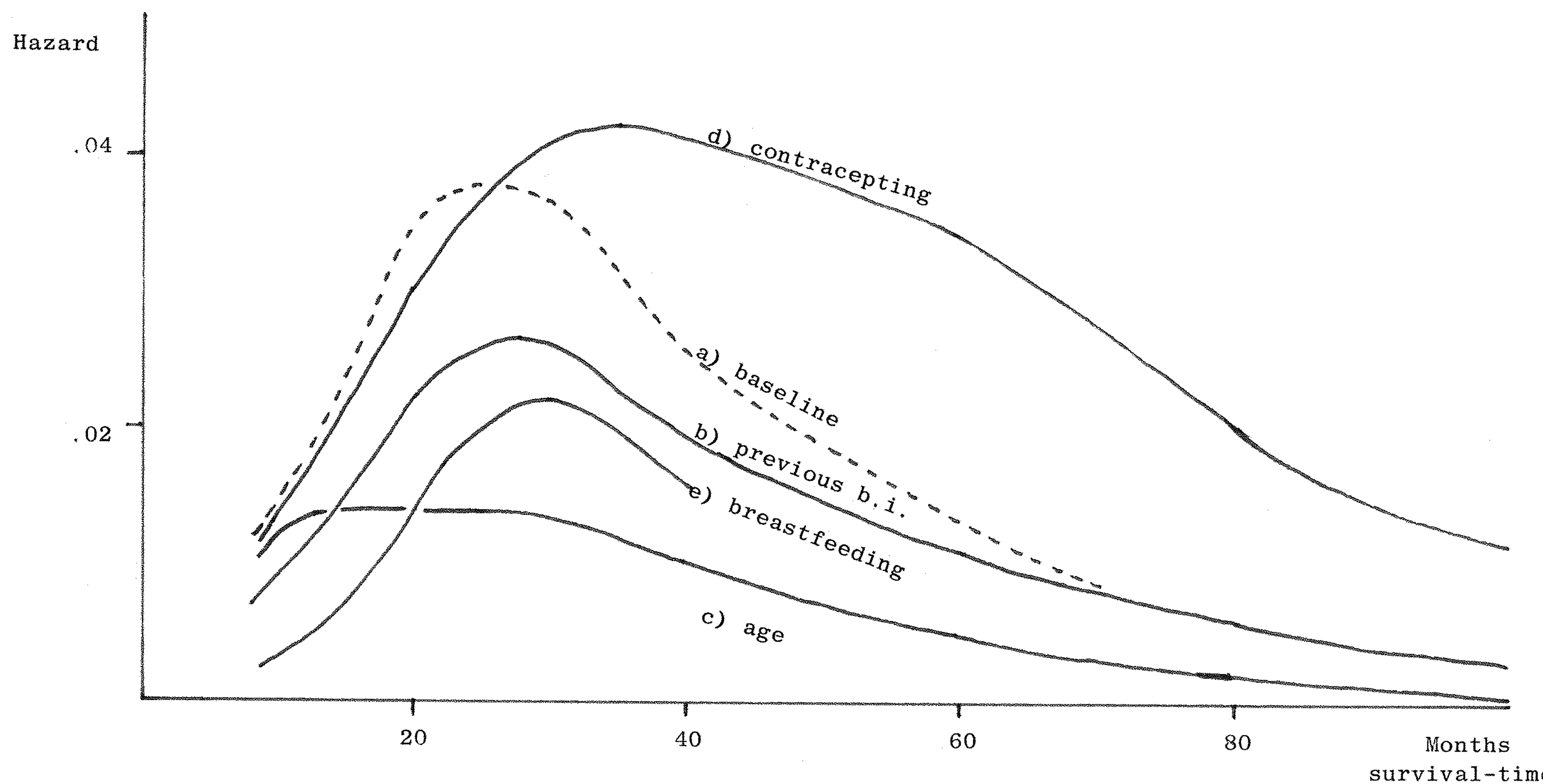


Figure 4.12 Hazards from Model (4.14) for Indonesia: a) the baseline hazard; b) when the length of the previous birth-interval is increased by 70 months; c) when age is increased by 15 years; d) when contracepting; e) when breastfeeding.

analysis (it will be recalled from section 4.2 above that birth-order and duration of motherhood effects are mainly the effects of previous fertility measured as the average of the lengths of prior birth-intervals).

The present research also contributes understanding about biasing mechanisms in the field of proportional hazards models in general.

The most important objective of further research should be to isolate the factors of current fertility which underlie previous fertility. In particular the impact of wasted pregnancies could be investigated with the present data and methods. It is also important to learn more about the causes of infecundity. However, the proportional hazards approach may not be the best framework within which to pursue these objectives, because of the influences of biasing mechanisms.

Another objective should be to apply the model to data from developed countries, (one would anticipate much greater birth-order effects); and to extend the model to first and second birth-intervals.

Finally, the model may serve as the basis of a demographic tool for assessing current levels of fertility and projecting future fertility. Further research could show how best to measure previous fertility, and how best to predict infecundity.

Appendix 4.A The approximate maximum likelihood solution

The probability that the i^{th} woman survives in her j^{th} birth-interval to survival-time t months is:

$$\exp\left\{ - \int_{o}^{t} \lambda_{ij}(s|\underline{\alpha})\,ds \right\} \tag{4.A1}$$

where $\lambda_{ij}(s|\underline{\alpha})ds$ is the hazard function times ds: that is, the probability that she has a birth in the interval (s, s+ds) given that she has already survived s months in the birth-interval; and where $\underline{\alpha}$ is a vector of parameters. The likelihood function may therefore be written:

$$L(\underline{\alpha}) = \prod_{i}\prod_{j} \lambda_{ij}^{\varepsilon_{ij}}(t_{ij}|\underline{\alpha}) \exp\left\{ -\int_{o}^{t_{ij}} \lambda_{ij}(s|\underline{\alpha})\,ds \right\} \tag{4.A2}$$

where t_{ij} is the survival-time at closure of the j^{th} birth-interval of the i^{th} woman and $\varepsilon_{ij}=1$ if closure is by a birth, and 0 if by interview. (In fact, dates in the survey data are recorded in century-month form, and so exact dates are taken to be at the mid-point of the century-month, except when a birth occurs in the same century-month as the interview, in which case the birth is taken to occur one-third of the way through the century month, to avoid bias.)

The proportional hazards model may be written (as in equation (4.1) above):

$$\ell n\ \lambda_{ij}(t|\underline{\alpha}) = \underline{\alpha}'\ \underline{X}_{ij}(t) \tag{4.A3}$$

where $\underline{X}_{ij}(t)$ is a set of covariate values for the j^{th} birth-interval of the i^{th} woman. Substituting (4.A3) into (4.A2) gives:

$$\ell n\ L(\underline{\alpha}) = \underline{\alpha}' \sum_{i}\sum_{j} \varepsilon_{ij}\ \underline{X}_{ij}(t_{ij}) - \sum_{i}\sum_{j} \int_{o}^{t_{ij}} \exp(\underline{\alpha}'\underline{X}_{ij}(s))\,ds \tag{4.A4}$$

Now let H be a partition of the space $\{i,j,s,\underline{X}\}$. Let $\delta_h=1$ if i,j,s and $\underline{X}$ are such that they correspond to the h^{th} subset of H. Let $\delta_h = 0$ otherwise. Thus δ_h is a function $\delta_h(i,j,s,\underline{X})$.

Denote

$$N_h = \sum_i \sum_j \int_o^{t_{ij}} \delta_h(i,j,s,\underline{X}_{ij}(s))ds$$

$$\underline{\overline{X}}_h = \frac{1}{N_h} \sum_i \sum_j \int_o^{t_{ij}} \delta_h(i,j,s,\underline{X}_{ij}(s))\ \underline{X}_{ij}(s)ds$$

$$S_h^2 = \frac{1}{N_h} \sum_i \sum_j \int_o^{t_{ij}} \delta_h(i,j,s,\underline{X}_{ij}(s))(\underline{X}_{ij}(s)-\underline{\overline{X}}_h)(\underline{X}_{ij}(s)-\underline{\overline{X}}_h)'ds$$

and $$\underline{X}_B = \sum_i \sum_j \varepsilon_{ij}\ \underline{X}_{ij}(t_{ij})$$

then if the partition H is chosen so that, within each subset h of the partition, $\underline{X}_{ij}(s)$ is approximately equal to $\underline{\overline{X}}_h$, then from (4.A4)

$$\ell n\ L(\underline{\alpha}) \simeq \underline{\alpha}'\underline{X}_B - \sum_h N_h(1+\tfrac{1}{2}\underline{\alpha}'S_h^2\underline{\alpha})e^{\underline{\alpha}'\underline{\overline{X}}_h} \qquad (4.A5)$$

and hence:

$$\frac{\partial \ell nL}{\partial \underline{\alpha}} = \underline{X}_B - \sum_h N_h \left[(1+\tfrac{1}{2}\alpha'S_h\alpha)\ \underline{\overline{X}}_h + S_h^2\alpha) \right] e^{\underline{\alpha}'\underline{\overline{X}}_h} \qquad (4.A6)$$

and

$$\frac{\partial^2 \ell nL}{\partial\underline{\alpha}\partial\underline{\alpha}'} = -\sum_h N_h \left[(1+\tfrac{1}{2}\underline{\alpha}'S_h\alpha)\underline{\overline{X}}_h\underline{\overline{X}}_h' + S_h^2\alpha\overline{X}'_h + \underline{\overline{X}}_h\underline{\alpha}'S_h^2 + S_h^2 \right] e^{\underline{\alpha}'\underline{\overline{X}}_h} \qquad (4.A7)$$

Thus only the information: $\underline{X}_B$, $\{N_h, \underline{\overline{X}}_h, S_h^2\}$ need be retained from the data. Equations (4.A5) to (4.A7) may be used to obtain approximate maximum likelihood estimates of the $\underline{\alpha}$ using the Newton Raphson procedure.

Empirically it was found that the approximation (4.A5) is adequate if the partition H subdivides the survival-time axis as follows:
6-9 months; 9-15 months; 15-60 months; 60-90 months; 90+ months.
Subdivisions with respect to the covariates was found to be unnecessary.

Appendix 4.B Modelling with cubic splines

It is reasonable to suppose that the effects of variates vary smoothly with survival-time. If so, then the proportional hazards model may be written:

$$\ell n\ \lambda(t|\underline{X}) = \underline{\alpha}'(t)\underline{X} \tag{4.B1}$$

where $\lambda(t|\underline{X})$ is the hazard function at survival-time t, given a vector of covariates $\underline{X}$, and where $\underline{\alpha}(t)$ is a vector of hazard functions. Thus $\underline{\alpha}(t)$ represents the effect of $\underline{X}$ at time t. The smooth functions $\underline{\alpha}(t)$ must be specified in some way, and then estimated.

Each function $\alpha_i(t)$ may be specified as a cubic spline. A cubic spline is a smooth function which is defined on a set of n positions or 'knots' $t_1 \ldots t_n$ along an axis. Between each pair of adjacent knots the function is a cubic polynomial. Before the first knot and after the last knot the function is a straight line. The coefficients of these (n+1) polynomials are chosen so that, at each knot, the function is continuous and twice differentiable at each of the knots. Specifying the values of $\alpha_{ij} = \alpha_i(t_j)$ at each of the knots $t_1 \ldots t_n$ then uniquely determines the spline function $\alpha_i(t)$ for all t. Of course these values α_{ij}, j=1...n are not known in advance in the present circumstances: they must be estimated, and they consequently form a set of n model parameters. The number and positions of the knots may however be set in advance to obtain the desired flexibility in $\alpha_i(t)$.

Now, for each j, a spline function $s_j(t)$ may be defined on the knots $t_1 \ldots t_n$ so that at each knot $s_j(t)$ is zero except at t_j where it is 1. It is easily shown that:

$$\alpha_i(t) = \sum_{j=1}^{n} \alpha_{ij} s_j(t) \tag{4.B2}$$

since the linear combination of the set of splines $\{s_j(t)\}$ possesses all the continuity properties of the individual $s_j(t)$, and since at the j^{th} knot only $s_j(t)$ is non-zero. Formula (4.B2) is extremely valuable for computational purposes since, substituting (4.B2) into (4.B1) gives:

$$\ell n\ \lambda(t|\underline{X}) = \underline{\beta}'\ \underline{Y} \qquad (4.B3)$$

where $\underline{\beta}'$ is the vector of parameters $(\alpha_{11}, \alpha_{12} \ldots \alpha_{1n}, \alpha_{21}, \alpha_{22} \ldots \alpha_{2n} \ldots)$ and where $\underline{Y}$ is the vector of variates $(X_1\ s_1(t), X_1 s_2(t) \ldots X_1 s_n(t), X_2 s_1(t), X_2 s_2(t) \ldots X_2 s_n(t) \ldots)$. Note that, at each survival-time t, the variates $\underline{Y}$ may be calculated before model fitting since the functions $s_j(t)$ do not depend on the unknown parameters $\underline{\beta}$. McNeil and Trussel (1977) explain how the coefficients of the $s_j(t)$ splines may be calculated. Note also that (4.B3) is now a linear model.

Empirically it was found that 4 knots were amply sufficient to accommodate the intricacies of the data, and that setting these at 10, 20, 40 and 80 months survival-time ensured that the hazard function tends to zero as $t \to 0$ or $t \to \infty$. Concentrating the knots at the shorter survival-times is sensible since this is where most of the exposure to risk of birth is situated.

Anderson and Senthilselvan (1980) use a cubic spline to model the baseline hazard in a proportional hazards model.

Appendix 4.C Relationship between the coefficient of LNPBI and the correlation between adjacent birth-interval lengths

Let $p(t_2|t_1)$ denote the instantaneous probability of a birth at survival-time t_2 conditional on t_1, the length of the previous birth interval. The coefficient $\alpha(t_2)$ of LNPBI ($\ell n(t_1$-6 months)) in model (4.12) is such that an increase in t_1 produces a decrease in the hazard $\lambda(t_2|t_1)$ for small t_2 and no change in $\lambda(t_2|t_1)$ for large t_2, and hence a decrease in $p(t_2|t_1)$ for small t_2 but an increase in $p(t_2|t_1)$ for large t_2. Hence for some value θ, roughly:

$$p(t_2|t_1) = p(t_2)\{1 + \theta(t_1-\mu_t)(t_2-\mu_t)\} \quad (4.C1)$$

where $p(t)$ is the unconditional instantaneous probability of a birth at t, and where μ_t is the unconditional expectation of t (it is assumed that t_1 and t_2 have the same unconditional distributions).

Now it follows from (4.C1) that the correlation between t_1 and t_2 is:

$$r_{t_1 t_2} = \theta\sigma_t^2 \quad (4.C2)$$

where σ_t^2 is the unconditional variance of t. It also follows from (4.C1) that:

$$\left[\frac{\partial \ell n\ p(t_2|t_1)}{\partial t_1}\right]_{t_1=\mu_t} = \theta(t_2-\mu_t) \quad (4.C3)$$

Now from model (4.12):

$$\left[\frac{\partial \ell n\ \lambda(t_2|t_1)}{\partial t_1}\right]_{t_1=\mu_t} = \frac{\alpha(t)}{\mu_t\text{-6 months}} \quad (4.C4)$$

and for small t_2 (say t_2 = 12 months)

$$\lambda(12 \text{ months} | t_1) \simeq p(12 \text{ months} | t_1) \tag{4.C5}$$

hence from (4.C3), (4.C4) and (4.C5):

$$\theta \simeq \frac{\alpha(12 \text{ months})}{\mu_t - 6 \text{ months}} \cdot \frac{1}{12 \text{ months} - \mu_t} \tag{4.C6}$$

and substituting (4.C6) into (4.C2) gives:

$$r_{t_1 t_2} \simeq \frac{\alpha(12 \text{ months})}{\mu_t - 6 \text{ months}} \cdot \frac{\sigma_t^2}{12 \text{ months} - \mu_t} \tag{4.C7}$$

Empirically $\mu_t \simeq 30$ months, $\sigma_t \simeq 12$ months. Substituting these values in (4.C7) gives:

$$r_{t_1 t_2} \simeq - \frac{\alpha(12 \text{ months})}{3} \tag{4.C8}$$

which is in agreement with table 4.7 and figure 4.6.

Appendix 4.D Biases

Suppose the hazard function $\lambda(t \mid \underline{X}, Y)$ at survival-time t depends only on observed factors $\underline{X}$ and unobserved factor Y, which for each individual do not alter within the birth interval, and suppose also, for ease of exposition, that this hazard function does not depend on t, so that:

$$\ell n\ \lambda(t \mid \underline{X}, Y) = \mu + \underline{\alpha}' \underline{X} + Y \qquad (4.D1)$$

Assume that, at t = 0, Y is distributed normally and independently of $\underline{X}$ with zero mean and variance σ_Y^2. Let T be the time to failure, then from (4.D1) it can easily be shown that at t = 0, the conditional expectation of T given X and Y is:

$$E_\circ(T \mid \underline{X}, Y) = \exp\left\{-(\mu + \underline{\alpha}' \underline{X} + Y)\right\} \qquad (4.D2)$$

and the conditional variance at t = 0 is:

$$V_\circ(T \mid \underline{X}, Y) = \exp\left\{-2(\mu + \underline{\alpha}' \underline{X} + Y)\right\} \qquad (4.D3)$$

Let Z = ℓn T and assume that at t = 0 the conditional distribution of Z given $\underline{X}$ and Y is approximately normal, then from equations (4.D2) and (4.D3), using the moment generating function of the normal distribution:

$$E_\circ (Z \mid \underline{X}, Y) = -\mu - \underline{\alpha}' \underline{X} - Y - \tfrac{1}{2} V_\circ (Z \mid \underline{X}, Y) \qquad (4.D4)$$

$$V_\circ (Z \mid \underline{X}, Y) = \ell n 2 \qquad (4.D5)$$

Hence from (4.D4) and (4.D5), Z may be expressed as:

$$Z = -\mu - \underline{\alpha}' \underline{X} - Y - \tfrac{1}{2} \sigma_\theta^2 + \theta \qquad (4.D6)$$

where, at t = 0 the noise term θ is distributed normally and independently of $\underline{X}$ and Y with zero mean and variance $\sigma_\theta^2 = \ell n$ Z. Now suppose that $\underline{X}$ and Y for each individual are approximately the same for the previous birth-interval as for the current birth-interval, and suppose that Z actually represents the length of the <u>previous</u> birth-interval. Then (4.D6) still holds approximately and Z contains information about the unmeasured factor Y.

Now the hazard at time t given $\underline{X}$ and Z is:

$$\lambda(t \mid \underline{X}, Z) = f(T = t \mid T \geq t, \underline{X}, Z)$$

where here, and below, f denotes a probability density,

$$= \int_{-\infty}^{\infty} f(T = t, Y \mid T \geq t, \underline{X}, Z)\, dY$$

using the addition law of probability,

$$= \int_{-\infty}^{\infty} f(T = t \mid T \geq t, \underline{X}, Y, Z)\; f(Y \mid T \geq t, \underline{X}, Z)\; dY$$

using the multiplication law of probability,

$$= \int_{-\infty}^{\infty} \lambda(t \mid \underline{X}, Y)\; f(Y \mid T \geq t, \underline{X}, Z)\, dY$$

assuming that the hazard does not depend on Z if $\underline{X}$ and Y are specified,

$$= \frac{\int_{-\infty}^{\infty} \lambda(t \mid \underline{X}, Y)\; f(T \geq t \mid \underline{X}, Y, Z)\; f(Y \mid \underline{X}, Z)\; dY}{\int_{-\infty}^{\infty} f(T \geq t \mid X, Y, Z)\; f(Y \mid \underline{X}, Z)\; dY}$$

Using a form of Bayes Theorem,

$$= \frac{\int_{-\infty}^{\infty} \lambda(t \mid \underline{X}, Y)\; f(T \geq t \mid \underline{X}, Y)\, f(Z \mid \underline{X}, Y)\; f(Y \mid \underline{X})\; dY}{\int_{-\infty}^{\infty} f(T \geq t \mid \underline{X}, Y)\; f(Z \mid \underline{X}, Y)\; f(Y \mid \underline{X})\, dY}$$

using again the assumption that the hazard does not depend on Z if X and Y are specified, and using Bayes Theorem again also,

$$= \frac{\int_{-\infty}^{\infty} \exp\left[\mu+\underline{\alpha}'\underline{X}+Y-te^{\mu+\underline{\alpha}'\underline{X}+Y} - \frac{1}{2\sigma_\theta^2}(Z+\mu+\underline{\alpha}'\underline{X}+Y+\frac{1}{2}\sigma_\theta^2)^2 + \frac{Y^2}{2\sigma_Y^2}\right] dY}{\int_{-\infty}^{\infty} \exp\left[-te^{\mu+\underline{\alpha}'\underline{X}+Y} - \frac{1}{2\sigma_\theta^2}(Z+\mu+\underline{\alpha}'\underline{X}+Y+\frac{1}{2}\sigma_\theta^2)^2 + \frac{Y^2}{2\sigma_Y^2}\right] dY}$$

using equations (4.D1) and (4.D6), elementary life-table theory and the distribution assumptions for Y and Θ; and rearranging gives:

$$\ell n\lambda(t\,|\underline{X},Z) = \frac{\mu+\underline{\alpha}'\underline{X}}{1+\frac{\sigma_Y^2}{\sigma_\theta^2}} - \frac{Z}{1+\frac{\sigma_\theta^2}{\sigma_Y^2}}$$

$$+ J\left[t\exp\left(\frac{\mu+\underline{\alpha}'\underline{X}}{1+\frac{\sigma_Y^2}{\sigma_\theta^2}} - \frac{Z}{1+\frac{\sigma_\theta^2}{\sigma_Y^2}}\right), \left(\frac{1}{\sigma_Y^2}+\frac{1}{\sigma_\theta^2}\right)^{-\frac{1}{2}}\right] \qquad (4.D7)$$

where

$$J(a,b) = \ell n\left\{\frac{\int_{-\infty}^{\infty}\exp(-ae^{bw+\frac{1}{2}b^2} - \frac{1}{2}w^2)\,dw}{\int_{-\infty}^{\infty}\exp(-ae^{bw-\frac{1}{2}b^2} - \frac{1}{2}w^2)\,dw}\right\} \qquad (4.D8)$$

It can be shown numerically, for moderate values of a and b, that

$$J(a,b) \simeq -ab^2$$

and hence, from (4.D7) and (4.D8):

$$\ell n \quad (t \mid \underline{X}, Z) = \frac{\mu + \underline{\alpha}'\underline{X}}{1 + \frac{\sigma_Y^2}{\sigma_\theta^2}} - \frac{Z}{1 + \frac{\sigma_\theta^2}{\sigma_Y^2}}$$

$$- t \exp\left(\frac{\mu + \underline{\alpha}'\underline{X}}{1 + \frac{\sigma_Y^2}{\sigma_\theta^2}} - \frac{Z}{1 + \frac{\sigma_\theta^2}{\sigma_Y^2}} \right) \cdot \frac{1}{\frac{1}{\sigma_Y^2} + \frac{1}{\sigma_\theta^2}} \tag{4.D9}$$

and linearising (4.D9) with respect to $\underline{X}$ and Z gives:

$$\ell n \ \lambda(t \mid \underline{X}, Z) = \left(\frac{\mu + \underline{\alpha}'\underline{X}}{1 + \frac{\sigma_Y^2}{\sigma_\theta^2}} - \frac{Z}{1 + \frac{\sigma_\theta^2}{\sigma_Y^2}} \right) \left(1 - \frac{t}{\frac{1}{\sigma_Y^2} + \frac{1}{\sigma_\theta^2}} \right) \tag{4.D10}$$

Thus, from (4.D10) it can be seen that if Y is a very important underlying factor, ($\sigma_\theta^2 << \sigma_Y^2$), then the $\underline{X}$ variables would produce only minor differentials in the hazard when controlling for Z, although an increasing downward bias would accrue to the estimated effects of the underlying factor with increasing survival time. When Y is an unimportant underlying factor, ($\sigma_\theta^2 >> \sigma_Y^2$) then the effects of the underlying factor would be badly underestimated, but the $\underline{X}$ variables would then produce differentials, which would diminish only slowly with survival-time.

Note that (4.D10) holds well only for moderate t, Z and σ_Y^2. However, similar sorts of trends would be produced even for somewhat extreme values of these elements.

The relative size of the noise component.

Let W denote the natural logarithm of the length of the birth-interval which immediately precedes the previous birth-interval, then, like (4.D6):

$$W = -\mu - \underline{\alpha}'\underline{X} - Y - \tfrac{1}{2}\sigma_\theta^2 + \theta_W \qquad (4.D11)$$

and assuming that the noise terms θ and θ_W are independent then it is easily shown, using (4.D6) and (4.D11), that the correlation between Z and W, given $\underline{X}$, is:

$$\rho_{ZW|\underline{X}} = \frac{1}{1 + \dfrac{\sigma_\theta^2}{\sigma_Y^2}} \qquad (4.D12)$$

Now empirically, approximately, from table 4.7, $\rho_{ZW|\underline{X}} \approx .1$, and so, from (4.D12):

$$\frac{\sigma_\theta^2}{\sigma_Y^2} \approx 9.0 \qquad (4.D13)$$

Substituting (4.D13) into (4.D10) gives, at t = 0:

$$\ell n\ \lambda(t \mid \underline{X}, Z) = .9(\mu + \underline{\alpha}'\underline{X}) - .1Z \qquad (4.D14)$$

This suggests that, at t = 0 the estimated $\underline{X}$ effects are a slight under-estimate of those which would result if Z was omitted, and the Z effect is about one tenth of the size of the effect of the underlying factor Y.

The effect of measuring previous fertility more accurately.

Now let $\overline{Z}$ denote the average of the natural logarithms of the lengths of n prior birth-intervals. Z in equation (4.D6) may be replaced by $\overline{Z}$ but the variance of the noise term would then be σ_θ^2/n. Consequently, when replacing Z in (4.D10) by $\overline{Z}$, σ_θ^2 should be replaced by σ_θ^2/n. Substituting (4.D13) in the result gives, approximately, for n less than about 6, at t = 0:

$$\ell n\ \lambda(t \mid \underline{X}, \overline{Z}) = (1 - .1n)(\mu + \underline{\alpha}'\underline{X} - .1nZ) \qquad (4.D15)$$

analogously to (4.D14). Thus it may be seen that increasing the number of birth-intervals in the measure of previous fertility, Z, has the effect of approximately proportionately increasing the estimated effect of that variable.

Selection due to infecundity.

Now suppose that, in addition to the mechanisms described above, there are individuals at t = 0 who are infecund; and suppose that the proportion infecund $q(X_1)$ depends only on X_1, where

$$\text{logit } q(X_1) = \beta + \gamma X_1 \tag{4.D16}$$

Now let λ denote the hazard when the infecund individuals are excluded, and let λ^* denote the hazard when they are included, The using (4.D1) and elementary life-table theory:

$$\lambda^* (t|\underline{X},Y) = \frac{\lambda(1 - q(X_1))e^{-\lambda t}}{q(X_1) + (1 - q(X_1))e^{-\lambda t}} \tag{4.D17}$$

where λ represents $\lambda(t|\underline{X},Y)$. Hence from (4.D17) using (4.D1) and (4.D16):

$$\begin{aligned} \ell n \ \lambda^* (t|\underline{X},Y) &= \mu + \underline{\alpha}'\underline{X} + Y - \ell n \left\{1+\exp(\beta+\gamma X_1+te^{\mu+\underline{\alpha}'\underline{X}+Y})\right\} \\ &\simeq \mu - \ell n\left\{1+\exp(\beta+te^{\mu})\right\} \\ &\quad + (\underline{\alpha}'\underline{X}+Y)\left\{1 - \frac{te^{\mu}}{1+\exp(-\beta-te^{\mu})}\right\} \\ &\quad - \frac{\gamma X_1}{1+\exp(-\beta-te^{\mu})} \end{aligned} \tag{4.D18}$$

using the Taylor series expansion in $\underline{X}$ and Y.

Thus from (4.D18), when t is small:

$$\ell n \ \lambda^*(t|\underline{X},Y) \simeq \mu - \ell n(1+e^{\beta}) + \underline{\alpha}'\underline{X}+Y - \frac{\gamma X_1}{1+e^{-\beta}} \tag{4.D19}$$

showing that initially the infecund women cause a negative bias in the baseline hazard and a bias proportional to $-\gamma$ in the effects of X_1.

Now suppose that initially only a very few individuals are infecund, so that β is large and negative and so that for some large t, say t_β,

$$t_\beta e^{\mu} = 1 + \exp(-\beta - t_\beta e^{\mu}) \tag{4.D20}$$

Then as t increases towards t_β, from (4.D18):

$$\ell n\ \lambda^*(t|\underline{X},Y) \to \mu - \frac{\gamma}{t_\beta e^{-\mu}} X_1 \tag{4.D21}$$

showing that, if initially only a few women are infecund, for moderately large t the effects of all variables except X_1 disappear, and the effect of X_1 is biased.

	Step 1		Step 2		Step 3	
Bangladesh	FBI	37.				
Columbia	FMA	24.	EDU (PBD,FMD,OBD)	6.	FMD,OBD (PBD)	4.
Indonesia	FBI	198	EDU	65.		
Jamaica	FMA	13.				
Jordan	FBI (FMD,OBD)	19.	PBD,FMD (OBD)	11.	EDU	13.
Kenya	EDU	22.	FBI	25.		
Korea	FMD (OBD)	246.	FBI,PBD (PBA)	35.	EDU	10.
Mexico	PBA	25.	FMD (OBD)	7.	EDU	8.
Sri Lanka	FBI	104.	PBA,FMA	24.		

Appendix Table 4.E1 The first three steps of a forwards selection amongst the linear main-effect terms listed in table 4.1for birth-order 1, with χ^2 values corresponding to the selected terms. Terms enclosed in parentheses have χ^2 values within 75% of that for the selected term, and are listed in order of decreasing χ^2. When more than one term is selected in a single step, this is due to logical relationships between the selected terms. Terms with $\chi^2 < 3.0$ are not shown. (Each main-effect term represents one degree of freedom).

	Step 1		Step 2		Step 3	
Bangladesh	ABI (PBI)	215.	FBA,PBA	12.	PBI	6.
Columbia	ABI (PBD,PBA,PBI)	129.	FBD,PBD (FMD)	96.	EDU	41.
Indonesia	ABI	251.	FBD,PBD	62.	FBI,FMD (FBA,PBA,OBD)	16.
Jamaica	ABI	42.	FMD	9.	EDU (FBA,PBA,FMA,OBD)	4.
Jordan	ABI (PBI,EDU)	55.	EDU	47.	OBD	13.
Kenya	PBI (ABI)	109.	ABI (EDU,FBD)	14.	EDU	11.
Korea	PBA (EDU)	187.	EDU	117.	PBI (ABI,FBA)	27.
Mexico	PBA (ABI)	197.	EDU	63.	ABI,FBA	64.
Sri Lanka	PBA	281.	ABI,FBA (PBI)	65.	EDU (FBD,PBD,DBD,FMD)	31.

Appendix Table 4.E2.The first three steps of a forwards selection amongst the linear main-effect terms listed in table 4.1for birth-order 3, with χ^2 values corresponding to the selected terms. Terms enclosed in parentheses have χ^2 values within 75% of that for the selected term, and are listed in order of decreasing χ^2. When more than one term is selected in a single step, this is due to logical relationships between the selected terms. Terms with $\chi^2 < 3.0$ are not shown. (Each main effect term represents one degree of freedom.)

	Step 1		Step 2		Step 3	
Bangladesh	ABI (PBA)	177.	FBD,PBD (PBI)	26.	PBI	20.
Columbia	PBD (PBA)	94.	PBA,OBD (FMD,ABI,FBD)	48.	EDU (FMA,FMD)	21.
Indonesia	ABI (PBA)	187.	FBD,PBD	52.	PBI	22.
Jamaica	ABI (PBI)	48.	FBD,PBD	22.	PBI	10.
Jordan	ABI	65.	EDU	35.		
Kenya	ABI (PBI)	46.	FBD,PBD	17.	PBI	12.
Korea	PBA (PBD)	160.	EDU (PBD,OBD)	81.	PBD,OBD (FMD)	54.
Mexico	ABI (PBA)	177.	EDU (FBA,PBA)	40.	FBA,PBA	25.
Sri Lanka	PBA (PBI)	233.	ABI,FBA (PBD,OBD)	58.	FBD,PBD,OBD	51.

Appendix Table 4.E3 The first three steps of a forwards selection amongst the linear main-effect terms listed in table 4.1 for birth-order 5, with χ^2 values corresponding to the selected terms. Terms enclosed in parentheses have χ^2 values within 75% of that for the selected term, and are listed in order of decreasing χ^2. When more than one term is selected in a single step, this is due to logical relationships between the selected terms. Terms with $\chi^2 < 3.0$ are not shown. (East main-effect term represents one degree of freedom).

Appendix Table 4.E4 χ^2 contributions of terms in model (4.6) corrected for remaining terms

	BOR	MOD	PBA	BOR.PBD	BOR.EDU
Bangladesh	74.	371.	97.	60.	2
Columbia	66.	298.	64.	50.	28.
Indonesia	96.	599.	179.	140.	91.
Jamaica	53.	105.	48.	50.	4.
Jordan	2.	214.	84.	15.	34.
Kenya	21.	316.	47.	7.	6.
Korea	151.	268.	363.	685.	166.
Mexico	32.	452.	274.	38.	54.
Sri Lanka	59.	574.	394.	82.	71.
d.f.	3	4	4	1	1

Appendix Table 4.E5 Parameter estimates from model(4.12)

	Bangladesh	Columbia	Indonesia	Jamaica	Jordan	Kenya	Korea	Mexico	Sri Lanka
$\mu(10)$	-4.91	-4.17	-4.99	-4.35	-4.37	-4.85	-6.45	-4.58	-4.96
$\mu(20)$	-3.45	-3.02	-3.49	-3.00	-3.00	-3.07	-3.76	-3.11	-3.40
$\mu(40)$	-3.06	-3.40	-3.18	-3.39	-2.91	-2.93	-2.69	-3.13	-3.18
$\mu(80)$	-4.23	-4.61	-4.48	-4.65	-4.18	-4.04	-4.19	-4.34	-3.36
$\alpha(10)$	-.81	-.40	-.70	-.17	-.39	-.52	-.12	-.49	-.49
$\alpha(20)$	-.41	-.44	-.41	-.23	-.14	-.24	-.44	-.32	-.38
$\alpha(40)$	-.13	-.21	-.13	-.13	-.09	-.10	-.26	-.11	-.19
$\alpha(80)$	-.02	-.09	-.09	-.02	-.12	-.13	.12	-.09	-.12
$\beta(10)$	-.002	-.014	.003	-.009	-.010	.019	.017	-.008	-.016
$\beta(20)$	-.020	-.032	-.023	-.035	-.025	-.007	-.027	-.040	-.035
$\beta(40)$	-.041	-.027	-.032	-.025	-.034	-.012	-.094	-.045	-.051
$\beta(80)$	-.085	-.037	-.066	-.039	-.075	-.234	-.175	-.062	-.094
η	.054	.039	.080	.094	.012	.011	.042	.045	.050
γ	-.006	-.033	-.015	-.018	-.009	.000	-.012	-.005	-.015
θ	-.008	-.003	-.011	-.009	-.006	-.005	-.032	-.005	-.008
δ	.006	-.053	.006	-.062	-.065	.008	-.060	-.045	-.012

Chapter 5 - Conclusions

This chapter discusses the aims, methods, results and conclusions of the research presented above in Chapters 1 to 4 under two headings: an overview of the contribution to age, period, cohort methodology; and suggestions for further research. Finally, section 5.3 closes this thesis with a few summary remarks.

5.1 Overview of contributions to age, period, cohort methodology.

Chapters 1 to 4 above represent a number of quite different avenues of exploration of age, period, cohort methodology: The first three of these chapters concentrate on highly aggregated demographic data from the fields of fertility, nuptiality, mortality and morbidity, and the fourth is concerned with individual level fertility data. The first three chapters are principally methodological, whereas the fourth is principally substantive in motivation. Chapters 1 and 3 discuss issues relating to the simultaneous incorporation of age, period and cohort into a model, and the models in chapters 2 and 4 do not explicitly include period and cohort dimensions simultaneously. Chapter 1 is primarily motivated by a need to <u>justify</u> simultaneous incorporation of age, period and cohort into a model; chapter 2 develops <u>applications</u> of a particular model; chapter 3 <u>explores models</u>; and chapter 4 <u>explores data</u> using these and related dimensions. Although chapters 1 to 4 do not systematically explore <u>all</u> avenues of potential research into age, period, cohort methodology, the variety of problems and approaches considered do facilitate a broad perspective on the subject. A brief revew of the contributions to age, period, cohort methodology contained in the above chapters now follows.

Purely as a means of data description, demographers often summarise variation across time-periods or cohorts using measures such as, in the context of fertility for example, the total fertility rate or completed family size. The opportunity afforded by the additive age, period, cohort model of simultaneously ascribing some variation to each of these three dimensions must have seemed to early age, period, cohort analysts a useful aid to description. In fact many of these early analyses contain interpretations which go no further than a descriptive account of the data. The recent accusations of statistical insupportability (Glenn, 1976) and illogicality (Goldstein, 1978 and 1979) aimed at these analyses would seem a serious blow to the apparently modest aspirations of those analysts. In chapter 1 above it is demonstrated that there is nothing illogical about describing the data in terms of separate components of variation ascribed to each dimension, and figure 1.1 demonstrates the type of variation which is described by estimated contrasts in the parameter estimates from the model. However, none of the analysts indicate that this is the type of variation being described.

It is also shown in chapter 1, that for purely descriptive purposes, the particular identification chosen for the parameters of the additive age, period, cohort model is immaterial. However, in most applications of the model an attempt is made to find some justifications for the chosen identification, and consequently it must be concluded either that such justification is misguided or that really the purpose of the analysis is not purely descriptive.

When using the additive age, period, cohort model for theoretical purposes, Goldstein's (1978, 1979) accusation of illogicality can be avoided provided that parameter estimates are interpreted in terms of non-interacting underlying factors which are each associated with at most one of age, period and cohort. When these assumptions are made, the effects of changes in the levels of factors are usually of central interest, but unfortunately cannot be estimated without reliable supplementary information on the underlying factors, owing to the identification problem. Various attempts to resolve the identification problem, many of which are unreliable or simply invalid, have been reported in the literature, and several of these are reviewed in chapter 1.

Thus it is clear that in many researches the use of the additive age, period, cohort model has been accompanied with confusion, much of which has been brought about by failure to distinguish between the descriptive and the theoretical purposes, which differ totally in respect of assumptions, identification and interpretation. In particular, the way of avoiding Goldstein's (1978, 1979) charges of illogicality differs between the two purposes. The main contribution of chapter 1 to age, period, cohort methodology is the drawing of this distinction.

The assumption of non-interaction (or additivity) between the underlying factors of the additive age, period, cohort model has led to Glenn's (1976) accusation of statistical insupportability. Glenn's (1976) argument essentially rests on the fact that any cohort factor effect or interaction can be re-expressed as a form of interaction between age, and period factors. The argument is valid, yet if a simple structure underlying the data can be found in terms of age, period and cohort factors, then it makes sense to tentatively interpret the cohort factors literally rather than in terms of some curiously constrained interaction between age and period factors for which there may be no clear theoretical explanation. Chapter 1 provides statistical tests of the additive model versus alternative formulations involving factors from only two dimensions interacting in a simple way. However in most circumstances there can be no guarantee that interactions between factors are not more complex, and consequently such tests may often be of limited value. Moreover, as noted in chapter 3, residual patterns in highly aggregated data tend to persist irrespective of the model, which immediately invalidates these strict statistical procedures. Nevertheless, the algebra developed for these tests does demonstrate how simple interactions are particularly or wholly confounded with age, period and cohort factor main effects.

The additivity assumption is put under further scrutiny in chapter 3 where it is shown, for a variety of data, that theoretical arguments do not lead to additive age, period and cohort factors effects. Simple cohort-experience models are theoretically much more satisfactory and often capable of fitting the data well. Moreover, they are not encumbered with interpretation-blocking identification problems. However Glenn's (1976) arguments still apply, and interpretations must necessarily be tentative. In fact interpretations from cohort-experience models should be held extremely tentatively, since usually there are good substantive reasons for expecting simple interactions (as expressed by the bimodel) between period and age factors to account for much of the variation in the data. In general it is difficult to choose between a cohort-experience model or the bimodel on grounds of either fit or interpretability. Cohort-experience models can make much stronger substantive assertions than the bimodel and this is both a strength and a weakness, since on the one hand a clear theoretical framework could prove a powerful tool in many areas including projection, but on the other hand these assertions could be wrong and consequently misleading.

Chapter 3 succeeds in demonstrating the difficulty in choosing between models, but some clear general results do emerge: that period factors tend to be more important than cohort factors; that there is little point in adding on period parameters to a cohort-oriented theoretical model age-structure, or changing some of the cohort parameters to period parameters in order to accommodate time-period influences, unless this is done with close and careful reference to the theoretical foundation of the original model; that (as noted above) additive age, period and cohort factors are not usually indicated by either theory or data; and that cohort-experience models can produce interesting interpretations but sometimes unstable parameter estimates.

Perhaps chapter 3 might be criticised for attempting to find general results applicable to a wide variety of aggregate level demographic data, when the only valid approach is to consider each data set independently of unrelated sets of data. However the intention of the chapter is not to present a formula for model derivation; indeed the results strongly suggest that pre-constructed models such as the additive age, period, cohort model should not automatically be applied wherever there is a suspicion of factors related to all three dimensions: hence the attraction of the cohort-experience models which are only loosely specified in general form. The main purpose of the chapter is to make clear what can realistically be expected of theoretical models of highly aggregated data, and to discuss some of the possible advantages and disadvantages of different approaches to modelling such data.

Two models in particular, both cohort-experience models, emerge from chapter 3 as being of particular interest: the modification of Hernes' (1972) nuptiality model (3.16) and the achieved fertility model (3.22). The nuptiality model (3.16) leads to some rather interesting interpretations. Specifically, male marriageability has been declining presumably because women no longer need to find a husband in order to be secure financially, owing to their improved employment and earnings potentials assisted by equal opportunities legislation. This has particularly affected the marriagability of older males, whose principle advantage in the marriage market in the past may have been financial. Contrastingly, female marriageability seems hardly to have been affected (apart from slight fluctuations due to the war), suggesting that economic and social conditions are not important ingredients of female marriagability. The pressure-to-marry component of the model also appears to be time-invariant, indicating that pressure-to-marry is simply determined by the anxiety of being 'left on the shelf', as Hernes' (1972) originally suggested.

The achieved fertility model (3.22) also produces some fascinating results. According to the parameter estimates from the model, up to (on average) the second child, each additional child achieved produces a halving of the fertility rate, but after (on average) the second child each additional child produces a ten-fold reduction in fertility. This suggests that couples do not look ahead to plan their fertility in accordance with a desired family size, but rather react spontaneously to the number of children they already have, drastically reducing fertility when (on average) two children have been achieved. One of the intriguing possibilities for this model is that turning points in fertility might be predicted on the basis of constant period factor effects. Although the interpretations from this model can at this stage be held only tentatively, the theoretical foundation of the model and the consequent parameter estimates are in close agreement with 'common sense', and this adds some degree of reassurance concerning at least the approximate validity of the model.

In many situations the analyst might wish to avoid the risk of making strong theoretical assumptions of uncertain validity, and a model such as the period-specific bimodel could provide a convenient alternative, being generally capable of succinctly capturing the data variation, at the same time making somewhat unspecific theoretical assertions. Chapter 2 shows how Gabriel's (1971) biplot may be used to display the parameter estimates from the bimodel. One of the advantages of plotting the data in this way is that trends across ages and across time-periods may be separately and explicitly represented, revealing clearly even very minor features which could easily be overlooked when inspecting the raw data by eye. A certain amount of familiarity with interpreting biplots is however required in order to make the most of this graphical technique, but interpretation is greatly assisted by constructing axes representing 'level' (e.g. the total fertility rate and 'structure' (e.g. the mean age of fertility) on the biplot. For a data table of dimensionality higher than two, the bimodel and the biplot can still be used to great effect, as is demonstrated in chapter 2 and in several places in chapter 3.

Chapter 2 also shows how the bimodel may be used as a basis for constructing simple linear regression models for any type of demographic schedule. By collecting together schedules of any given type from a variety of countries and time-periods, and applying the bimodel, two or three standard structures may be empirically determined, and these may then be used as the 'independent' variables in a regression model for the given type of schedule. The model may then be used to improve the quality of other data, requiring only two (or at most three) relevant items of information from a population to estimate its regression parameters and hence to complete its schedule. Each population used in the construction of the regression model, and also all other populations for which there is at least some relevant data, may be presented on the biplot, revealing possibly clusters of populations and simultaneously trends over time.

The utility of the bimodel in model generation is enhanced still further by the possibility of using cubic splines to graduate the extracted standard structures thereby permitting, say, quinquennial age-specific data to form the basis of a single-year-of-age specific regression model. This has been done in chapter 2 using quinquennial age-specific legitimate fertility rates from a variety of populations. The resulting regression model is shown to be similar in structure to the usual model for such data: that of Coale and Trussell (1974). This technique of model generation is a powerful demographic tool, and could find wide application especially in connection with deficient data and for projection.

The above discussion concerns aggregate level data. At the individual level differentials may be examined at much greater depth and interpretations need not be so speculative although in order to explore many dimensions simultaneously it is often convenient to model variation across interval-level variables using only linear or quadratic components, rather than with a whole set of dummy covariates as in the additive age, period, cohort model for example. Thus in chapter 4, only linear or quadratic terms in the many dimensions constructable from the birth-history data are considered. Those dimensions include many types of 'cohort' variables (e.g. birth-cohort, marriage-cohort and motherhood-cohort) and 'age' variables (e.g. age, duration of marriage, duration of motherhood, duration since previous birth) as well as 'age-at-entry' variabes (e.g. age-at-marriage, age-at-motherhood). Thus many sources of confounding of the age period, cohort type are present when using several of these variables simultaneously. By restricting attention to only the linear components of these variables, some variables become completely confounded with others. This seems at first sight a severe problem, but it could alternatively be considered an advantage since some variables are automatically controlled for when others are explicitly included in the model. This should cause no problems provided that interpretations are expressed appropriately, although of course it may be impossible to ascertain whether period or cohort factors are responsible for some aspects of the data.

The main purpose of chapter 4 however is not to systematically explore the use of age-period-cohort models in connection with individual level data in the same way as was done with aggregate level data in chapter 3, although the results of chapter 4 certainly feed back into age, period, cohort type methodology. The purpose of chapter 4 is to explore birth-history data from diverse developing countries with the hope of discovering patterns which are not country-specific, which might lead to fundamental hypotheses applicable to developing countries in general.

The analysis of chapter 4 demonstrates strong patterns in fertility common to all countries examined (although Korea conforms less well than the remaining countries). The variables which dominate a woman's fertility at a given age are: some measure of her previous pace of fertility (such as the length of her previous birth-interval, or the average length of her previous inter-birth intervals); and her age. These results may seem unsurprising, but it should be noted that these results also show what is <u>not</u> important in determining fertility for these populations, for example: the various cohort variables; duration-of-marriage and motherhood; age-at-marriage and motherhood. More surprisingly, even parity and time-period are of minor importance compared to age and the pace of previous fertility, although there are signs that parity is becoming more important. Education, which has been shown to be of major importance in many other analyses of fertility data, also appears as a minor variable here.

As with aggregate level data, variables constructed from dates cannot be considered to have a direct causative effect on fertility; rather they represent measures of underlying causal factors. It is reasonable to question the usefulness of such variables when more proximate variables may be available in the data. The use of date-related variables may be justified on the grounds that they reflect a number of underlying factors, many of which may not be available or even known. However chapter 4 well illustrates the biasing effect of measurement errors: the effect of including more prior birth-intervals when calculating the pace of a woman's previous fertility is to decrease the measurement errors on the factors underlying 'pace', and to consequently decrease the errors-in-variable bias and increase the magnitude of the effects. In fact statistical arguments suggest that if the factors underlying 'pace' were measured accurately, their effects could be about ten times the size of those estimated using the measure defined by the length of the previous birth-interval. Interestingly, when using the average length of prior birth-intervals as the 'pace' variable, precision increases in proportion to the number of prior birth-intervals included in the average, causing an apparent duration-of-motherhood effect which is constant across birth-order. It is shown in chapter 4 that selectivity mechanisms operating within the birth-interval produce errors-in-variables biases which increase with time-since-previous-birth. This result is not context-specific, and the principle is applicable in all fields in which life-table models are used.

The interpretation of the 'pace' variable is that it measures underlying factors such as biological fecundity and coital frequency. Surprisingly, contraceptive usage and length of breastfeeding evidently do not underlie the pace variable to a noticeable extent, although they both have substantial effects on fertility. In particular, some of the effects of education appear to operate through the contraception variable. The age variable has little effect shortly following the previous birth but, as time since the previous birth increases, the age variable becomes increasingly important. The only plausible explanation for this is that there is substantial heterogeneity in fecundity associated with age, possibly due to an age-related risk of impairment to fecundity occurring as a by-product of the physiological changes which take place during pregnancy, childbirth, and lactation. This is an important new suggestion which requires further work to properly substantiate.

Before this chapter is drawn to a close, the following section delineates areas for further research suggested by the foregoing results.

5.2 Suggestions for further research.

Both methodological and substantive areas for further research are suggested by the work presented above. On the methodological front it would be interesting to find further applications of the bimodel technique of model generation described in chapter 2. In particular this technique could prove useful in constructing model life-tables. The use of component-type models of mortality life-tables is however not new: Bourgeois-Pichat (1962), Hogan and McNeil (1979), Hobcraft (1979), and Zaba (1979) have used similar types of component life-table model. Nevertheless, the bimodel and its developments in chapter 2 could prove superior to all of these since: firstly, the Singular Value Decomposition used in estimating the bimodel is very efficient, and very widely available; secondly, cubic splines may be used to graduate extracted components; and thirdly, the biplot allows both age and country parameters to be simultaneously displayed. It would also be interesting to consider how the bimodel could be applied to migration data - an area of demography which has not been investigated at all in the present thesis.

Still on the methodological front, further work should be done to discover an optimum way of combining information from the birth-history to provide a measure of current fertility. The log of a weighted geometric mean of prior birth-interval lengths, suggested in chapter 4, is a useful starting point for such an investigation. For developed countries, and increasingly for developing countries, parity should be brought into account (this is also suggested by the achieved fertility model of chapter 3). As well as leading to better estimates of current fertility, such a model would be useful for projection.

The analysis of chapter 4 also opens up the way for investigation of biasing mechanisms in proportional hazards models. Although this area has implications well beyond the demographic context, proportional hazards models are becoming increasingly utilised by demographers, and consequently further research in this area should have important consequences for some demographic analyses.

On the substantive front, the cohort-experience models of chapter 3 are of considerable interest, in particular the modification of Hernes' (1972) nuptiality model (3.16) and the achieved fertility model (3.22). As has been stressed above, at this stage these models can be held only tentatively, and further research should therefore concentrate on discovering their applicability to other populations, and to finding support for them from suitable individual level data. The breadth and depth of available data on fertility should prove a useful testing ground for the achieved fertility model. This model does not easily accommodate changing age-patterns of entry into exposure to childbearing, and consequently it might be informative to apply and if necessary adapt the model for use with legitimate fertility data, or perhaps more usefully with age-at- and duration-of- motherhood data. The basic precept of the model is that fertility is a spontaneous response to parity, and consequently the most incisive test of the model would be in conjunction with parity specific data; this could also produce an interesting extension to the model. Unfortunately, neither duration-of-motherhood nor parity specific data of a suitable form are available in a long time-series for England and Wales, although it may be possible to find some way of utilising such data as is available. The O.P.C.S. Longitudinal Study could represent a valuable alternative source of data.

Chapter 4 raises some important questions about the presence of infecund or subfecund individuals in each parity group. The hypothesis that physiological changes wrought during pregnancy, childbirth and lactation sometimes result in lasting impairment to fecundity especially for older women, deserves further scrutiny, even though only a small proportion of women may be affected. The difficulty in persuing this line of research is that evidence for such a phenomenon is necessarily circumstantial, since involuntary infecundity would not normally be directly measurable. Nevertheless it would be interesting to estimate, albeit somewhat tentatively, differentials in subfecundity with respect to country, parity, age and other background variables.

It is also important to extend the work of chapter 4 to investigate the reasons for differences between countries in the baseline hazard for birth-intervals, which persist despite making several controls, and even though differentials on these controls are similar across countries.

5.3 Closing remarks.

In conclusion, it may be said that the work contained in this thesis demonstrates the usefulness and the limitations of models involving date-related variables such as age, period and cohort. The focus of much of the research was methodological, yet interesting substantive hypotheses have emerged from chapters 3 and 4. The potential of the bimodel and the cohort-experience type of model, and the comparative lack of utility of the additive age, period, cohort model, in conjunction with aggregate level demographic data, have been established. Errors-in-variables biases associated with date-related variables have been demonstrated in connection with individual level data, yet despite these biases such variables prove to be a powerful source of control. The work suggests several avenues for further development and exploration which could lead to important contributions to demographic methodology and understanding.

Anderson, J.A., and Senthilselvan, A.(1980). "Smooth estimates for the hazard function." J.R.S.S(B), 42, No. 3, 322 - 327.

Barrett, J.C. (1973). "Age, time and cohort factors in mortality from cancer of the cervix." J. Hyg., Camb., 71, 253 - 259.

Barrett, J. C. (1978a). "A method of mortality analysis: application to breast cancer." Revue Epidémiologie et Santé Publique 26, 419 - 425.

Barrett, J. C. (1978b). "The redundant factor method and bladder cancer mortality." J. Epidemiology and Community Health, 32, No. 4, 314 - 316.

Barrett, J. C. (1979). "Duration, period and cohort measures in marital fertility." Population et Famille, 47, 69 - 79.

Beard, R. E. (1963). In "Acturaial methods of mortality analysis." Proceedings of the Royal Society of London, Series B., 159, 56 - 64.

Becker, G. S. (1960). "An economic analysis of fertility" in Universities National Bureau Committee for Economic Research, "Demographic and economic change in developed countries", Princeton: Princeton University Press.

Becker, G. S. (1965) "A theory of the allocation of time", Economic Journal 75, 493 - 517.

Benzécri, J. P. (1976) "L'analyse des données. T.11: L'analyse des correspondances." Dunod, Paris.

Bloom, D. (1980). "Age patterns of women at first birth, essays in labour economics and demography." Unpublished Ph.D dissertation, Princeton University. Ann Arbor, Mich: University Microfilm.

Bock, R. D. (1975). "Multivariate statistical methods in behavioural research". New York: McGraw-Hill.

Bourgeois-Pichat, J. (1962). "Factor analysis and sex-age-specific death rates: a contribution to the study of the dimensions of mortality". United Nations Population Bulletin, 6, 147 - 201.

Bradu, D., and Gabriel, K. R. (1978) "The biplot as a diagnostic tool for models of two-way tables." Technometrics 20, No. 1, 47 - 68.

Brass, W. (1960). "The graduation of fertility distributions by polynomial functions". Pop. Stud. 14, No. 2, 148 - 162.

Brass, W. (1971). "On the scale of mortality". in Brass, W., (ed) "Biological aspects of demography", 69 - 110. London: Taylor and Francis.

Brass, W. (1974a). "Mortality models and their uses in demography". Transactions of the Faculty of Actuaries (Edinburgh) 33, 123 - 142.

Brass, W. (1974b). "Perspectives in population prediction: illustrated by the statistics of England and Wales" (with discussion). J.R.S.S.(A)., 137, 532 - 583.

Brass, W. (1980). "On the reliability of population forecasts". Paper presented at a meeting of the Social Statistics Section of the R.S.S.

Brass, W., Coale, A.J., Demeny, P., Heisel, D., Lorimer, F., Romanuik, A., and Van de Walle, E. (1968). "The demography of tropical Africa." Princeton: Princeton University Press.

Braun, H.I. (1980) "Regression-like analysis of birth interval sequences." Demography 17, No. 2, 207 - 223.

Braun, H. I., and Hoem, J. M. (1979) "Modelling cohabitational birth intervals in the current Danish population". Working paper No. 24, Laboratory of Actuarial Mathematics, University of Copenhagen.

Brouard, N. (1980). "A factor analysis and typology of the survey on fertility and contraception in Martinique, 1975". Unpulished paper, Institut National d'Etude Demographique.

Bumpass, L. L., Rindfuss, R. R., and Janosik, R. B. (1978). "Age and marital status at first birth and the pace of subsequent fertility." Demography 15, No. 1, 78 - 86.

Butz, W. P., and Ward, M. P. (1977), "The emergence of countercyclical U.S. fertility" Santa Monica, Cal., Rand Monograph R-1605-NIH.

Burtz, W. P. and Ward, M.P. (1979). "Is U.S. fertility temporarily low? Evidence on the timing component of period fertility rates". Population and Development Review 5, 663 - 668.

Case, R. A. M. (1956a). "Cohort analysis of mortality rates as an historical or narrative technique". Brit. J. Prev. and Soc. Med. 10, 159 - 171.

Case, R. A. M. (1956b). "Cohort analysis of cancer mortality in England and Wales, 1911 - 54 by site and sex". Brit. J. Prev. and Soc. Med. 10, 172 - 199.

Casterline, J. and Hobcraft, J. N. (1981). "Speed of reproduction: evidence from nine W.F.S. surveys". World Fertility Survey technical report No. 1629.

Coale, A. J. (1971) "Age patterns of marriage". Pop. Stud. 25, 193 - 214

Coale, A. J. and Demeny, P. (1966). "Regional model life-tables and stable populations. Princeton: Princeton University Press.

Coale, A. J. and McNeil, D. R. (1972). "The distribution by age of the frequency of first marriage in a female cohort". J. Amer. Stat. Assoc. 67, 743 - 749.

Coale, A. J. and Trussell, T. J. (1974). "Model fertility schedules: variations in the age structure of childbearing in human populations" Pop. Index 40, No. 2, 185 - 258.

Congdon, P. (1980). "Forecasting births in Greater London: an application of the Easterlin Hypothesis". Pop. Stud. 34, No. 2., 267 - 278.

Cox, D. R. (1972). "Regression models and life tables". J.R.S.S.(B), 34, 187 - 202.

Cramer, H., and Wold, H. (1935). "Mortality variations in Sweden: a study in graduation and forecasting". Skandinavisk Aktuarietidskrift, 11, 161 - 241.

Derrick, V. P. A. (1927). "Observations on errors of age in the population statistics of England and Wales and (2) the changes in mortality indicated by the national records". Journal of the Institute of Actuaries, 58, 117 - 146.

Duchêne, J., and Gillet-de Stefano, S. (1974). "Ajustment analytique des courbes de fécondité générale". Population et famille 32, No. 2, 53 - 93.

Easterlin, R. A. (1968). "Population labour force and long swings in economic growth". New York: National Bureau of Economic Research.

Easterlin, R. A. (1973). "Relative Economic Status and the American fertility swing". In "Family economic behaviour" (Eleanor B., Sheldon, ed.), J. B. Lippincott Company, Philadelphia, 170 - 223.

Easterlin, R. A. and Condran, G. A. (1976). " A note on the recent fertility swing in Australia, Canada, England and Wales and the United States", in H. Richards (ed), "Population, factor movements and economic development: studies presented to Brinley Thomas", Cardiff: University of Wales Press.

Ermisch, J. (1979). "The relevance of the 'Easterlin Hypothesis' and the 'New Home Economics' to fertility movements in Great Britain". Pop. Stud. 33, No. 1, 39 - 58.

Ermisch, J. (1981) "Economic opportunities, marriage squeezes and the propensity to marry: an economic analysis of period marriage rates in England and Wales". Pop. Stud. 35, No. 3., 347 - 356.

Ewbank, D. (1974). "An examination of several applications of the standard pattern of age at first marriage". Unpublished Ph.D dissertation, Princeton University, Princeton.

Farid, S. M. (1973). "On the pattern of cohort fertility". Pop. Stud. 27, No. 1, 159 - 168.

Farkas, G. (1977). "Cohort, age and period effects upon the employment of white females: evidence for 1957 - 1968". Demography 14, No. 1, 33 - 42.

Fienberg, S. E. and Mason, W. M. (1979). "Identification and estimation of age-period-cohort models in the analysis of discrete archival data". In Sociological Methology 1979. (K. F. Schuessler. ed.), 1 - 67, Jossey-Bass, San Francisco.

Finnäs, F., and Hoem, J. M. (1980). "Starting age and subsequent birth intervals in cohabitational unions in current Danish cohorts, 1975". Demography 17, No. 3, 275 - 295.

Freedman, R. and Casterline J. (1981), "Nuptiality and fertility in Taiwan". Proceedings of the I.U.S.S.P. conference on nuptiality and fertility, 61 - 100, L. T. Ruzicka (ed.), Ordina, Leige, Belgium.

Frost, W. D. (1939). "The age selection of mortality from tuberculosis in successive decades". American Journal of Hygiene 30, (3A), 91 - 96.

Gabriel, K. R. (1971). "The biplot graphic display of matrices with application to principal component analysis". Biometrika 58, No. 3., 453 - 467.

Gabriel, K. R., and Zamir, S. (1979). "Lower rank approximation of matrices by least squares with any choice of weights". Technometrics 21, No. 4., 489 - 498.

Garenne, M. (1980). "Use of correspondence analysis for the analysis of fertility determinants: application to four WFS countries". Paper prepared for the fourth meeting of the United Nations Working Group on Comparative Analysis of World Fertility Survey Data, Geneva, November, 1980.

Gilks, W. R. (1979). "An examination of a model of marital fertility depending on both age and duration of marriage". Unpublished M.Sc. thesis, University of Southampton.

Glass, D. V. (1976). "Recent and prospective trends in fertility in developed countries". Review lecture, Philosophical Transactions of the Royal Society of London (B), Biological Sciences, 274, 1 - 52.

Glenn, N. D. (1976). "Cohort analysts futile quest: statistical attempts to separate age, period and cohort effects". Amer. Soc. Rev. 41, 900 - 904.

Goldstein, H. (1978). Review of "Sociological Methodology" in Quantitative Sociology Newsletter, Autumn, 80 - 81.

Goldstein, H. (1979). "Age, period and cohort effects - a confounded confusion". Bulletin in Applied Statistics, 6, No. 1, 19 - 24.

Greenberg, B. G., Wright, J. J., Sheps, C. G. (1950). "A technique for analysing some factors affecting the incidence of syphilis". J.A.S.A. 45, 373 - 399.

Hadwiger, H. (1940). "Eine analytische reproduktionsfunktion für biologische gesamtheiten". Skandinavisk Aktuarietidskrift 23, 101 - 113.

Hastings, D. W., and Robinson, J. G. (1973). "A re-examination of Hernes' model on the process of entry into first marriage for United States women, cohorts 1891 - 1945". Amer. Soc. Rev. 38, 138 - 142.

Hayes, J. G., and Halliday, J. (1974). "The least squares fitting of cubic spline surfaces to general data sets". J. Inst. Math. Applic. 14, 89 - 103.

Henry, L. (1961). "Some data on Natural Fertility". Eugenics Quarterly, 8, no. 2, 81 - 91.

Hernes, G. (1972). "The process of entry into first marriage". Amer. Soc. Rev. 37, 173 - 182.

Hobcraft, J. N. (1979). "Model life-tables for developing countries: an interim report". A report prepared on behalf of the Population Division of the Department of Economic and Social Affairs of the United Nations Secretariat.

Hobcraft, J. N. and Gilks, W. R. (1981). "Age, period and cohort analysis in mortality studies". I.U.S.S.P. meeting in Dakar, Senegal (in press).

Hobcraft, J. N., Menken, J. and Preston, S. (1979). "Age, period and cohort effects in demography: a review", in H.H. Winsborough and O. D. Duncan (eds.) "Analyzing longitudinal data for age, period and cohort effects", Academic Press, New York: 1982, and in Population Index 48, No.1, (1982), 4 - 43.

Hobcraft, J. N. and Rodriguez, G. (1980). "Methodological issues in life-table analysis of birth-histories". Paper presented at a seminar on the analysis of maternity histories, London, 9 - 11 April, 1980, I.U.S.S.P. (C.C.A.F.).

Hoem, J. M., Madsen, D., Nielsen, J. L., Ohlsen, E. -M., Hansen, H. O., and Rennermalm, B. (1981) "Experiments in modelling recent Danish fertility curves". Demography 18, No. 2, 231 - 243.

Hogan, H. R. and McNeil, D. R. (1979). "On fitting statistical relations to age-specific mortality data". Paper presented to the annual conference of the Population Association of America, Philadelphia, 1979.

Isaac, L., Cutrights, P., Jackson, E., and Kelly, W. (1979). "Period effects on race-and parity - specific birth probabilities of American women 1917 - 76; a new measure of fertility". Unpublished Paper, University of Texas, Austin.

Jain, A. K., and Bongaarts, J. (1980). "Socio-biological factors in exposure to child-bearing: breastfeeding and its fertility effects". World Fertility Survey Conference, July, 1980, London.

Johnson, N. (1949). "Systems of frequency curves generated by methods of translation". Biometrika 36, 149 - 176.

Jones, E. F. (1982). "Ways in which childbearing affects women's employment: evidence from the U.S. 1975 National Fertility Study". Pop. Stud. 36, No. 1, 5 - 14.

Jones, R. E. (1980). "Further evidence on the decline in infant mortality in pre-industrial England: North Shropshire 1561 - 1810". Pop. Stud. 34, No. 2, 239 - 250.

Kendall, M. G. (1979). "The World Fertility Survey: current status and findings". Population Reports, Series M, No. 3.

Kendall, M. G. and Stuart, A. (1951). "The advanced theory of statistics". 2: "Inference and relationship". Third edition, Griffin, London.

Kermack, W. O., McKendrick, A. G., and McKinlay, P. L. (1934). "Death-rates in Great Britain and Sweden: expression of specific mortality rates as products of two factors, and some consequences thereof". Journal of Hygiene 34, 433 - 457.

Kiernan, K. E., and Diamond, I. (1982). "Family-of-origin and educational influences on age at first birth: the experiences of a British birth cohort". CPS Research Paper No. 82, Centre of Population Studies, London School of Hygiene and Tropical Medicine.

Le Bras, H. (1979) (translated by J. Hobcraft). "Model life-tables". Unpublished manuscript, United Nations.

Ledermann, S., and Breas, J. (1959). "Les dimension de la mortalité". Population 14, 637 - 682.

Lee, R. D. (1974). "The formal dynamics of controlled population and the Echo, the Boom and the Bust". Demography 11, 563 - 585.

Lee, R. D. (1976). "Demographic forecasting and the Easterlin Hypothesis". Population and Development Review 2, 459 - 468.

Lee, R. D. (1977). "Target fertility, contraception and aggregate rates: toward a formal synthesis". Demography 14, No. 4, 455 - 479.

Lee, R. D. (1978). "Causes and consequences of age structure fluctuations: the Easterlin Hypothesis". Conference on Economic and Demographic Change, Issues for the 1980s. I.U.S.S.P. Helsinki.

Lee, R. D. (1980). "Aiming at a moving target: period fertility and changing reproductive goals". Pop. Stud. 34, No. 2, 205 - 226

Marini, M. M. (1981) "Effects of the timing of marriage and first birth on fertility". Journal of Marriage and the Family 43, 27 - 45.

Mason, K. O., Mason, W. M., Winsborough, H. H. and Poole, W. K. (1973). "Some methodological issues in cohort analysis of archival data". Amer. Soc. Rev. 38, 242 - 258.

Mason, K. O., Mason, W. M., Winsborough, H. H. and Poole, W. K. (1976). "Reply to Glenn". Amer. Soc. Rev. 41, 905.

Mason, W. M. and Smith, H. L. (1979). "Age, period, cohort analysis and the study of deaths from pulmonary tuberculosis", in H. H. Winsborough and O. D. Duncan (eds.), "Analyzing longitudinal data for age, period and cohort effects", Academic Press, New York, 1982.

Mazur, D.P.(1963). "A demographic model for estimating age-order specific fertility rates". J.A.S.A, 58, 774 - 788.

McDonald, P., Ruzicka, L. T. and Caldwell, J. C. (1980). "Interrelations between nuptiality and fertility: evidence from the World Fertility Survey". World Fertility Survey Conference, July, 1980, London.

McNeil, D. R., Trussell, T. J. and Turner, J.C. (1977) "Spline interpolation of demographic data". Demography 14, No. 2., 245 - 252.

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.

Menken, J., Trussell, J., Stempel, D., Babakol, O. (1981). "Proportional hazards of life table models: an illustrative analysis of socio-demographic influences on marriage dissolution in the United States". Demography 18, No. 2, 181 - 200.

Miller, M. D. (1949). "Elements of graduation". Chicago: American Institute of Actuaries.

Mincer, J. (1963). "Market prices, opportunity costs, and income effects". in C. Christ et al (eds.), "Measurement in economics: studies in mathematical economics in memory of Yeyuda Grunfeld", Palo Alto, Cal.: Stanford University Press.

Mitra, S. (1967). "The pattern of age-specific fertility rates". Demography 4, 894 - 906.

Mitra, S. and Romanuik, A. (1973). "Pearsonian type I curve and its fertility projection potentials". Demography 10, No. 3, 351 - 365.

Murphy, E., and Nagnur, D. (1972). " A Gompertz that fits: applications to Canadian fertility patterns". Demography 9, No. 1, 35 - 50.

Murphy, M. (1982). "Gompertz and Gompertz relational models for forecasting fertility: an empirical exploration". CPS Research Paper No. 82, Centre for Population Studies, London School of Hygiene and Tropical Medecine.

O.P.C.S. (1977a). Unpublished table of first-marriage and the single population, by single years of age, for each sex, since the 1900 birth cohorts. Office of Population Censuses and Surveys.

O.P.C.S. (1977b). "Legitimate live-birth rates (per 1000 women) at successive durations of marriage, for women married once only; age at marriage and calendar years of marriage, 1920 - 75". Table 10.3, Birth Statistics volume, Series FM1, No. 3, Office of Population, Censuses and Surveys.

O.P.C.S. (1979). Table of fertility rates for all women in England and Wales for single years of age between 15 and 49, and single-year time-periods from 1938 to 1979. Derived from an unpublished table supplied by the Office of Population, Censuses and Surveys London.

Osmond, C. and Gardner, M. J. (1982). "Age, period and cohort models applied to cancer mortality rates". Statistics in medicine 1, 245 - 259.

Page, H. J. (1976). "Patterns underlying fertility schedules: a decomposition by both age and marriage duration." Populations Studies 30, No. 3, 85 - 106.

Pittenger, D. B. (1980). "Some models of age-specific fertility rates". Environment and Planning A, 12, 1145 - 1164.

Preston, S. H. and McDonald, J. (1979). "The incidence of divorce within cohorts of American marriages contracted since the Civil War". Demography, 16, No. 1, 1 - 26.

Pullum, T. W. (1978). "Parameterising age, period and cohort effects: an application to U.S. delinquency rates, 1964 - 73". In "Sociological methodology 1978". (K. F. Schuessler, ed.), 116 - 140. Jossey-Bass, San Francisco.

Pullum, T. W. (1980). "Separating age, period and cohort effects in white U.S. fertility, 1920 - 70". Social Science Research 9, 225 - 244.

Rindfuss, R. R. and Bumpass, L. L. (1978). "Age and the sociology of fertility: how old is too old?" In K. E. Taeuber, L. L. Bumpass and J. A. Sweet (eds.). "Social Demography" Academic Press, New York.

Rodriguez, G. and Hobcraft, J. (1980). "Illustrative analysis: life table analysis of birth intervals in Columbia". W.F.S. Scientific Reports, No. 16.

Romaniuk, A.(1973). "A three parameter model for birth projections". Pop. Stud. 27, 467 - 478.

Ruzicka, L. T. (1976). "Age at marriage and timing of first birth". Population Studies 30, No. 3, 527 - 538.

Ryder, N. B. (1978), "A model of fertility by planning status". Demography 15, No. 4., 433 - 458.

Ryder, N. B. (1980). "Components of temporal variations in American fertility", in R. W. Hiorns (ed.), "Demographic patterns in developed societies", 11 - 54. London: Taylor and Francis Ltd.

Sacher, G. A. (1957). Discussion in "Mammalian aspects of basic mechanisms of radiobiology", 121 - 125 and 137 - 142. NAS-NRC. Publications No. 513, Report No. 21.

Sacher, G. A. (1960). "Analysis of life tables with secular terms". In the "Biology of aging", Amer. Inst. Biol. Sci. Symposium 6, 253.

Sacher, G. A. (1977). "Life table modification and life prolongation", in Finch, C. E. and Hayflick, L. (eds), "Handbook of the biology of aging", 532 - 638, New York: Van Nostrand Reinbold.

Samuelson, P. A. (1976). "An economist's non-linear model of self-generated fertility waves". Pop. Stud. 30, 243 - 247.

Sanderson, W. C. (1979). "Towards an economic analysis of the baby boom: a test of the Easterlin Hypothesis." Unpublished paper, Standford University.

Smith, D. M. and Keyfitz, M. (1977). "Mathematical demography: selected papers". Biomathematics 6, Berlin-Heidelberg: Springer-Verlag.

Springett, V. H. (1950. "A comparitive study of tuberculosis mortality rates". Journal of Hygiene, 48, No. 3, 361 - 395.

Stoto, M. and Menken, J. (1977). "Estimation of a parametric model for the distribution of birth intervals based on censored data". Princeton University, Office of Population Research, preliminary report.

Talwar, P. P. (1974). "Model fertility patterns for population projections". Presented to the Annual Meeting of the Population Association of America, New York.

Thurston, R. (1970). "An analysis of recent Swedish nuptiality". Unpublished paper, Office of Population Research, Princeton University.

Townsend, J. L. (1978). "Smoking and lung cancer: a cohort data study of men and women in England and Wales, 1935 - 70". J.R.S.S.(A), 141, No. 1, 95 - 107.

United Nations (1965). "Legitimate live-birth rates specific for age of mother: latest available year". United Nations Demographic Yearbook for 1965, table 25, p. 568 - 573.

United Nations (1975). "Legitimate live-birth rates specific for age of mother: latest available year." United Nations Demographic Yearbook for 1975, table 34, p. 776 - 778.

Vaupel, J. W., Manton, K. G., and Stollard, E. (1979). "The impact of heterogeneity in individual frailty on the dynamics of mortality". Demography 16, No. 3, 439 - 54.

Wilkie, J. R. (1981). "The trend toward delayed parenthood". Journal of Marriage and the Family, 43, 583 - 591.

Wolfenden, H. H. (1954). "Population Statistics and their compilation". First published 1925. Chicago: University of Chicago Press.

Wunsch, G. (1966). "Courbes des Gompertz et perspectives de fécondité". Recherches Economiques de Louvain 32, 457 - 468.

Wunsch, G. (1979). "Effect of changes in nuptiality on mortality in Western Europe". Paper presented at the IUSSP seminar on nuptiality and fertility, January, 1979, Bruges.

Yule, G. A. (1906). "On the changes in the marriage and birth rates in England and Wales during the past half century". J.R.S.S. 69, 88 - 147.

Zaba, B. (1979). "The four parameter logit life-table system". Pop. Stud. 33, No. 1. 79 - 100.